

Pricing under Random Information Flow and the Theory of Information Pricing

A thesis presented for the degree of
Doctor of Philosophy of Imperial College London
by

Yan Tai LAW

Department of Mathematics
Imperial College London
180 Queen's Gate, London SW7 2BZ

OCTOBER 2011

Declaration

I certify that this thesis, and the research to which it refers, are the product of my own work, and that any ideas or quotations from the work of other people, published or otherwise, are fully acknowledged in accordance with the standard referencing practices of the discipline.

Signed:

Copyright

Copyright in text of this thesis rests with the Author. Copies (by any process) either in full, or of extracts, may be made **only** in accordance with instructions given by the Author and lodged in the doctorate thesis archive of the college central library. Details may be obtained from the Librarian. This page must form part of any such copies made. Further copies (by any process) of copies made in accordance with such instructions may not be made without the permission (in writing) of the Author.

The ownership of any intellectual property rights which may be described in this thesis is vested in Imperial College, subject to any prior agreement to the contrary, and may not be made available for use by third parties without the written permission of the University, which will prescribe the terms and conditions of any such agreement. Further information on the conditions under which disclosures and exploitation may take place is available from the Imperial College registry.

Acknowledgments

First and foremost, I would like to thank my supervisor, Dorje Brody, who introduced me to this interesting field of research and gave me the opportunity to be his student. I am indebted to the endless amount of effort put into me; he taught me so much and has always been more than a mentor to me. I have been inspired and learned tremendously from this work, both in financial mathematics and thoughts.

I would like to express my gratitude to all the people who supported me to complete this work over this period, with particular thanks to M. Davis, L. Hughston, S. Miao, K. O'Brien, T. Allen, P. Karasinski, M. Soloveitchik, G. Butti, S. Kiselev, R. Coleman and G. Salkin. Also to all my colleagues at Imperial College London who have made this a great atmosphere to work in.

I thank HSBC for support in the early stages of my work, and for giving me the environment to learn many valuable things outside academia.

I acknowledge the Department of Mathematics, EPSRC, the Institute of Mathematical Sciences, an Imperial College Research Impulse Grant, and HSBC for financial support.

Last but not least, this has been a long journey full of ups and downs, which could not have been possible without the constant support and encouragement of my family and my wife Chloe.

Abstract

This thesis presents a mathematical formulation of informational inhomogeneity in financial markets, with emphasis on its impact on asset volatility, the notion of information extraction, and the role of information providers. We begin with a brief review of the BHM framework, which models the market filtration by an information process consisting of a signal and a noise term, such that the signal-to-noise ratio is determined by the information flow rate. Motivated by the observations that valuable information is rarely circulated homogeneously across financial markets, and that the information flow rate is typically random, we introduce, in the first part of the thesis, an extension of the BHM approach that leads to the simplest class of stochastic volatility models. In this extended framework we derive closed form expressions: for (a) asset price processes; (b) pricing formulae for options; and (c) option deltas. We show that the model can be calibrated to fit volatility surfaces reasonably well, and that it can be used effectively to model information manipulation. In the second part we introduce a framework for the valuation of information. In particular, a new formulation of the utility-indifference argument is introduced and used as a basis for pricing. We regard information as a quantity that converts *a priori* distributions into *a posteriori* distributions. The amount of information can thus be quantified by relative entropy. The key to our theory is to equate the maximised *a posteriori* utility with the *a posteriori* expectation of the utility of the *a priori* optimal strategy. This formulation leads to one price for a given quantity of upside, and another for a given quantity of downside information. Various intuitive, as well as counterintuitive implications (for example, price of information is not necessarily an increasing function of the volume of information) of our theory are discussed in detail.

Table of contents

Acknowledgments	4
Abstract	5
List of figures	11
Introduction	12
1 Information-Based Asset Pricing	18
1.1 Market filtration	18
1.2 Asset information and price formation	20
1.3 Price dynamics	21
1.4 Securities and derivative pricing	25
2 Stochastic Information Revealing Rate	27
2.1 Motivation	27
2.2 Setup	28
2.3 Meaning of the information process	29
2.4 Sensitivity analysis	30
2.5 Quantisation of information flow rate	32
2.6 Price process with a random information flow rate	33
2.6.1 Markovian property of ξ_t	34
2.6.2 Conditional probability	35
2.7 Credit-risky bond dynamics	36
2.8 Volatility dynamics	40

2.9	Information-flow rate process	42
2.10	Analysis of path behaviour	42
2.11	Mutual information analysis	46
2.11.1	Limiting property of mutual information	47
3	Options on Credit-risky Bonds with Random Information Flows	50
3.1	Option pricing formula	50
3.2	Measure change technique	51
3.3	Sensitivity analysis	55
3.4	Stochastic volatility	56
3.5	Remarks on the random information and the BHM mixture	57
3.6	Information manipulation	58
4	Theory of Information Pricing	61
4.1	Introduction	61
4.2	Heuristic illustration	63
4.2.1	Optimal portfolio choice for an uninformed investor	64
4.2.2	Optimal portfolio choice for informed investor	66
4.3	Utility indifference pricing of information	67
4.4	Information dependent price	68
4.5	Information independent price	70
4.6	Separation of density functions	71
4.7	Utility indifference information pricing with continuous payoff	72
4.7.1	Unconstrained scenario	76
4.7.2	Constrained scenario	77
4.8	Analysis of expected loss at maximum information cost	80
4.9	Cost of information in continuous time	81
4.9.1	Utility optimisation	81
4.9.2	Utility optimisation with information	84
4.10	Cost for extra noisy information	86
4.11	Discussion	88

	8
Conclusion	90
Appendices	91
A Risk Neutral Measure for the Cost of Information	91
A.1 Presence of arbitrage in the Grossman-Stiglitz formula	91
A.2 Finiteness of the terminal cost in discrete setting	92
A.3 Gaussian case	97
B Filtration Enlargement	100
B.1 Martingale preserving probability measure	101
B.2 Strong predictable representation property	104
C Utility Optimisation in Incomplete Market via Convex Duality	105
C.1 No arbitrage condition	105
C.2 Dual formulation	106
C.3 Exponential utility optimisation	108
References	117

List of Figures

- 2.1 The unconditional expectation $\mathbb{E}[g_t(\sigma)]$ of the Fisher information for different values of t and σ . The horizontal axis represents the time remaining to the maturity. The vertical axis represents the values of σ , ranging from 0 to 2. The numerical result is obtained by averaging over 1,000 sample paths. The parameters are chosen such that $\{x_i\} = \{0, 0.5, 1\}$, $\{p_i\} = \{0.1, 0.15, 0.75\}$, and $T = 1$. Where the value of $\mathbb{E}[g_t(\sigma)]$ is high, the model is on average sensitive to the choice of the specification of the parameter σ 32
- 2.2 Simulated bond price (left), volatility (top right) and Vol-of-Vol (bottom right) processes in uncertain information model with $\bar{\sigma} = \{0.6, 0.8\}$, $q_k = \{0.5, 0.5\}$, $X_T = \{0, 1\}$, $p_i = \{0.2, 0.8\}$, $r = 0\%$ and $T = 1$ 43
- 2.3 Simulated bond price (left), volatility (top right) and Vol-of-Vol (bottom right) processes in uncertain information model with $\sigma = \{0.4, 1\}$, $q_k = \{0.5, 0.5\}$, $X_T = \{0, 1\}$, $p_i = \{0.2, 0.8\}$, $r = 0\%$ and $T = 1$ 44
- 2.4 Each line represents the average of 5,000 sample paths of volatility processes. The two solid lines are that of the original BHM model with the top and bottom line taking $\sigma = 0.9$ and $\sigma = 0.5$ respectively, $X_T = \{0, 1\}$, $p_i = \{0.2, 0.8\}$, $r = 0\%$ and $T = 5$. The dashed lines are mean volatilities of the uncertain information model taking values, from top to bottom respectively, $\sigma_k = \{0.5, 0.9\}$ with $q_k = \{0, 1\}$, $q_k = \{0.1, 0.9\}$, $q_k = \{0.2, 0.8\}$, ..., and $q_k = \{1, 0\}$ 45
- 2.5 Magnification of the averages of volatility processes. Each line represents the average of 5,000 paths of volatility processes. The solid lines are that of the original BHM model with the top and bottom line taking $\sigma = 0.9$ and $\sigma = 0.5$, respectively, $X_T = \{0, 1\}$, $p_i = \{0.2, 0.8\}$, $r = 0\%$ and $T = 5$. The dashed lines are mean volatilities of the uncertain information model taking the values, from top to bottom respectively, $\sigma_k = \{0.5, 0.9\}$ with $q_k = \{0, 1\}$, $q_k = \{0.1, 0.9\}$, $q_k = \{0.2, 0.8\}$, ..., and $q_k = \{1, 0\}$ 46

- 2.6 Mutual information under the random σ model. The vertical axis represents the mutual information and the horizontal axis represents the time elapsed. We let $\sigma_k = \{0.5, 0.9\}$ with probabilities, from top to bottom at early times, $q_k = \{0, 1\}$, $q_k = \{0.1, 0.9\}, \dots, q_k = \{1, 0\}$ respectively; $X_T = \{0, 1\}$, $p_i = \{0.2, 0.8\}$, and $T = 5$ 49
- 3.1 Call price against strike for different option maturities under the BHM (left) and uncertain information (right) models. We see that options with longer maturities are less expensive in the random information model, for any strike price. This is because the additional uncertainty in σ indeed reduces volatility in the initial stage of a long maturity option. As the options approach maturity the two models become the same. 55
- 3.2 *The implied BHM volatility.* The implied volatility surface resulting from calibrating each strike and maturity of the random information model to the BHM model is shown. The parameters are set to be $X_T = \{0, 1\}$, $p_i = \{0.2, 0.8\}$, $\sigma_{\text{average}} = 1.5$, $\sigma_{\text{Rand}} = \{0.3, 2.7\}$, $q_k = \{0.5, 0.5\}$, $r = 0\%$ and $T \in [0, 2]$ 57
- 3.3 Left: Sample paths of the BHM, i.e. bond price against time elapsed, with $\sigma = 1$ (blue) and $\sigma = -1$ (red) using same Brownian bridges; and the corresponding skew processes associated with $X_T(\omega) = 1$ (Top right) and $X_T(\omega) = 0$ (Bottom right). Here we set $X_T = \{0, 1\}$, $p_i = \{0.2, 0.8\}$, $r = 0\%$ and $T = 5$ 59
- 4.1 Left: Expected utilities $\mathbb{E}_\pi[U(W_p^*)]$ for $\xi = 0, 1, 2$ as a function of the cost c are shown here as three horizontal lines. For each value of ξ we also plot $\mathbb{E}_\pi[U(W_{\pi,c}^*)]$, when short selling is strictly prohibited. When c is sufficiently small, the latter utilities are higher, but as the cost increases, they decrease. At $c = c^*$ the two utility functions for each ξ agree and determine the unique equilibrium price for which the investor is indifferent. Here we have chosen parameters to be $\alpha = 0.1$, $s_0 = 0.5$, $W_0 = 5$, $p = 0.4$, $q = 0.7$, and $\delta = 0.95$. The numerical values for the cost for each ξ are plotted as a histogram on the right side. In this example, we have $c^*(0) = 1.51$, $c^*(1) = 0.71$, and $c^*(2) = 1.65$. The average cost is $\bar{c}^* = 1.10$ 69
- 4.2 The Bhattacharyya spherical distance between p and π for difference values of ξ (the left top curve corresponds to $\xi = 0$, left middle to $\xi = 2$, and left bottom to $\xi = 1$). The horizontal axis represents different values of p 72
- 4.3 Left: Cost $c^*(\xi)$ of information as a function of the realised value of ξ . Right: Relative entropy $I(\xi) = I(\pi_\xi|p)$ as a function of the realised value of ξ . The parameter values used here are: $s_0 = 1$, $\delta = 0.95$, $\alpha = 0.1$, $W_0 = 10$, $\mu = 1.1$, $\sigma^2 = 0.2$, $\mu_\epsilon = 0$, and $\sigma_\epsilon^2 = 0.2$. The red curves correspond to pieces of downside information (i.e., values of ξ that indicate decrease in s_1 as compared to the *a priori* expectation) and the blue curves correspond to upside information. 74

- 4.4 The cost $c^*(I)$ as a function of the quantity of information I , when there is no constraint for short selling. The solid lines in all four plots (they are identical) correspond to the parameter values used in figure 4.3. While keeping other parameters fixed, the following changes in parameter values are used, represented by dashed lines in the plots. Top left: α has been increased to $\alpha = 0.2$; top right: σ^2 has been increased to $\sigma^2 = 0.4$; bottom left: σ_ϵ^2 has been increased to $\sigma_\epsilon^2 = 0.4$; and bottom right: δ has been decreased to $\delta = 0.92$. The red curves correspond to downside information, and the blue curves correspond to upside information. 75
- 4.5 Left: Cost $c^*(\xi)$ of information as a function of the realised value of ξ when short selling is strictly prohibited. Right: Relative entropy $I(\xi) = I(\pi_\xi|p)$ as a function of the realised value of ξ . The cost $c^*(\xi)$ measures the usefulness of information, while relative entropy $I(\xi)$ measures the quantity of information. The parameter values used here are: $s_0 = 1$, $\delta = 0.95$, $\alpha = 0.1$, $W_0 = 10$, $\mu = 1.1$, $\sigma^2 = 0.2$, $\mu_\epsilon = 0$, and $\sigma_\epsilon^2 = 0.2$. The red curves correspond to pieces of downside information (i.e. values of ξ that indicate decrease in s_1 as compared to the *a priori* expectation) and the blue curves correspond to upside information. 76
- 4.6 The cost $c^*(I)$ as a function of the quantity of information I , when short selling is strictly prohibited. The solid lines in all four plots (they are identical) correspond to the parameter values used in figure 4.5. While keeping other parameters fixed, the following changes in parameter values are used, represented by dashed lines in the plots. Top left: α has been increased to $\alpha = 0.2$; top right: σ^2 has been increased to $\sigma^2 = 0.4$; bottom left: σ_ϵ^2 has been increased to $\sigma_\epsilon^2 = 0.4$; and bottom right: δ has been decreased to $\delta = 0.92$. The red curves correspond to downside information, and the blue curves correspond to upside information. 78
- 4.7 Interpolation of the information price $c^*(I)$ from the strictly constrained case to the unconstrained case. Top row, from left: (i) borrowing is strictly forbidden; (ii) borrowing is capped at $W_0/2$; (iii) borrowing is capped at W_0 . Bottom row, from left: (i) borrowing is capped at $2W_0$; (ii) borrowing is capped at $5W_0$; (iii) unlimited borrowing permitted. 79

Introduction

Asset pricing with random information flow

Various aspects of the role of information in asset pricing have been studied extensively in the literature (e.g., asymmetric information, filtration enlargement) for at least half a century [1, 27, 28, 31, 50, 52]. The value of an asset is based on the expectation of its future returns—whether investing in the asset will bring the desired returns. Such an expectation is determined by the available information gathered up to the time at which the valuation is made. However, it is rarely the case one would actually be able to obtain the full information that is in principle available in the market: it is more usually the case that different investors receive different amounts of information, and this in turn contributes in a significant way to market volatility. Indeed, it is reasonable to take the view that the problems associated with modelling (a) information inhomogeneity, and (b) volatility structure are very closely related. To understand the volatility structure in the market, we need to take a step back from traditional asset pricing or volatility modelling, and take as a fresh starting point the modelling of information.

In classical asset pricing models, the normal point of embarkation is usually the specification of a model for the price process of an asset, which is typically defined on a prescribed probability space equipped with the augmented filtration generated by a (multi-dimensional) Brownian motion to which the price process is taken to be adapted. The consideration of such a Brownian filtration sometimes give rise to the question whether the Brownian motion really contains any useful information—is it not just noise, after all? Indeed, we often speak of an asset price as being “driven” by noise. So it seems paradoxical that asset price movements can be “informative” on the one hand, and yet merely driven by “noise” on the other. In light of this, it would be more intuitive to begin the modelling of asset price processes and associated derivative prices by explicitly modelling the filtration—that is to say, by the modelling of information.

One promising way forward to this end has been provided in the framework developed by Brody, Hughston and Macrina [17] (hereafter denoted BHM) in several works and in collaborations with various co-authors. The idea of the BHM framework is that the price process of an asset is viewed as the output of a series of factors brought into play by the market, rather than an input. The filtration available to a financial market is an important ingredient, indeed a key ingredient, for price formation—a point that tends to be neglected in mainstream asset pricing. Movements of prices are, to a large extent, due to changes in either (i) the “pricing kernel”, or (ii) the information available to the market. The former means in essence that the representative investor has changed his attitude to risk and his level of impatience, and hence is adjusting his position to accommodate these changes in “preferences”. The second source of price movement is due to information update in the market: when new information becomes available, price movements reflect the changes in investor expectations of the returns or cash flows likely to be offered by the securities under consideration. Anderson *et al.* [4] used high-frequency data to show that news is one of the driving factors of price dynamics, and makes a contribution to conditional mean jumps. More recently, Bollen *et al.* [8] reported that the sentiment analysis of Twitter texts can be used to predict future movements of stock indices. Independently, Brody *et al.* [12] demonstrated a similar predictability feature of the financial market using internet blog data. All these researches provide evidence by empirical research that there exists an information inhomogeneity in the financial market.

The BHM framework adopts mathematical ideas in a fundamentally novel way from the general theory of the filtering of random systems (see, e.g., Kallianpur [36] and Wonham [56]). According to the BHM methodology, the price process of an asset that delivers a single cash flow at some specified terminal date is driven by the dynamics of the conditional (*a posteriori*) probability of the terminal cash flow, given the information in the filtration up to that time. The price process of an asset is derived with the specification of an information process that generates the filtration of the probability space. The information process aims to take into account the two sources affecting price movements: the signal concerning the asset return, and the associated noise that obscures the signal. The price is then given by the discounted risk-adjusted expectation of the cash flow, conditional on the available information.

In the BHM model, one needs to specify the information flow rate parameter that describes the rate of revelation of information across the market. However, measuring the information flow in the market is somehow a very difficult task—in part because information can be rather intangible—and there also exists the problem of information inhomogeneity, where part of the market may not be able to absorb certain news as quickly as some other parts of

the market are able to. These problems lead eventually to the suggestion that a very natural improvement or extension of the BHM framework would be to admit within the modeling framework the possibility of a random information flow rate.

We shall in what follows discuss in detail the analysis of such random information flow rate models, and derive the relevant conditional probability necessary for asset pricing. The dynamics of the bond price process thus obtained will be shown to be consistent with the BHM model in the sense that the random information flow model reduces to the BHM if the distribution of the information flow random variable collapses to a point. An information-theoretic analysis will be presented in some detail to highlight the learning rate of information under the auspices of such a model. The model for the risky asset, resulting from the use of random information flow rate, can in fact be viewed as a BHM-mixture model, in a way analogous to the log-normal mixture models sometimes used as extensions of the Black-Scholes model. Here, however, the mixture arises as a consequence; not as a model postulate. We are thus able to identify an information-theoretic justification for such mixture models that are otherwise introduced in the literature often on an *ad hoc* basis [9, 43].

Perhaps one of the most interesting phenomena arising from the random information flow model concerns the volatility process. With an uncertain information flow rate, the bond volatility process begins lower and ends higher, when compared to the corresponding situation in the original BHM model. Economically, this can be explained by the suggestion that since information flows are uncertain the result is a reduction in the willingness of market participants to engage in market activity. As more information is revealed towards maturity, the bond volatility is maintained because there is still activity due to uncertain but nontrivial potential profitable opportunities.

The information flow parameter in this sense controls the volatility of the bond process. A major consequence of the random information flow model is that it allows for the possibility of modelling stochastic volatility in option pricing. We shall work out the pricing of an option on a credit-risky bond, under the random information model, where the solution is no more complicated computationally than the Black-Scholes option pricing formula. We shall show by calibrating the BHM model to the random information flow model that we can produce a volatility surface as observed in option markets. This approach allows for the pricing, from an intuitive setting in an information-driven context, of derivatives that are explicitly sensitive to future volatility levels.

The method proposed here for modelling information flows to derive the stochasticity of volatility offers possible solutions to many of the existing problems in traditional stochastic volatility modelling as in Heston [32], Hull *et al.* [33], Stein *et al.* [51] as well as many

others. Jäckel [34] has provided a concise summary of some such models. The problem of these models is that the structure of the volatility is typically modelled in a rather *ad hoc* manner, and is somehow not very obvious to justify. For example, in order to calibrate the Heston model, one would need the mean reverting rate and the volatility of the volatility (vol-of-vol) to be extremely low. The set of calibrated parameters are often hard to interpret economically, and perhaps this is why the choice of a model is often largely dependent on the ease with which it can be calibrated to market data, rather than the degree to which it fits the structure of the contract.

Finally, the investigation of the random information flow model leads to the concept of “price manipulation” by an act carried out by some market agent (e.g., an institution) capable of manipulating the information flow rate. This idea of information disparity is further quantified in the section of the thesis on information pricing.

Information cost and measurement

Under an uncertain information flow one can nevertheless find fair prices for securities. However, due to these uncertainties regarding the final outcome, the conditional volatility process becomes very different and there exists the possibility of information manipulation. Disparity of information alone can affect the expectation of the assets, and hence the present value. Therefore, there will be a general desire among market participants to acquire more information, even at a cost. If one were able to obtain extra information to obtain a better expectation of the future return of an asset, then it seems reasonable (perhaps through the adaptation of some generalised principle of the absence of arbitrage) that the required piece of information should be fairly priced. In a sense, information can be regarded as a “perishable” asset that can be bought or sold while still “fresh”.

The fact that information can be regarded as tradable object has been envisaged by Gleick [25], in which the basic unit of information, the bit, is used to represent atoms for constructing an amount of information. In our view, information acts like a kind of commodity that can be consumed or traded, but will perish rather quickly, like bread or fish, and must be used before its “best-by date”, or discarded.

The first question we may ask is: what is the cost for a piece of information? This turns out to be a nontrivial problem. Indeed, before one can meaningfully ask this question, the definition of information needs to be understood in the first place. For any piece of information, it could at least in principle have almost any possible effect on the stock price:

it could be inside information, the revelation of which would alter the stock price; or it could be based on some comprehensive historical analysis of the market, which has no direct connection to any changes to the company under consideration; or it could be pure speculation or “gossip” made up by essentially random sources.

The problem of distinguishing reliable knowledge from rumour—or equivalently to distinguish information from noise—has been a subject of study for many years. In a paper by Hayek [31] the author discusses the importance to society of knowledge about specific tasks, and that there should exist a price system to filter out the irrelevant parts, such that the right skill set is accordingly valued more highly. However, unless we know the “bottom-line” result, it is difficult to tell which piece of information is actually valuable. Sometimes only time can tell the truth.

As in the BHM framework, we consider a piece of information to consist of two parts, namely, (i) a signal term and (ii) a noise term. As will be shown in chapter 2, the signal-to-noise ratio plays a significant role in determining the quality of the information. This parameter describes the information flow in the market: nonetheless, it also drives the volatility of the price process. It is shown in chapter 2 that if one uses a manipulated signal-to-noise ratio, the derived price process can move in a direction completely opposite to the process with a true signal-to-noise ratio. In this case we consider the information to be unreliable. For simplicity, we assume that the information on hand is reliable.

Therefore, all information is useful but with different degrees of usefulness. This is consistent with the idea of the price system in Hayek [31]. The notion of usefulness crucially depends on whether it changes one’s prior belief, and to what extent one can leverage on it. For the former, an investor is willing to pay for a piece of information that is not already held by the general public. It is the change of a stock price in adjusting to this information that creates profit opportunities. Such “surprises” can be quantified using information measures such as the relative entropy. A cost function can then be associated to each new bit of information, for downside and upside. This is shown in chapter 4.

We consider information as an asset that is nonreplicable. Hence for its pricing we propose to adopt the method of utility indifference. The idea is that the informational cost should balance out the profit gained for having this extra information, otherwise the market will fall into disequilibrium. We show that using the utility indifference pricing formula proposed by Grossman and Stiglitz [28] can lead to a form of arbitrage in the sense that with a modest cost for the information an investor can gain, in extreme cases, unbounded upside return. The key shortcoming of the celebrated work of Grossman and Stiglitz [28] is that the cost for information is independent of the quality of information. We introduce an aug-

mented utility indifference price such that it depends on the quality, i.e. opportunities it may provide, of information being released, and one that is free of arbitrage in the sense that a piece of information that offers unbounded benefit, according to our theory, costs infinity. In fact, this statement is made more precise in Appendix A to show the existence of a martingale family. It is also worth noting that our utility indifference information prices have different characteristics for upside and downside information, and we shall demonstrate the sensitivity of the prices to changes with respect to various model parameters.

The layout of the thesis will be as follows. In chapter 1 we review the basic background material of the information-based asset pricing framework that is presented in Brody *et al.* [15, 17]. In chapter 2, we introduce the notion of a random information flow rate, and discuss the necessary conditions for this modelling setup to be economically viable. We derive the bond pricing formula and its dynamics, drawing attention to the volatility process and the information-learning rate. In chapter 3, the option pricing formula and the associated option delta that determines the sensitivity are worked out in detail. The use of the random information flow can be viewed as the simplest stochastic-volatility extension of the BHM model, leading to a better-fitted implied volatility surface. Unlike conventional stochastic volatility models, where the stochasticity is introduced typically in an *ad hoc* manner, we are able to justify our approach from empirical facts concerning the lack of knowledge of the flow rate parameter.

In the second half of the thesis, we present the main argument for the theory of information pricing, and demonstrate how the theory works by means of simple examples. Specifically, in Section 4.4 we propose a new pricing formula for information, such that the price is dependent on the quality of information being offered. Of course, from the viewpoint of information provider, it is not possible to guarantee the quality of information; our pricing formula thus can be understood as providing a pre-agreement between information provider and information receiver. Various intuitive, as well as counterintuitive observations (for example, price of information is not necessarily an increasing function of the volume of information) that follow from our theory are discussed in detail in chapter 4 by means of simple examples. In particular, we consider the behaviour of the cost structure associated with various different degrees of market constraints regarding the possibility of short selling (for example, if short selling is strictly forbidden, then one can avoid loss from information concerning a “downside” price move, but cannot make profit out of this; whereas if short selling is allowed, the situation is quite different). Finally, in Section 4.10 we discuss briefly our theory to the continuous-time setup by extending the work of Amendinger *et al.* [3] (for reference purposes their framework is reviewed in Appendix B).

Chapter 1

Information-Based Asset Pricing

The purpose of this chapter is to review the information-based asset pricing framework developed by Brody, Hughston, and Macrina [11, 13, 15, 17, 42] (BHM), which constitutes the background of the most of the material in the first part of this thesis. The BHM framework was initially developed in the context of providing a solution to the pricing and hedging of credit risky assets, where the focus was placed on calculating the conditional expectation of a future cash flow taking various possible states. However, the framework naturally found applications in asset classes well beyond credit-risky instruments, reaching areas such as equity assets, commodities, reinsurance, statistical arbitrage, and stochastic volatility modelling. The BHM framework is perhaps unconventional in that the stochastic process describing the asset dynamics is *derived*, rather than imposed, from more primitive underlying concepts, namely, the modelling of the actual cash flows and the associated flow of information. This is in sharp contrast with the more traditional approach following the seminal work of Black and Scholes [7], but nevertheless is fully consistent with the by now familiar arbitrage-free pricing theory, as presented in [6, 19, 29, 30, 49].

1.1 Market filtration

We consider financial markets with the specification of a filtered probability space denoted by $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{0 \leq t \leq T < \infty}, \mathbb{Q})$, where $\mathbb{Q}$ denotes the risk-neutral measure and the filtration $\{\mathcal{F}_t\}$ is understood to be the market filtration. The market is not assumed to be complete, the absence of arbitrage and the existence of an established pricing kernel imply the existence of a unique risk-neutral measure. We further assume that the interest rate system is deterministic, and let $\{P_{tT}\}_{0 \leq t \leq T < \infty}$ denote the discount function at time t . That the

interest rate system is deterministic implies that the discount function can be expressed in terms of a ratio of initial discount functions:

$$P_{tT} = \frac{P_{0T}}{P_{0t}}.$$

The discount function system is assumed to be differentiable in T and strictly decreasing, and to satisfy the conditions $0 < P_{0t} \leq 1$ and $\lim_{t \rightarrow \infty} P_{0t} = 0$. Under these assumptions, it follows that if an integrable random variable X_T represents the cash flow associated with an asset paid at time T , then the value of the asset at time t is given by:

$$B_{tT} = P_{tT} \mathbb{E}[X_T \mid \mathcal{F}_t]. \quad (1.1)$$

Since X_T represents the value of the future cash flow, market participants do not have direct access to its value before time T . In other words, X_T is only $\mathcal{F}_T$ -measurable. For simplicity, let us assume that the random variable X_T takes discrete values $\{x_1, \dots, x_n\}$ with *a priori* probabilities $\{p_1, \dots, p_n\}$.

Information available to market participants up to time t can be modelled by an information process. The information process has the property such that maximum information in a suitably defined sense is only revealed at time T ; otherwise we have a noisy signal. In particular, let us consider an $\{\mathcal{F}_t\}$ -adapted market information process of the form:

$$\xi_t = \sigma X_T t + \beta_{tT}, \quad (1.2)$$

where the process $\{\beta_{tT}\}_{0 \leq t \leq T}$ is a standard Brownian bridge on the interval $[0, T]$ independent of X_T , and σ is the information revealing rate parameter, which is assumed to be deterministic.

The Brownian bridge $\{\beta_{tT}\}$ is a Gaussian process satisfying $\beta_{0T} = \beta_{TT} = 0$ with mean

$$\mathbb{E}[\beta_{tT}] = 0,$$

and covariance

$$\mathbb{E}[\beta_{sT} \beta_{tT}] = \frac{s(T-t)}{T}$$

for all $0 \leq s \leq t \leq T$. The Brownian bridge represents the “noise” over the time interval $[0, T]$, since it is assumed to be independent of the “signal” term X_T . An intuitive way of understanding noise is to think of it as representing rumours, speculations, and misinfor-

mation. Naturally there are many such noises circulating in financial markets affecting the price dynamics in a random manner. The idea here is to use the Gaussian process $\{\beta_{tT}\}$ to model the aggregate of these different noise contributions.

The construction of the information process implies that information regarding the future cash flow is gradually revealed from zero, and reach a maximum at the terminal time when the value of X_T is revealed. This idea can be illustrated more explicitly by the consideration of the mutual information between X_T and ξ_t , that is, the amount of information contained in ξ_t about the value of X_T , which monotonically increases in time and reaches its maximum at T . The role of the information flow rate parameter σ then becomes more apparent in that the growth rate of the mutual information (the “learning curve”) increases as σ increases. More details on the analysis of mutual information will be presented in the next chapter.

We assume that the market filtration $\{\mathcal{F}_t\}$ is generated by the market information process: $\mathcal{F}_t = \sigma(\{\xi_s\}_{0 \leq s \leq t})$. The cash flow X_T is not $\mathcal{F}_t$ measurable for $t < T$. The bridge process $\{\beta_{tT}\}$ is not adapted to $\{\mathcal{F}_t\}$ and thus is not directly accessible to market participants. This reflects the fact that until the cash flow is paid, the market participants cannot distinguish signal from noise in the market.

1.2 Asset information and price formation

It can be shown that the information process in this setting is a Markov process [17], hence the problem of bond-price valuation in (1.1) is reduced to determining

$$B_{tT} = P_{tT} \mathbb{E}[X_T \mid \xi_t].$$

An alternative way to express the above equality is to write

$$B_{tT} = P_{tT} \sum_{i=1}^n x_i \pi_{it}(x_i \mid \xi_t),$$

where

$$\pi_{it} = \mathbb{Q}(X_T = x_i \mid \xi_t)$$

is the conditional probability of X_T given the information ξ_t .

Using the Bayes formula we deduce that

$$\pi_{it} = \frac{p_i \rho(\xi_t | X_T = x_i)}{\sum_j p_j \rho(\xi_t | X_T = x_j)}.$$

Here, the conditional density of ξ_t can be deduced from the relation

$$\mathbb{Q}(\xi_t < \sigma X_T t + \beta_{tT} | X_T) = \mathbb{Q}(\beta_{tT} < \xi_t - \sigma X_T t | X_T).$$

Recall that conditional on $X_T = x_i$, ξ_t is a Gaussian random variable with mean $\sigma x_i t$ and variance $t(T - t)/T$. Hence we have

$$\rho(\xi_t | X_T = x_i) = \frac{1}{\sqrt{2\pi t(T - t)/T}} \exp\left(-\frac{1}{2} \frac{(\xi_t - \sigma x_i t)^2}{t(T - t)/T}\right).$$

It follows that the bond price in the original BHM model takes the form

$$B_{tT} = P_{tT} \frac{\sum_i x_i p_i \exp[\frac{T}{T-t}(\sigma x_i \xi_t - \frac{1}{2} \sigma^2 x_i^2 t)]}{\sum_i p_i \exp[\frac{T}{T-t}(\sigma x_i \xi_t - \frac{1}{2} \sigma^2 x_i^2 t)]}.$$

Remark: The bond price is derived through the specification of the information process and an application of the Bayes theorem. One can think of the bond price as the noisy *a posteriori* expectation of the future cash flow, suitably discounted.

1.3 Price dynamics

The behaviour of the asset price can be studied by working out the stochastic dynamics of the conditional probability

$$\pi_{it} = \frac{p_i \exp[\frac{T}{T-t}(\sigma x_i \xi_t - \frac{1}{2} \sigma^2 x_i^2 t)]}{\sum_i p_i \exp[\frac{T}{T-t}(\sigma x_i \xi_t - \frac{1}{2} \sigma^2 x_i^2 t)]}.$$

This can be achieved by an application of Ito's lemma. Writing $X_{tT} = \mathbb{E}[X_T | \xi_t]$ for the conditional expectation, we have the dynamic of π_{it} as

$$d\pi_{it} = \frac{\sigma T}{T - t} (x_i - X_{tT}) \pi_{it} dW_t, \quad (1.3)$$

where the process $\{W_t\}_{0 \leq t \leq T}$ defined by

$$W_t = \xi_t + \int_0^t \frac{\xi_s}{T-s} ds - \sigma T \int_0^t \frac{X_{sT}}{T-s} ds \quad (1.4)$$

is an $\{\mathcal{F}_t\}$ -Brownian motion. The fact that the innovation process $\{W_t\}$ is a Brownian motion can be verified by showing that $\mathbb{E}[W_u|\mathcal{F}_t] = W_t$ for $0 \leq t \leq u \leq T$ and that $(dW_t)^2 = dt$, hence by virtue of Lévy's criterion $\{W_t\}$ is a Brownian motion.

An alternative perspective of the problem is to view this as a nonlinear filtering problem, with the solution given by the Fujisaki-Kallianpur-Kunita (FKK) theory (see [23, 49]). Under the filtering setting, given a probability space and a background filtration $\{\mathcal{G}_t\}$, we consider an observation process of the form

$$\xi_t = \int_0^t \mu_s ds + Y_t, \quad (1.5)$$

and an unobservable process

$$z_t = z_0 + \int_0^t \phi_s ds + M_t, \quad (1.6)$$

where $\{Y_t\}$ is a Brownian motion and $\{M_t\}$ a $\mathcal{G}_t$ -martingale. All the processes $\{\mu_t\}$, $\{Y_t\}$, $\{\phi_t\}$ and $\{M_t\}$ in equations (1.5) and (1.6) are assumed to be adapted to the filtration $\{\mathcal{G}_t\}$. The problem is to estimate z_t given the observation $\{\xi_t\}$. The knowledge for the estimation comes from the observation of $\{\xi_t\}$, hence generating the filtration $\{\mathcal{F}_t\}$. Denote an estimate by

$$\hat{z}_t = \mathbb{E}[z_t|\mathcal{F}_t].$$

The basic result of nonlinear filtering theory states that the dynamics of the estimate is of the following form

$$d\hat{z}_t = \hat{\phi}_t + [(\widehat{\mu z})_t - \hat{\mu}_t \hat{z}_t] dW_t, \quad (1.7)$$

where the *innovation process*

$$W_t = \xi_t - \int_0^t \hat{\mu}_s ds \quad (1.8)$$

is an $\{\mathcal{F}_t\}$ -Brownian Motion.

To use the result above, we can rewrite ξ_t by considering the stochastic differential form of the Brownian bridge

$$d\beta_{tT} = -\frac{\beta_{tT}}{T-t}dt + dY_t$$

for $0 \leq t < T$, where $\{Y_t\}$ a standard Brownian motion. And take note of the fact that $\xi_t = \sigma X_T t + \beta_{tT}$:

$$d\xi_t = \frac{\sigma X_T T - \xi_t}{T-t}dt + dY_t.$$

Comparing this to equation (1.5) we deduce that

$$\mu_t = \frac{\sigma X_T T - \xi_t}{T-t}.$$

The estimate of this process is therefore

$$\hat{\mu}_t = \frac{\sigma X_{tT} T - \xi_t}{T-t},$$

where $X_{tT} = \mathbb{E}[X_T | \mathcal{F}_t]$. Substituting the above expression for $\hat{\mu}_t$ into equation (1.8), we obtain equation (1.4). For the dynamics of $\{\pi_{it}\}$, we consider the case when our unobservable process takes the form

$$z_t = \mathbb{1}\{X_T = x_i\},$$

where $\mathbb{1}\{\cdot\}$ denotes the indicator function. It follows that

$$\begin{aligned} \hat{z}_t &= \mathbb{E}[\mathbb{1}\{X_T = x_i\} | \mathcal{F}_t] \\ &= \mathbb{Q}(X_T = x_i | \mathcal{F}_t), \end{aligned}$$

which by definition is just π_{it} . Hence

$$(\widehat{\mu z})_t = \frac{\sigma x_i T - \xi_t}{T-t} \pi_{it}.$$

Substituting this, $\hat{z}_t$ and $\hat{\mu}_t$ into equation (1.7), we get equation (1.3). This also allows us to deduce that $\{W_t\}$, represented by equation (1.4), is an $\mathcal{F}_t$ -Brownian motion.

Finally, we can write down the discount bond dynamics

$$dB_{tT} = r_t B_{tT} dt + \Sigma_{tT} dW_t,$$

where the short rate takes the form

$$r_t = -\frac{\partial \ln P_{0t}}{\partial t},$$

and the absolute bond volatility Σ_{tT} is given by

$$\Sigma_{tT} = \frac{\sigma T}{T-t} P_{tT} V_{tT},$$

where the conditional variance of X_T is defined as

$$V_{tT} = \sum_{i=1}^n (x_i - \mathbb{E}[X_T | \xi_t])^2 \pi_{it},$$

in which we see that the information flow rate parameter σ controls the overall magnitude of the absolute volatility. The important result is that the bond price process is derived from the information process together with the specification of the cash flow X_T . Information from the market is somehow filtered into knowledge of the price dynamics. This is intuitive, as any price movement is closely related to updates of market information. Anderson *et al.* [4] has demonstrated, for instance, with high frequency historical data, macroeconomic news is a driving factor of the price dynamics, and has direct contributions to conditional mean jumps.

Conversely, one can extract the information status from the price process. In terms of filtrations, the information contained in $\{\xi_t\}$ is equivalent to that in $\{B_{tT}\}$, i.e.

$$\sigma(\{\xi_s\}_{0 \leq s \leq t}) = \sigma(\{B_{sT}\}_{0 \leq s \leq t}).$$

This follows from the fact that if we define $B(t, \xi)$ such that $B_{tT} = B(t, \xi_t)$. Then

$$\left. \frac{\partial B(t, \xi)}{\partial \xi} \right|_{\xi=\xi_t} = \frac{\sigma T}{T-t} \frac{\sum_i p_i [x_i - B_{tT}/P_{tT}]^2 \exp \left[\frac{T}{T-t} (\sigma x_i \xi_t - \frac{1}{2} \sigma^2 x_i^2 t) \right]}{\sum_i p_i \exp \left[\frac{T}{T-t} (\sigma x_i \xi_t - \frac{1}{2} \sigma^2 x_i^2 t) \right]}$$

is a positive quantity. Therefore, $B(t, \xi_t)$ is monotonically increasing in ξ_t , and hence invertible. Economically, this means that the observed price or volatility has already incorporated the structure of news, uncertainty and expectation of the future cash flow.

1.4 Securities and derivative pricing

This simple setting already allows us to price a range of different products including options on bonds, barrier options, credit default swaps, and basket of credit risky bonds, and so on. The price of a vanilla European option, in particular, can be derived in closed form, and the structure is similar to that of the Black-Scholes. To sketch the procedure for pricing a European option on a credit-risky bond that matures at time T , with exercise date $t < T$, one considers the price of an option with the payoff $(B_{tT} - K)^+$ at time t :

$$C_0 = P_{0t} \mathbb{E}[(B_{tT} - K)^+],$$

where B_{tT} is the bond price at the date the option matures, and K is the strike price. This equality can be rewritten in terms of the conditional probability π_{it} as follows

$$C_0 = P_{0t} \mathbb{E} \left[\left(\sum_i P_{tT} \pi_{it} x_i - K \right)^+ \right]. \quad (1.9)$$

For simplicity of notation, we denote the numerator of the conditional probability π_{it} by

$$p_{it} = p_i \exp \left[\frac{T}{T-t} \left(x_i \sigma \xi_t - \frac{1}{2} x_i^2 \sigma^2 t \right) \right].$$

Consequently we have the following:

$$\pi_{it} = \frac{p_{it}}{\Phi_t} \quad \text{where} \quad \Phi_t = \sum_i p_{it}.$$

Hence equation (1.9) can be rewritten as:

$$C_0 = P_{0t} \mathbb{E} \left[\frac{1}{\Phi_t} \left(\sum_i (P_{tT} x_i - K) p_{it} \right)^+ \right].$$

The term Φ_t^{-1} is shown to be a $\mathbb{Q}$ -martingale in [17] and can be used to effect a change of measure. The call option price is therefore given by

$$C_0 = P_{0t} \mathbb{E}^{\mathbb{B}_T} \left[\left(\sum_{i=0}^n (P_{tT} x_i - K) p_{it} \right)^+ \right],$$

where p_{it} is the unnormalised conditional probability of X_T given ξ_t and $\mathbb{B}_T$ denotes the bridge measure. Essentially, the density of this bridge measure is the normalising factor of

the conditional expectation, and ξ_t is a $\mathbb{B}_T$ -Gaussian random variable.

If we consider a digital bond for simplicity, so that X_T takes the values $\{x_0, x_1\}$ with probabilities $\{p_0, p_1\}$, respectively, the initial price of a call option becomes

$$C_0 = P_{0t} \mathbb{E}^{\mathbb{B}_T} \left[((P_{tT}x_0 - K) p_{0t} + (P_{tT}x_1 - K) p_{1t})^+ \right].$$

Depending on the values of $P_{tT}x_i$ the option can expire in-the-money or out-of-the-money for certain. The only nontrivial case is when $P_{tT}x_1 \geq K \geq P_{tT}x_0$, where the option can expire either in or out-of-the-money, depending on whether $\xi_t > \xi^*$, where ξ^* is the unique critical value of ξ_t such that $B_{tT} = K$. In this case, we can compute the expectation explicitly, since ξ_t is a $\mathbb{B}_T$ -Gaussian random variable. By performing the Gaussian integration explicitly, we arrive at the BHM option pricing formula

$$C_0 = P_{0t} \sum_{i=0}^n (P_{tT}x_i - K) N(\sigma x_i \sqrt{\tau} - \bar{Z}),$$

where $N(a)$ is the Gaussian cumulative distribution function up to a , $\tau = tT/(T - t)$ and $Z = \xi_t^*/\sqrt{t(T - t)/T}$. The final result is no more complicated computationally than the Black-Scholes option formula.

Chapter 2

Stochastic Information Revealing Rate

The purpose of this section is to develop a model that extends the original BHM model reviewed in chapter 1 such that the volatility parameter is assumed stochastic. We shall begin by arguing the economic rationale of considering such an extension, which is augmented by a sensitivity analysis of the original BHM model on the volatility parameter. We then derive the bond pricing formula in the stochastic volatility setup, and investigate in some detail the dynamical behaviour of the price process both analytically and numerically.

2.1 Motivation

The fact that it is market information that affects price dynamics has been emphasised by many authors [2, 20, 40, 52], and has also been demonstrated against market data (e.g., [4, 8, 12]). The aim of the BHM approach is to bring the mathematical abstraction of financial modelling at the level of the specification of market filtration, i.e. the flow of information. In this way, price process can be derived as an emergent phenomenon, rather than postulated from the outset, as we have illustrated in the previous chapter.

One of the simplifying assumptions in the original BHM model (1.2) and the various generalisations of it that have appeared in the literature [16, 42] is that the information flow-rate parameter σ is taken to be $\mathcal{F}_0$ -measurable. In a more realistic setup, however, market participants have little knowledge about the value of σ . In fact, in most cases the information flow rate is not measurable even after the value of X_T is revealed. The main issue addressed in this and the next chapter is therefore to extend the original BHM model to allow for σ to be a random variable. Such an extension constitutes the simplest stochastic volatility model within the BHM framework, which is shown in the next chapter to give sufficient

flexibility to calibrate volatility surfaces.

The information revealing rate reflects the learning rate of the terminal value of the bond. In other word, it is the growth rate of the filtration, or the information set, as time passes. It is possible that the growth rate is not measurable, and we only have a range of uncertain values. This parameter σ is in fact fairly difficult to determine—at which rate market participants “learn” is not straightforward to measure. Hence it is reasonable to assume that σ is not measurable for any $t < T$, and possibly even when $t = T$. We begin our random information/signal model with σ being a random variable that is fixed at the initial time. This is equivalent to having a range of “guesses” for $\sigma = \{\sigma_k\}_{1 \leq k \leq m}$, each with a probability assigned. The price process is thus expected to behave like a mixture of the constant σ model—the BHM mixture model. This has a certain similarity to the log-normal mixture models, in particular, the uncertain-volatility model in Mercurio *et al.* [43]. Our approach of having a fixed random variable σ is the first step towards addressing stochastic volatility in the information-based pricing framework. It is aimed to tackle the essence of the problem, and provide a tractable model for the asset price.

As shown in chapter 1, the rate at which information is revealed in the market determines the magnitude of the volatility. Thus σ can be viewed as the “volatility parameter” for the bond price. In the stochastic setting, the volatility can be thought of as reflecting mixed views of expectations contained in the market prices of options of different strikes. The calibration of the implied volatility skew therefore can be used to extract this information from the market prices. This is different from the more conventional idea of modelling the volatility explicitly with a structural model; instead, we “filter” it from the factor that drives it.

2.2 Setup

Let us formally introduce the stochastic information model. As before, our financial market is specified by the probability space $(\Omega, \mathcal{F}, \mathbb{Q})$; we assume the existence of a preferred pricing kernel, and the absence of arbitrage. The deterministic discounting system $\{P_{tT}\}$ is as introduced before. The information process at time t that generates the market filtration $\mathcal{F}_t$ is defined by

$$\xi_t = \bar{\sigma} X_T t + \beta_{tT}, \quad (2.1)$$

where X_T and β_{tT} are random variables denoting the terminal payoff and the Brownian bridge process, respectively. The information revealing rate parameter $\bar{\sigma}$ is now a random

variable taking values $\{\sigma_1, \dots, \sigma_m\}$ with probabilities $\{q_1, \dots, q_m\}$. All variables are assumed independent. For the moment we take $\bar{\sigma}$ to be a discrete random variable for simplicity. However, as we show below, for the process (2.1) to form market filtration $\bar{\sigma}$ in fact has to be discrete.

2.3 Meaning of the information process

Before we proceed to investigate properties of the model with a random information flow rate in detail, let us first comment on the interpretation of the information process (2.1). In general, in a financial market, even if we make the simplifying assumption that the only relevant market factor is the cash flow X_T , there are plenty of information sources available for X_T , each being obscured by noise. We can therefore represent each of the information source in the signal-plus-noise form (1.2) and write

$$\begin{cases} \xi_t^1 = \sigma_1 X_T t + \beta_{tT}^1 \\ \vdots \\ \xi_t^n = \sigma_n X_T t + \beta_{tT}^n \end{cases} \quad (2.2)$$

for the family of information processes available in the market concerning the impending cash flow X_T . The various noise processes $\{\beta_{tT}^i\}_{i=1,\dots,n}$ in general may be mutually correlated (with correlation matrix ρ), but they are all independent of X_T .

An important point to observe now is the fact that the aggregate information processes (2.2) is somewhat redundant; the information relevant to the cash flow X_T contained in (2.2) can be represented in the form of a single information process (1.2), with the choice

$$\sigma^2 = \frac{\sum_i^n \sigma_i^2 \rho_{ii}^{-1} - 2 \sum_{i \neq j} \sigma_i \sigma_j \rho_{ij}^{-1}}{\det(\rho)} \quad (2.3)$$

for the *effective information flow rate*, and

$$\beta_{tT} = \frac{1}{\sigma} \left(\sum_{i,j}^N \sigma_i \rho_{ij}^{-1} \beta_{tT}^i \right)$$

for the *effective noise*. Here ρ_{ij}^{-1} denotes the ij element of the inverse correlation matrix.

Putting the matter differently, the filtration generated jointly by the set of information processes (2.2) is equivalent to the filtration generated jointly by the single information process (1.2) and a family of noise processes given by combinations of $\{\beta_{tT}^i\}$. However, since the

noise terms are independent of X_T , they make no contribution towards the pricing of the asset. We can therefore discard them altogether and represent the totality of ‘relevant’ information in the form of a single information process (2.1).

Remark: In the case of a pair of information processes for the same market factor X_T , the construction of an effective information process is used effectively in [13] to characterise the behaviour of an informed trader having access to additional noisy information.

2.4 Sensitivity analysis

It is of interest to identify the sensitivity of the BHM model (1.2) to the specification of the information flow-rate parameter σ , given the fact that the value of σ is usually unknown. Often one considers the option vega as a measure of parameter sensitivity, but here we are interested in a global measure of parameter sensitivity. The result will be useful in identifying the region in the parameter space for which a misspecification of the flow-rate parameter σ yields significant errors in pricing a range of products, not just vanilla options.

A universal measure of sensitivity in statistical analysis is given by the Fisher information [22]. To work out the Fisher information associated with the parameter σ we proceed as follows. For simplicity, let us assume that the cash flow X_T takes discrete values $\{x_i\}$ with *a priori* probability $\{p_i\}$. To determine the information measure of Fisher we need the expression for the *a posteriori* probability $\pi_{it} = \mathbb{Q}(X_T = x_i | \mathcal{F}_t)$. This is given by

$$\pi_{it} = \frac{p_i \exp\left(\frac{T}{T-t} \left(\sigma x_i \xi_t - \frac{1}{2} \sigma^2 x_i^2 t\right)\right)}{\sum p_i \exp\left(\frac{T}{T-t} \left(\sigma x_i \xi_t - \frac{1}{2} \sigma^2 x_i^2 t\right)\right)}.$$

We can therefore regard $\{\pi_{it}\} = \{\pi_{it}(\sigma)\}$ as a one-parameter family of probabilities. The Fisher information $g_t(\sigma)$ associated with the parameter σ is then defined by the expression:

$$g_t(\sigma) = \sum_i \frac{1}{\pi_{it}(\sigma)} \left(\frac{\partial \pi_{it}(\sigma)}{\partial \sigma} \right)^2.$$

Remark: It is interesting that the Fisher information in the case of the BHM model (1.2) has the interpretation in terms of the conditional variance:

$$g_t(\sigma) = \sigma t \left(\frac{T}{T-t} \right)^2 \text{Var}(X_T \beta_{tT} | \xi_t). \quad (2.4)$$

To see this, let us define

$$p_{it}(\sigma) = p_i \exp \left(\frac{T}{T-t} \left(\sigma x_i \xi_t - \frac{1}{2} \sigma^2 x_i^2 t \right) \right),$$

so that $\pi_{it} = p_{it} / \sum_i p_{it}$. We then have

$$\frac{\partial p_{it}(\sigma)}{\partial \sigma} = \frac{T}{T-t} (x_i \xi_t - \sigma x_i^2 t) p_{it}(\sigma),$$

from which it follows that

$$\begin{aligned} \frac{\partial \pi_{it}}{\partial \sigma} &= \frac{1}{(\sum_i p_{it})^2} \left[\frac{\partial p_{it}}{\partial \sigma} \sum_i p_{it} - p_{it} \sum_i \frac{\partial p_{it}}{\partial \sigma} \right] \\ &= \frac{T}{T-t} p_{it} \left(\frac{x_i \xi_t - \sigma x_i^2 t}{\sum_i p_{it}} - \frac{\sum_i p_{it} (x_i \xi_t - \sigma x_i^2 t)}{(\sum_i p_{it})^2} \right). \end{aligned}$$

Therefore, we obtain

$$g_t(\sigma) = \left(\frac{T}{T-t} \right)^2 \frac{\sum_i p_{it} \left(x_i \xi_t - \sigma x_i^2 t - \frac{\sum_i p_{it} (x_i \xi_t - \sigma x_i^2 t)}{\sum_i p_{it}} \right)^2}{\sum_i p_{it}},$$

and by substitution of (1.2) in here we deduce (2.4).

The reason that the Fisher information $g_t(\sigma)$ measures the parameter sensitivity follows from the celebrated Cramér-Rao inequality, which shows that the variance of the parameter estimate is bounded below by the inverse of the Fisher information. Additionally, as observed by Rao [47], the separation, i.e. the divergence measure associated with a pair of models characterised by $\pi_{it}(\sigma)$ and $\pi_{it}(\sigma')$ is determined by the integral:

$$D_t(\sigma, \sigma') = \int_{\sigma}^{\sigma'} \sqrt{g_t(u)} \, du. \quad (2.5)$$

In figure 2.1 we show the numerical plot of the expectation $\mathbb{E}[g_t(\sigma)]$ of the Fisher information for different values of t , averaged over 1,000 sample paths. The parameters are chosen such that $\{x_i\} = \{0, 0.5, 1\}$, $\{p_i\} = \{0.1, 0.15, 0.75\}$, and $T = 1$. We observe that the Fisher information is increasing in σ . It is evident from the plot that the model is quite sensitive to the choice of the information flow rate parameter σ up until about two thirds of the way into the duration of the contract. The price of a generic derivative on a credit risky bond, with maturity not too close to zero but also shorter than two-third of the bond maturity, is thus likely to be sensitive to the specification of the parameter σ in the

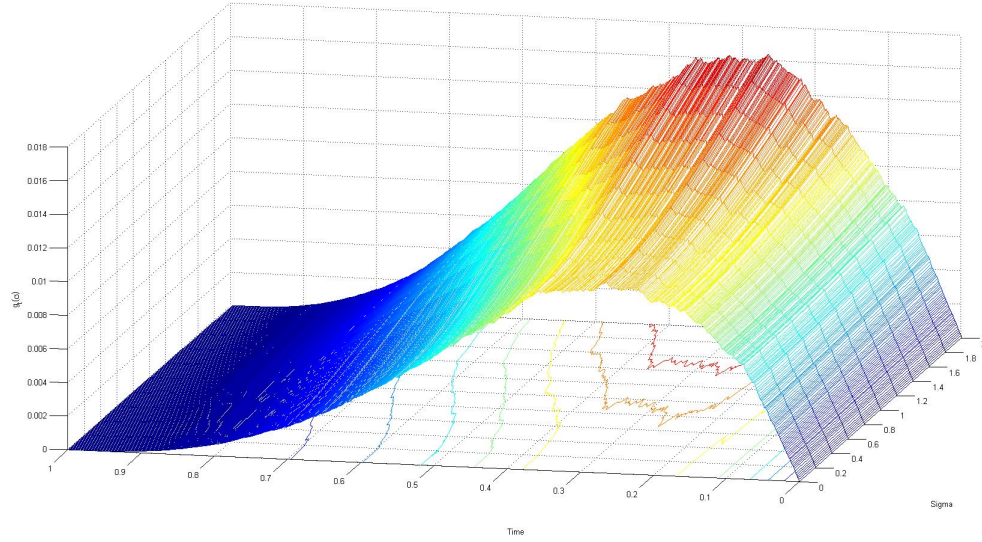


Figure 2.1: The unconditional expectation $\mathbb{E}[g_t(\sigma)]$ of the Fisher information for different values of t and σ . The horizontal axis represents the time remaining to the maturity. The vertical axis represents the values of σ , ranging from 0 to 2. The numerical result is obtained by averaging over 1,000 sample paths. The parameters are chosen such that $\{x_i\} = \{0, 0.5, 1\}$, $\{p_i\} = \{0.1, 0.15, 0.75\}$, and $T = 1$. Where the value of $\mathbb{E}[g_t(\sigma)]$ is high, the model is on average sensitive to the choice of the specification of the parameter σ .

most basic version of the BHM model (1.2). This sensitivity can be made more robust by allowing σ to be random.

2.5 Quantisation of information flow rate

As indicated above, the complicated way in which the market information flow rate σ depends on individual information flow rates $\{\sigma_i\}$, as represented in (2.3), suggests that it is not reasonable to assume that market participants have access to the value of σ . Indeed, in many realistic setups it is unlikely that market participants will ever learn the value of σ , even after the value of X_T is revealed. More generally, the value of σ can change over time in a random manner. Here we shall be considering the simplest such situation in which σ is given by a fixed random variable independent of X_T and $\{\beta_{tT}\}$. We observe, however, that whether σ is a fixed random variable or a stochastic process, the $\mathcal{F}_T$ -measurability of X_T enforces a stringent constraint on the allowable choices for σ .

To see this, let us consider a simple example. Suppose that the random cash flow X_T takes three possible values $\{0, 0.5, 1\}$ with *a priori* probabilities $\{p_1, p_2, p_3\}$, and σ takes two

possible values $\{0.5, 1\}$ with *a priori* probabilities $\{q_1, q_2\}$. Then there are four possible realised values for the terminal information ξ_T :

$$\xi_T = \sigma T X_T = \begin{cases} 0 & \Rightarrow X_T = 0, \quad \sigma = 0.5 \text{ or } 1; \\ 0.25T & \Rightarrow (X_T, \sigma) = (0.5, 0.5); \\ 0.5T & \Rightarrow (X_T, \sigma) = (0.5, 1) \text{ or } (1, 0.5); \\ T & \Rightarrow (X_T, \sigma) = (1, 1). \end{cases} \quad (2.6)$$

In this case, if the realised value of ξ_T happens to be 0, $0.25T$, or T , then we can unambiguously determine the value of X_T irrespective of what the value of σ might have been, whereas if the outcome happens to be $0.5T$, then the value of X_T could be 0.5 or 1.

This example thus illustrates the fact that once the cash flow random variable is given, the information flow rate cannot take an arbitrary random variable. In particular, σ can not be a continuous random variable. We define the following admissible set for σ :

$$\Sigma(X_T) = \{\sigma : X_T \text{ is } \mathcal{F}_T\text{-measurable}\}.$$

In the rest of this chapter we shall always assume that $\sigma \in \Sigma(X_T)$, unless specified otherwise. More generally, if $\sigma = \sigma(t)$ is time dependent, then $\sigma(t)$ for $t < T$ can be arbitrary, but the constraint discussed here for the value of $\sigma(T)$ is still applicable. Therefore, there is a kind of “quantisation” condition imposed on σ . For the same token, in the random- σ environment, the cash flow X_T cannot be a continuous variable—this does not pose real constraints on our approach because cash flows are in reality never continuous.

Remark: it is an unsolved problem to determine all “compatible” pairs (X_T, σ) such that σX_T “reveals” X_T .

2.6 Price process with a random information flow rate

Bearing in mind the quantisation condition imposed by the measurability of X_T , we proceed to consider the random- σ extension of the BHM model. Specifically, we let X_T take the values $\{x_i\}_{i=1,\dots,n}$ with *a priori* probabilities $\{p_i\}_{i=1,\dots,n}$; and let $\bar{\sigma}$ take the values $\{\sigma_k\}_{k=1,\dots,m}$ with *a priori* probabilities $\{q_k\}_{k=1,\dots,m}$. These random variables are chosen such that degenerate situations like the example given in (2.6) are excluded, and hence X_T is ensured to be $\mathcal{F}_T$ -measurable.

We shall now determine the price of a credit-risky bond, with the new information process.

By the fundamental theorem of asset pricing, the bond price at time t is

$$\begin{aligned} B_{tT} &= P_{tT} \mathbb{E}[X_T \mid \xi_t] \\ &= P_{tT} \sum_i^n x_i \pi_{it} \end{aligned}$$

where $\pi_{it} = \mathbb{Q}(X_T = x_i \mid \mathcal{F}_t)$ is the conditional probability for the random variable X_T . The calculation of the bond price can be simplified somewhat if the information process $\{\xi_t\}$ is Markovian. This, indeed, turns out to be the case, as we show below.

2.6.1 Markovian property of ξ_t

To show that the information process satisfies the Markov property, we need to verify that

$$\mathbb{Q}(\xi_t \leq x \mid \mathcal{F}_s) = \mathbb{Q}(\xi_t \leq x \mid \xi_s)$$

for all $x \in \mathbb{R}$ and all s, t such that $0 \leq s \leq t \leq T$. For this we shall follow closely the argument of [17]. It suffices to show that

$$\mathbb{Q}(\xi_t \leq x \mid \xi_s, \xi_{s_1}, \dots, \xi_{s_k}) = \mathbb{Q}(\xi_t \leq x \mid \xi_s)$$

for any collection of times $t, s, s_1, \dots, s_k$ such that $T \geq t > s > s_1 > \dots > s_k > 0$. Using the definition of Brownian bridge we note that

$$\mathbb{E} \left[\beta_{tT} \left(\frac{\beta_{sT}}{s} - \frac{\beta_{s_1T}}{s_1} \right) \right] = 0,$$

i.e. β_{tT} and $(\beta_{sT}/s - \beta_{s_1T}/s_1)$ are independent. More generally, $(\beta_{sT}/s - \beta_{s_1T}/s_1)$ and $(\beta_{s_2T}/s_2 - \beta_{s_3T}/s_3)$ are independent. We have the information process

$$\xi_t = \bar{\sigma} t X_T + \beta_{tT},$$

where both $\bar{\sigma}$ and X_T are time-independent random variables, and X_T is $\mathcal{F}_T$ measurable. Hence we can write

$$\frac{\xi_s}{s} - \frac{\xi_{s_1}}{s_1} = \frac{\beta_{sT}}{s} - \frac{\beta_{s_1T}}{s_1}$$

and

$$\begin{aligned}\mathbb{Q}(\xi_t \leq x \mid \xi_s, \xi_{s_1}, \dots, \xi_{s_k}) &= \mathbb{Q}\left(\xi_t \leq x \mid \xi_s, \frac{\xi_s}{s} - \frac{\xi_{s_1}}{s_1}, \dots, \frac{\xi_{s_{k-1}}}{s_{k-1}} - \frac{\xi_{s_k}}{s_k}\right) \\ &= \mathbb{Q}\left(\xi_t \leq x \mid \xi_s, \frac{\beta_{sT}}{s} - \frac{\beta_{s_1T}}{s_1}, \dots, \frac{\beta_{s_{k-1}T}}{s_{k-1}} - \frac{\beta_{s_kT}}{s_k}\right).\end{aligned}$$

However, since ξ_s and ξ_t are independent of the remaining variables that depend only on the noise, i.e. $\beta_{sT}/s - \beta_{s_1T}/s_1, \dots, \beta_{s_{k-1}T}/s_{k-1} - \beta_{s_kT}/s_k$, the desired Markov property follows immediately.

Therefore, our problem is now reduced to calculating

$$B_{tT} = P_{tT}\mathbb{E}[X_T \mid \xi_t].$$

This is equivalent to determining the conditional probability

$$\pi_{it} = \mathbb{Q}(X_T = x_i \mid \xi_t).$$

2.6.2 Conditional probability

Given the information process $\xi_t = \bar{\sigma}tX_T + \beta_{tT}$, where X_T takes values x_i with probabilities p_i for $i = \{1, \dots, n\}$ and $\bar{\sigma}$ takes values σ_k with probabilities q_k for $k = \{1, \dots, m\}$, there are several ways of deriving the posterior conditional probability of the random variable X_T using the Bayes formula. The key result that we shall be using is as follows:

$$\mathbb{Q}(X_T = x_i \mid \xi_t) = \sum_k \mathbb{Q}(X_T = x_i, \bar{\sigma} = \sigma_k \mid \xi_t) \quad (2.7)$$

$$= \sum_k \mathbb{Q}(X_T = x_i \mid \xi_t, \bar{\sigma} = \sigma_k) \mathbb{Q}(\bar{\sigma} = \sigma_k) \quad (2.8)$$

$$= \frac{\sum_k p_i q_k \mathbb{Q}(\xi_t \mid X_T = x_i, \bar{\sigma} = \sigma_k)}{\sum_i \sum_k p_i q_k \mathbb{Q}(\xi_t \mid X_T = x_i, \bar{\sigma} = \sigma_k)} \quad (2.9)$$

We derive this result in three different ways. First, on account of the Bayes formula we have

$$\begin{aligned}\mathbb{Q}(X_T = x_i \mid \xi_t) &= \frac{\mathbb{Q}(X_T = x_i) \mathbb{Q}(\xi_t \mid X_T = x_i)}{\sum_i \mathbb{Q}(X_T = x_i) \mathbb{Q}(\xi_t \mid X_T = x_i)} \\ &= \frac{\sum_k p_i q_k \mathbb{Q}(\xi_t \mid X_T = x_i, \bar{\sigma} = \sigma_k)}{\sum_i \sum_k p_i q_k \mathbb{Q}(\xi_t \mid X_T = x_i, \bar{\sigma} = \sigma_k)}.\end{aligned}$$

An alternative way of deriving from (2.7) to (2.9) is as follows:

$$\begin{aligned}
\sum_k \mathbb{Q}(X_T = x_i, \bar{\sigma} = \sigma_k \mid \xi_t) &= \frac{\sum_k \mathbb{Q}(\xi_t, X_T = x_i, \bar{\sigma} = \sigma_k)}{\mathbb{Q}(\xi_t)} \\
&= \frac{\sum_k \mathbb{Q}(\xi_t \mid X_T = x_i, \bar{\sigma} = \sigma_k) \mathbb{Q}(x_i) \mathbb{Q}(\sigma_k)}{\sum_i \mathbb{Q}(\xi_t \mid X_T = x_i) \mathbb{P}(x_i)} \\
&= \frac{\sum_k p_i q_k \mathbb{Q}(\xi_t \mid X_T = x_i, \bar{\sigma} = \sigma_k)}{\sum_i \sum_k p_i q_k \mathbb{Q}(\xi_t \mid X_T = x_i, \bar{\sigma} = \sigma_k)}.
\end{aligned}$$

The first equality holds because ξ_t does not depend on the index k . The second equality holds is because X_T and σ are independent. Finally, yet another way of establishing the step from (2.8) to (2.9) is as follows:

$$\begin{aligned}
\sum_k \mathbb{Q}(\sigma_k) \mathbb{Q}(X_T = x_i \mid \xi_t, \sigma_k) &= \sum_k \mathbb{Q}(\sigma_k) \frac{\mathbb{Q}(\xi_t, X_T = x_i, \sigma_k)}{\mathbb{Q}(\xi_t, \sigma_k)} \\
&= \sum_k \mathbb{Q}(\sigma_k) \frac{\mathbb{Q}(\sigma_k \mid \xi_t, X_T = x_i) \mathbb{Q}(\xi_t, x_i)}{\mathbb{Q}(\sigma_k \mid \xi_t) \mathbb{Q}(\xi_t)} \\
&= \frac{\mathbb{Q}(\xi_t, X_T = x_i)}{\mathbb{Q}(\xi_t)} \sum_k \mathbb{Q}(\sigma_k) \\
&= \frac{\mathbb{Q}(\xi_t \mid X_T = x_i) \mathbb{Q}(x_i)}{\mathbb{Q}(\xi_t)} \\
&= \frac{\sum_k p_i q_k \mathbb{Q}(\xi_t \mid X_T = x_i, \bar{\sigma} = \sigma_k)}{\sum_i \sum_k p_i q_k \mathbb{Q}(\xi_t \mid X_T = x_i, \bar{\sigma} = \sigma_k)}.
\end{aligned}$$

On the second line, we have used the property that X_T and σ are independent, which implies that $\mathbb{Q}(\sigma_k \mid \xi_t, X_T = x_i) = \mathbb{Q}(\sigma_k \mid \xi_t)$ and the result follows as required.

To obtain an explicit expression for $\mathbb{Q}(X_T \mid \xi_t)$, we note that conditional on $X_T = x_i$ and $\bar{\sigma} = \sigma_k$ the random variable $\xi_t = \bar{\sigma} X_T t + \beta_{tT}$ has essentially the same probability law as the Brownian bridge β_{tT} . Hence

$$\mathbb{Q}(\xi_t \mid X_T = x_i, \bar{\sigma} = \sigma_k) = \frac{1}{\sqrt{2\pi t(T-t)/T}} \exp \left(-\frac{1}{2} \frac{(\xi_t - \sigma_k x_i t)^2}{t(T-t)/T} \right).$$

2.7 Credit-risky bond dynamics

Under the present model, the value of a credit-risky bond is given by $B_{tT} = P_{tT} \mathbb{E}[X_T \mid \xi_t]$, where we have assumed a deterministic discounting system P_{tT} that is differentiable and

strictly decreasing. Let us write

$$\begin{aligned} X_{tT} &= \mathbb{E}[X_T \mid \xi_t] \\ &= \sum_i^n x_i \pi_{it}. \end{aligned}$$

We have for the conditional probability $\{\pi_{it}\}$ the following expression:

$$\pi_{it} = \frac{\sum_{k=1}^m p_i q_k \exp \left[\frac{T}{T-t} \left(x_i \sigma_k \xi_t - \frac{1}{2} x_i^2 \sigma_k^2 t \right) \right]}{\sum_{s=1}^n \sum_{r=1}^m p_s q_r \exp \left[\frac{T}{T-t} \left(x_s \sigma_r \xi_t - \frac{1}{2} x_s^2 \sigma_r^2 t \right) \right]}. \quad (2.10)$$

Thus the Bond price is given by

$$B_{tT} = P_{tT} \frac{\sum_i^n \sum_{k=1}^m x_i p_i q_k \exp \left[\frac{T}{T-t} \left(x_i \sigma_k \xi_t - \frac{1}{2} x_i^2 \sigma_k^2 t \right) \right]}{\sum_{s=1}^n \sum_{r=1}^m p_s q_r \exp \left[\frac{T}{T-t} \left(x_s \sigma_r \xi_t - \frac{1}{2} x_s^2 \sigma_r^2 t \right) \right]}.$$

Having in hand the present value of a credit-risky bond that is formulated based on the information process $\{\xi_t\}$, we would like to examine the dynamics of the bond price over time. To see this we can work out the dynamic of the conditional probability using Ito's lemma. Let us write $p_{ikt} = p_{ik}(t, \xi)$, where

$$p_{ik}(t, \xi) = p_i q_k \exp \left[\frac{T}{T-t} \left(x_i \sigma_k \xi - \frac{1}{2} x_i^2 \sigma_k^2 t \right) \right].$$

Diffusing this, we get

$$dp_{ikt} = \frac{\sigma_k T}{T-t} x_i \left(d\xi_t + \frac{\xi_t}{T-t} dt \right) p_{ikt}.$$

Summing both sides over the index k , we have

$$d \sum_k p_{ikt} = \frac{T}{T-t} x_i \left(d\xi_t + \frac{\xi_t}{T-t} dt \right) \sum_k \sigma_k p_{ikt}.$$

Thus by use of Ito's quotient rule, the dynamics of π_{it} can be calculated as follows:

$$\begin{aligned} \frac{d\pi_{it}}{\pi_{it}} &= \frac{1}{\pi_{it}} d \left(\frac{\sum_k p_{ikt}}{\sum_i \sum_k p_{ikt}} \right) \\ &= \frac{T}{T-t} x_i \left(d\xi_t + \frac{\xi_t}{T-t} dt \right) \frac{\sum_k \sigma_k p_{ikt}}{\sum_k p_{ikt}} - \frac{T}{T-t} \left(d\xi_t + \frac{\xi_t}{T-t} dt \right) \frac{\sum_i \sum_k x_i \sigma_k p_{ikt}}{\sum_i \sum_k p_{ikt}} \\ &\quad + \frac{T^2}{(T-t)^2} dt \left(\frac{\sum_i \sum_k x_i \sigma_k p_{ikt}}{\sum_i \sum_k p_{ikt}} \right)^2 - \frac{T^2}{(T-t)^2} x_i dt \frac{\sum_k \sigma_k p_{ikt}}{\sum_k p_{ikt}} \frac{\sum_i \sum_k \sigma_k x_i p_{ikt}}{\sum_i \sum_k p_{ikt}}. \end{aligned}$$

Now let us denote by

$$\mathbf{S} = \begin{pmatrix} \sigma_1 \\ \vdots \\ \sigma_m \end{pmatrix} \quad \text{and} \quad \mathbf{X} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

the column vectors for σ_k and x_i , respectively; and an $n \times m$ matrix $\mathbf{P}$, whose ij^{th} element is given by

$$\frac{p_{ijt}}{\sum_i \sum_k p_{ikt}}. \quad (2.11)$$

Therefore, we can write

$$\frac{\sum_i \sum_k x_i \sigma_k p_{ikt}}{\sum_i \sum_k p_{ikt}} = \mathbf{X}^T \mathbf{P} \mathbf{S} \quad \text{and} \quad \frac{\sum_k \sigma_k p_{ikt}}{\sum_i \sum_k p_{ikt}} = \mathbf{P} \mathbf{S}.$$

Evidently, $\mathbf{X}^T \mathbf{P} \mathbf{S}$ is a scalar, and $\mathbf{P} \mathbf{S}$ is a column vector of n elements. As for the dynamics of $\{\pi_{it}\}$ we thus have

$$\begin{aligned} \frac{d\pi_{it}}{\pi_{it}} &= \frac{T}{T-t} d\xi_t \left[\frac{\sum_k \sigma_k p_{ikt}}{\sum_k p_{ikt}} \frac{\sum_i \sum_k p_{ikt}}{\sum_i \sum_k p_{ikt}} x_i - \frac{\sum_i \sum_k x_i \sigma_k p_{ikt}}{\sum_i \sum_k p_{ikt}} \right] \\ &+ \frac{T}{T-t} dt \left[\frac{\sum_k \sigma_k p_{ikt}}{\sum_k p_{ikt}} \frac{\sum_i \sum_k p_{ikt}}{\sum_i \sum_k p_{ikt}} \frac{\xi_t x_i}{T-t} - \frac{\sum_i \sum_k x_i \sigma_k p_{ikt}}{\sum_i \sum_k p_{ikt}} \frac{\xi_t}{T-t} \right. \\ &\left. + \left(\frac{\sum_i \sum_k x_i \sigma_k p_{ikt}}{\sum_i \sum_k p_{ikt}} \right)^2 \frac{T}{T-t} - \frac{\sum_k \sigma_k p_{ikt}}{\sum_k p_{ikt}} \frac{\sum_i \sum_k p_{ikt}}{\sum_i \sum_k p_{ikt}} x_i \frac{\sum_i \sum_k x_i \sigma_k p_{ikt}}{\sum_i \sum_k p_{ikt}} \frac{T}{T-t} \right] \\ &= \frac{T}{T-t} \left[d\xi_t (\mathbf{P} \mathbf{S} \cdot \pi_{it}^{-1} x_i - \mathbf{X}^T \mathbf{P} \mathbf{S}) + dt \frac{1}{T-t} (\mathbf{P} \mathbf{S} \cdot \pi_{it}^{-1} x_i \xi_t - \mathbf{X}^T \mathbf{P} \mathbf{S} \cdot \xi_t \right. \\ &\quad \left. + (\mathbf{X}^T \mathbf{P} \mathbf{S})^2 T - \mathbf{P} \mathbf{S} \cdot \pi_{it}^{-1} x_i T \cdot \mathbf{X}^T \mathbf{P} \mathbf{S}) \right] \\ &= \frac{T}{T-t} [\mathbf{P} \mathbf{S} \cdot \pi_{it}^{-1} x_i - \mathbf{X}^T \mathbf{P} \mathbf{S}] dW_t, \end{aligned} \quad (2.12)$$

where

$$dW_t = d\xi_t + \frac{\xi_t}{T-t} dt - \frac{T}{T-t} \mathbf{X}^T \mathbf{P} \mathbf{S} dt. \quad (2.13)$$

The dynamics of the conditional probability, given by (2.12), describes how the posterior probability evolves, and this in turn determines the bond price dynamic. We shall now

show that the process $\{W_t\}_{0 \leq t \leq T}$ is an $\{\mathcal{F}_t^\xi\}$ -martingale. For $t \leq u$:

$$\begin{aligned} \mathbb{E}[W_u \mid \mathcal{F}_t^\xi] &= W_t + \mathbb{E}[(\xi_u - \xi_t) \mid \xi_t] + \mathbb{E}\left[\int_t^u \frac{\xi_s}{T-s} ds \mid \xi_t\right] - T \mathbb{E}\left[\int_t^u \frac{\mathbf{X}^T \mathbf{P} \mathbf{S}}{T-s} ds \mid \xi_t\right] \\ &= W_t + \mathbb{E}[\bar{\sigma} X_T u + \beta_{uT} \mid \xi_t] - \mathbb{E}[\bar{\sigma} X_T t + \beta_{tT} \mid \xi_t] + \mathbb{E}[\bar{\sigma} X_T \mid \xi_t] \int_t^u \frac{s}{T-s} ds \\ &\quad + \mathbb{E}\left[\int_t^u \frac{\beta_{sT}}{T-s} ds \mid \xi_t\right] - \mathbb{E}[\mathbf{X}^T \mathbf{P} \mathbf{S} \mid \xi_t] \int_t^u \frac{T}{T-s} ds. \end{aligned} \quad (2.14)$$

The quantity $\mathbf{X}^T \mathbf{P} \mathbf{S}$ is ξ_t measurable hence we can write

$$\mathbb{E}[\mathbf{X}^T \mathbf{P} \mathbf{S} \mid \xi_t] = \mathbf{X}^T \mathbf{P} \mathbf{S}.$$

On the other hand, the term $\mathbf{X}^T \mathbf{P} \mathbf{S}$ can be interpreted as the *a posteriori* expectation of X_T and $\bar{\sigma}$,

$$\mathbf{X}^T \mathbf{P} \mathbf{S} = \mathbb{E}[\bar{\sigma} X_T \mid \xi_t].$$

Hence all the terms with X_T in (2.14) cancel, and we have

$$\mathbb{E}[W_u \mid \mathcal{F}_t^\xi] = W_t + \mathbb{E}[\beta_{uT} \mid \xi_t] - \mathbb{E}[\beta_{tT} \mid \xi_t] + \int_t^u \frac{1}{T-s} \mathbb{E}[\beta_{sT} \mid \xi_t] ds. \quad (2.15)$$

Using the tower property of conditional expectation and the independence of $\{\beta_{iT}\}$ and σX_{iT} , we deduce that

$$\mathbb{E}[\beta_{uT} \mid \xi_t] = \mathbb{E}[\mathbb{E}[\beta_{uT} \mid \beta_{iT}] \mid \xi_t].$$

Also using the relation that

$$\mathbb{E}\left[\frac{\beta_{uT}}{T-u} - \frac{\beta_{tT}}{T-t} \mid \beta_{iT}\right] = 0,$$

we conclude

$$\mathbb{E}[\beta_{uT} \mid \xi_t] = \mathbb{E}[\mathbb{E}[\beta_{uT} \mid \beta_{iT}] \mid \xi_t] = \frac{T-u}{T-t} \mathbb{E}[\beta_{iT} \mid \xi_t].$$

Substituting this expression into (2.15) we obtain the martingale relation

$$\mathbb{E}[W_u \mid \mathcal{F}_t^\xi] = W_t.$$

In addition, we have

$$(dW_t)^2 = (d\xi_t)^2 = (d\beta_{iT})^2 = dt.$$

It follows, by virtue of the Lévy criterion, that $\{W_t\}$ is an $\{\mathcal{F}_t\}$ -Brownian motion.

The process $\{W_t\}$ is indeed the process that “drives” the asset price dynamics. It depends in a highly nontrivial way on both the signal component of the information flow, as well as on the noise term. We are now in a position to assess the stochastic dynamics of the Bond price. Since

$$B_{tT} = P_{tT} \sum_i x_i \pi_{it},$$

we have

$$\begin{aligned} dB_{tT} &= r_t B_{tT} dt + P_{tT} \frac{T}{T-t} \left[\sum_i x_i^2 \cdot \mathbf{P}\mathbf{S} - \mathbf{X}^T \mathbf{P}\mathbf{S} \cdot \sum_i x_i \pi_{it} \right] dW_t \\ &= r_t B_{tT} dt + P_{tT} \frac{T}{T-t} \left[\mathbb{E}[X_T^2 \bar{\sigma} \mid \xi_t] - \mathbb{E}[X_T \mid \xi_t] \cdot \mathbb{E}[X_T \bar{\sigma} \mid \xi_t] \right] dW_t \\ &= r_t B_{tT} dt + P_{tT} \frac{T}{T-t} V_{tT}^\sigma dW_t. \end{aligned}$$

We find that the absolute volatility of the bond price process is controlled by the covariance of X_T and $X_T \bar{\sigma}$:

$$V_{tT}^\sigma = \mathbb{E}[X_T^2 \bar{\sigma} \mid \xi_t] - \mathbb{E}[X_T \mid \xi_t] \cdot \mathbb{E}[X_T \bar{\sigma} \mid \xi_t].$$

In the case that σ and X_T are independent we have:

$$\begin{aligned} V_{tT}^\sigma &= \mathbb{E}[\bar{\sigma} \mid \xi_t] (\mathbb{E}[X_T^2 \mid \xi_t] - \mathbb{E}^2[X_T \mid \xi_t]) \\ &= \mathbb{E}[\bar{\sigma} \mid \xi_t] V_{tT}, \end{aligned}$$

where V_{tT} is the conditional variance of X_T given ξ_t . We see that if σ is constant then our model reduces to the original bond pricing formula in the BHM framework [15].

2.8 Volatility dynamics

Let us now work out the “Vol-of-Vol” for the bond price to examine the rate of decay of V_{tT}^σ . This is of interest in connection with the pricing of options on realised volatility or variance swaps. For notational simplicity, we write $\mathbb{E}_t[\cdot] \equiv \mathbb{E}[\cdot \mid \xi_t]$. Recall that V_{tT}^σ is defined by the covariance of X_T and $X_T \bar{\sigma}$. Let us express this in terms of p_{ikt} :

$$V_{tT}^\sigma = \frac{\sum_i \sum_k x_i^2 \sigma_k p_{ikt}}{\sum_i \sum_k p_{ikt}} - \frac{\sum_i \sum_k x_i p_{ikt}}{\sum_i \sum_k p_{ikt}} \frac{\sum_i \sum_k x_i \sigma_k p_{ikt}}{\sum_i \sum_k p_{ikt}},$$

and define the joint conditional probability of X_T and $\bar{\sigma}$ by

$$\pi_{ikt}^\sigma \equiv \frac{p_{ikt}}{\sum_l \sum_m p_{lmt}}. \quad (2.16)$$

Following a line of calculation similar to the one performed in the previous section, we deduce that

$$d\pi_{ikt}^\sigma = \frac{T}{T-t} (x_i \bar{\sigma} - \mathbb{E}_t[X_T \bar{\sigma}]) \pi_{ikt}^\sigma dW_t, \quad (2.17)$$

where dW_t is defined as before. This equation governs the dynamics of the joint probability for X_T and $\bar{\sigma}$. The conditional covariance process can be rewritten as:

$$V_{tT}^\sigma = \sum_i \sum_k x_i^2 \sigma_k \pi_{ikt}^\sigma - X_{tT} \sum_i \sum_k x_i \sigma_k \pi_{ikt}^\sigma,$$

where $X_{tT} \equiv \mathbb{E}[X_T | \xi_t]$ is the conditional expectation of X_T . Applying Ito's lemma, we obtain

$$dV_{tT}^\sigma = \sum_i \sum_k x_i \sigma_k [-\pi_{ikt}^\sigma dX_{tT} + x_i d\pi_{ikt}^\sigma - X_{tT} d\pi_{ikt}^\sigma - (dX_{tT})(d\pi_{ikt}^\sigma)].$$

On the other hand, for the dynamics of $\{X_{tT}\}$ we have

$$dX_{tT} = \frac{T}{T-t} V_{tT}^\sigma dW_t.$$

Combining these, we get

$$dV_{tT}^\sigma = -\left(\frac{T}{T-t}\right)^2 V_{tT}^\sigma V_{tT}^{X_T \sigma} dt + \frac{T}{T-t} \kappa_{tT}^\sigma dW_t, \quad (2.18)$$

where

$$V_{tT}^{X_T \sigma} \equiv \mathbb{E}_t[X_T^2 \bar{\sigma}^2] - (\mathbb{E}_t[X_T \bar{\sigma}])^2$$

is the conditional variance of the random variable $X_T \bar{\sigma}$, and

$$\kappa_{tT}^\sigma \equiv -\mathbb{E}_t[X_T](V_{tT}^\sigma + V_{tT}^{X_T \sigma}) + \text{Cov}_t(X_T \bar{\sigma}, X_T^2 \bar{\sigma}),$$

is the conditional skewness of the terminal payoff. The Vol-of-Vol, $\{\Sigma_{tT}^{(2)}\}$, is thus given by

$$\Sigma_{tT}^{(2)} = \left(\frac{T}{T-t}\right)^2 P_{tT} \kappa_{tT}^\sigma.$$

2.9 Information-flow rate process

It is worth noting that the dynamics of the conditional expectation of the information flow rate, $\bar{\sigma}$, can be read off from the foregoing analysis. This follows from the fact that we have obtained the joint density for the random variables X_T and $\bar{\sigma}$, as characterised by the probability matrix $\mathbf{P}$; thus by reduction to the marginal probability we obtain the conditional density for $\bar{\sigma}$. It should be evident from the structure of the information process

$$\xi_t = \bar{\sigma} X_T t + \beta_{tT} \quad (2.19)$$

that the roles of X_T and $\bar{\sigma}$ are in fact virtually interchangeable, except that while X_T is necessarily $\{\mathcal{F}_T^\xi\}$ -measurable, $\bar{\sigma}$ need not be so. If we let $\{\pi_{it}^X\}$ and $\{\pi_{kt}^{\bar{\sigma}}\}$ denote, respectively, the two marginal probabilities for X_T and $\bar{\sigma}$, then the measurability properties imply that $\lim_{t \rightarrow T} \pi_{it}^X = \delta_{ij}$ for some index j corresponding to the realised value x_j of the cash flow X_T , while in general

$$\lim_{t \rightarrow T} \pi_{it}^{\bar{\sigma}} \neq \delta_{kl} \quad (2.20)$$

for any index l .

In any event, by summing the rows of the matrix $\mathbf{P}$ we obtain the marginal probability:

$$\pi_{kt}^{\bar{\sigma}} = \frac{\sum_{i=1}^n p_i q_k \exp \left[\frac{T}{T-t} \left(x_i \sigma_k \xi_t - \frac{1}{2} x_i^2 \sigma_k^2 t \right) \right]}{\sum_{i=1}^n \sum_{k=1}^m p_i q_k \exp \left[\frac{T}{T-t} \left(x_i \sigma_k \xi_t - \frac{1}{2} x_i^2 \sigma_k^2 t \right) \right]}.$$

It follows that the conditional expectation $\bar{\sigma}_t = \mathbb{E}[\bar{\sigma} \mid \xi_t]$ of the flow-rate parameter is given by

$$\bar{\sigma}_t = \frac{\sum_{k=1}^m \sum_{i=1}^n \sigma_k p_i q_k \exp \left[\frac{T}{T-t} \left(x_i \sigma_k \xi_t - \frac{1}{2} x_i^2 \sigma_k^2 t \right) \right]}{\sum_{i=1}^n \sum_{k=1}^m p_i q_k \exp \left[\frac{T}{T-t} \left(x_i \sigma_k \xi_t - \frac{1}{2} x_i^2 \sigma_k^2 t \right) \right]}.$$

2.10 Analysis of path behaviour

In this section we analyse the path behaviour of a digital bond under the “uncertain information” model through simulation studies. This provides us with a better and more intuitive understanding of the characteristics of the model under study. In figures 2.2 and 2.3 we have shown simulations of sample paths of the bond price processes, the corresponding averaged (over five sample paths) volatility process, and the averaged (again, over five sample paths) vol-of-vol process. In these plots, we have set $r = 0\%$ for simplicity; and the maturity of the digital bond is $T = 1$ year. The information flow rate in figure 2.2 is chosen

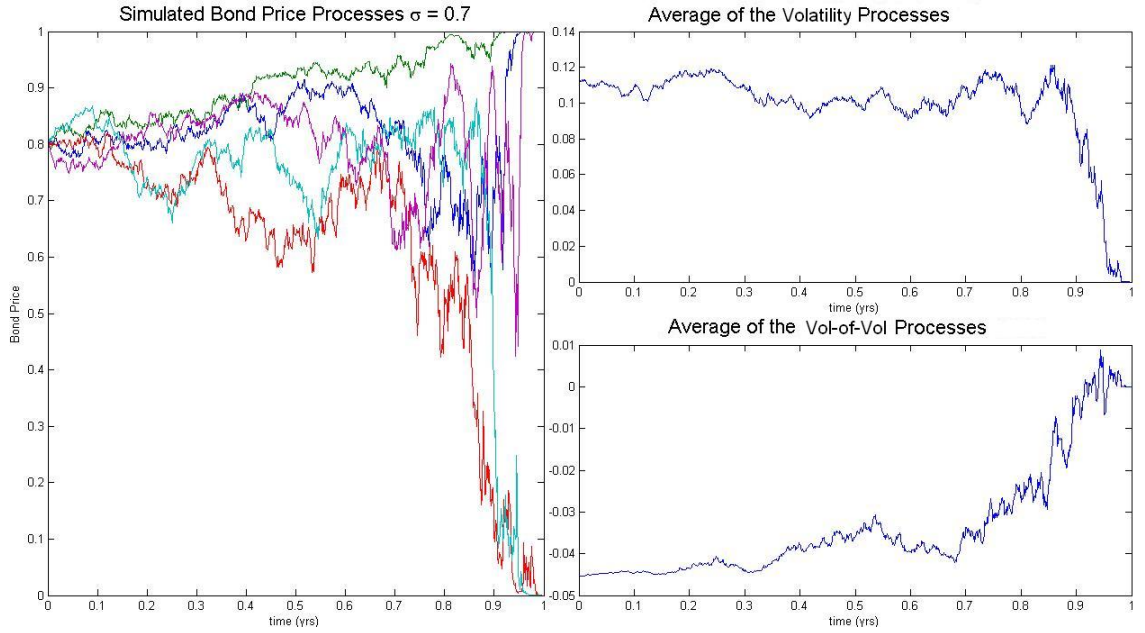


Figure 2.2: Simulated bond price (left), volatility (top right) and Vol-of-Vol (bottom right) processes in uncertain information model with $\bar{\sigma} = \{0.6, 0.8\}$, $q_k = \{0.5, 0.5\}$, $X_T = \{0, 1\}$, $p_i = \{0.2, 0.8\}$, $r = 0\%$ and $T = 1$.

to be a binary random variable taking values $\{0.6, 0.8\}$ with an equal *a priori* probability; whereas in figure 2.3, $\bar{\sigma}$ takes the values $\{0.4, 1.0\}$ also with an equal *a priori* probability. Hence in both cases we have $\mathbb{E}[\bar{\sigma}] = 0.7$, but they have different standard deviations.

One interesting observation to be drawn from the present simulation studies, as compared to the original BHM model, is that the degree of variation of the sample paths, i.e. the variability, is smaller than those observed in the BHM model with the same σ value. For example, although one of the sample paths in figure 2.3 has the value $\sigma = 1$, which in the BHM model would have caused the paths to reach their terminal values at about four-fifth of the way, such a large variation is not present in our uncertain information model. This is because the conditional expectation of X_T involves products of the *a priori* probability $\{q_k\}$, and this “dampens” the overall variability. If we set one of the q_k ’s equal to zero, which reduces the model to the original BHM, then the damping effect disappears. This behaviour is plausible because market participants are uncertain about the true information flow rate; their knowledge of the information is made additionally fuzzy by the uncertainty in $\bar{\sigma}$, and hence it generally takes a longer time to discover the true terminal cash flow. This uncertainty will, consequently, reduce the overall volatility for the majority of the duration of the bond. The reduction in volatility becomes apparent when we look at the average of the processes.

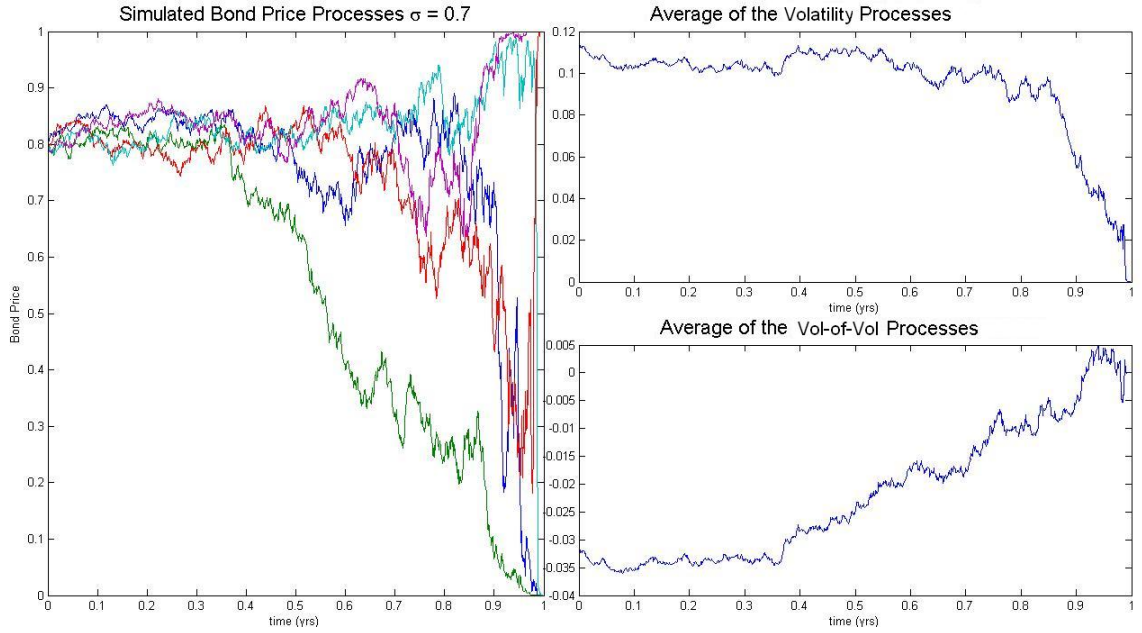


Figure 2.3: Simulated bond price (left), volatility (top right) and Vol-of-Vol (bottom right) processes in uncertain information model with $\sigma = \{0.4, 1\}$, $q_k = \{0.5, 0.5\}$, $X_T = \{0, 1\}$, $p_i = \{0.2, 0.8\}$, $r = 0\%$ and $T = 1$.

The volatility of the processes in some sense reflects the “learning rate” of the terminal value of the bond price (or, equivalently, the value of X_T). This learning rate is somewhat slowed down in the uncertain information model, hence resulting in the dampening of the volatility. If the standard deviation of $\bar{\sigma}$ is very small, then the volatility process almost matches that of the BHM model. However, as the standard deviation of $\bar{\sigma}$ is made wider so that the true information becomes less clear, this leads to a reduction in the price volatility. Nevertheless, as time passes, the value of X_T must eventually be revealed, and this leads to a subtle behaviour in the averaged volatility.

In view of this, we have investigated further the behaviour of the volatility processes. Because in our example here the random variable $\bar{\sigma}$ is chosen such that it is not measurable for any $t \in [0, T]$, we can expect some degrees of uncertainty to sustain until the very last moment. Indeed, the behaviour of the price process in the uncertain information model is not entirely counterintuitive; we expect, for example, that the price process behaves like a mixture of the BHM with the probabilities $\{q_k\}$ denoting the weights of the possible scenarios of σ_k . That this intuition is more or less correct, albeit there are subtle details, is illustrated in figure 2.4, in which the mean volatility $\bar{\Sigma}_{tT} = \mathbb{E}[\Sigma_{tT}]$ associated with various distributions for $\bar{\sigma}$ are compared. The two solid lines represent $\bar{\Sigma}_{tT}$ corresponding to the original BHM model, with the top line at $t = 0$ taking the value $\sigma = 0.9$ and the bottom line at $t = 0$ taking the value $\sigma = 0.5$. The dashed lines are mean volatilities generated

by our uncertain information model. They take values $\sigma_k = \{0.5, 0.9\}$ with the following probabilities: from the top at $t = 0$, $q_k = \{0, 1\}$; $q_k = \{0.1, 0.9\}$; $q_k = \{0.2, 0.8\}$; $q_k = \{0.3, 0.7\}$, and so on, until $q_k = \{1, 0\}$ for the bottom dashed line that coincides with the solid line.

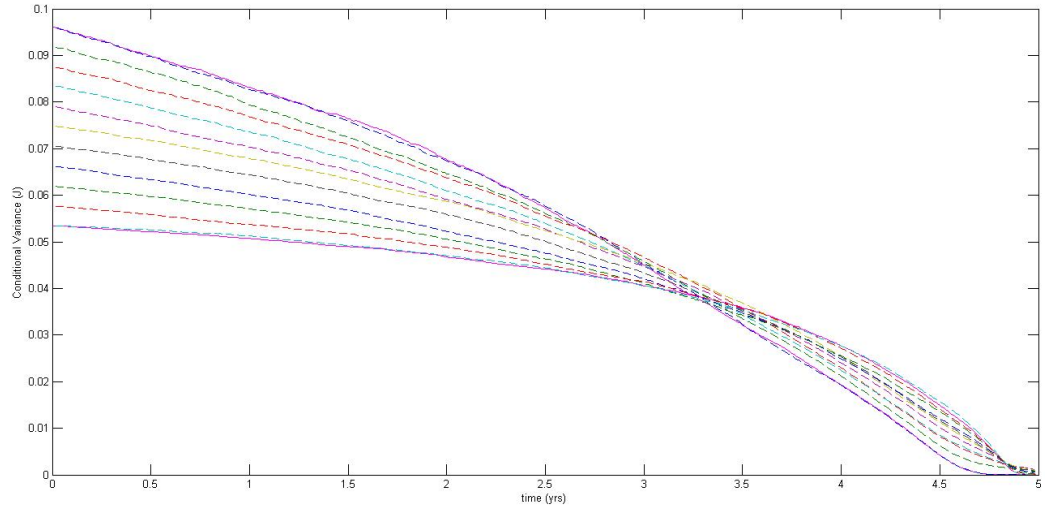


Figure 2.4: Each line represents the average of 5,000 sample paths of volatility processes. The two solid lines are that of the original BHM model with the top and bottom line taking $\sigma = 0.9$ and $\sigma = 0.5$ respectively, $X_T = \{0, 1\}$, $p_i = \{0.2, 0.8\}$, $r = 0\%$ and $T = 5$. The dashed lines are mean volatilities of the uncertain information model taking values, from top to bottom respectively, $\sigma_k = \{0.5, 0.9\}$ with $q_k = \{0, 1\}$, $q_k = \{0.1, 0.9\}$, $q_k = \{0.2, 0.8\}$, $\dots$, and $q_k = \{1, 0\}$.

The result reveals the following features: First, we see that when $\sigma = 0.5$ the average volatility starts with a lower value as compared to that of $\sigma = 0.9$; but eventually overrides the latter because information is revealed towards the end, and hence causing higher volatility. Second, we see that as we shift the mean of σ from 0.9 to 0.5, the mean volatility shifts accordingly. The graph shows that the mean volatility of the random σ model behaves like a weighted sum of the BHM mean volatilities, with weights given by the *a priori* probability $\{q_k\}$. In this regard, the random information model studied here can be viewed as a kind of BHM mixture model.

We also draw attention to the limiting behaviour of the mean volatility as $t \rightarrow T$. In figure 2.5 we see clearly that the mean volatility of the random information model does not reach zero when t is very close to T , even though that of the BHM model has. This indicates that there are “last minute surprises” because information regarding the terminal value is still uncertain, due to the additional uncertainty in $\bar{\sigma}$. This is the essence of the random information flow rate model, as it truly describes the uncertainty in the market up until the

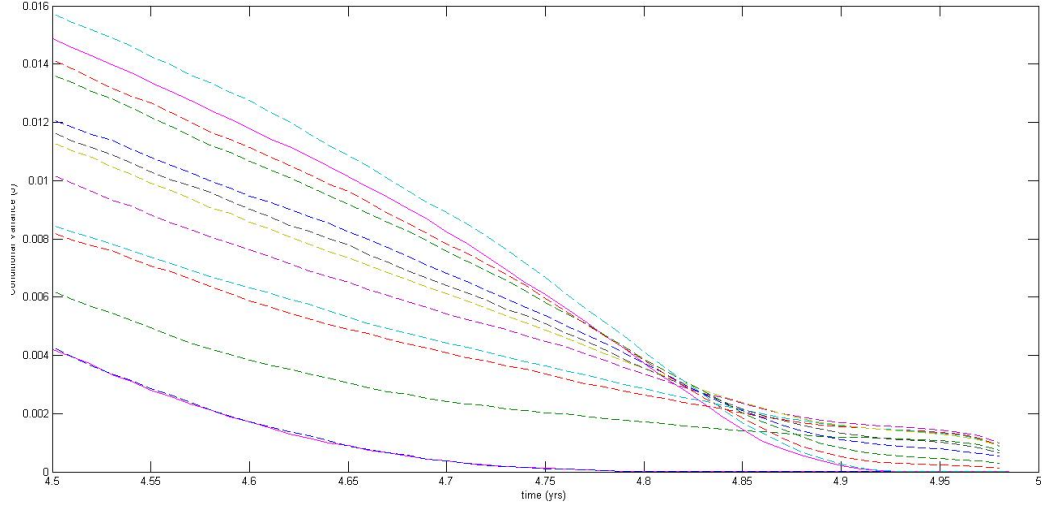


Figure 2.5: Magnification of the averages of volatility processes. Each line represents the average of 5,000 paths of volatility processes. The solid lines are that of the original BHM model with the top and bottom line taking $\sigma = 0.9$ and $\sigma = 0.5$, respectively, $X_T = \{0, 1\}$, $p_i = \{0.2, 0.8\}$, $r = 0\%$ and $T = 5$. The dashed lines are mean volatilities of the uncertain information model taking the values, from top to bottom respectively, $\sigma_k = \{0.5, 0.9\}$ with $q_k = \{0, 1\}$, $q_k = \{0.1, 0.9\}$, $q_k = \{0.2, 0.8\}$, $\dots$, and $q_k = \{1, 0\}$.

last moment. Note, however, that values of mean volatility above $t \gtrsim 0.996T$ were not available numerically, due to the existence of a cancellation of large numbers that is not recognised in the numerical code developed for our simulation studies.

We remark that from the expression (2.10) of the dynamics of the conditional probability process, $d\pi_{it}$ has an overall scaling term $T/(T-t)$ which diverges as $t \rightarrow T$. While the terms multiplying to this approach zero faster than $T-t$, making the expression well defined analytically, this divergence causes an issue in the simulation studies as one discretise the time steps. We observe that for sufficiently large values of σ_k , i.e. $\sigma_k > 0.7$, the average path can have a slight hump when t is very close to the maturity date, even though the terminal value stays unchanged.

2.11 Mutual information analysis

An alternative way of investigating properties of the price process is to study the behaviour of mutual information between B_{tT} and X_T , or, equivalently, between ξ_t and X_T . This is the quantity that measures the amount of information contained in the asset price about the value of the impending cash flow (cf. [24, 38] for further detail on the concept of mutual

information), and thus represents how much the market has learned about the value of X_T [13]. The mutual information is obtained by determining the expression:

$$J(\xi_t, X_T) = \sum_i^n \int_{-\infty}^{\infty} \rho_{\xi X}(x, i) \ln \left(\frac{\rho_{\xi X}(x, i)}{\rho_{\xi}(x) \rho_X(i)} \right) dx,$$

where

$$\rho_{\xi X}(x, i) = \frac{d}{dx} \mathbb{Q}[(\xi_t < x) \cap (X_T = x_i)]$$

is the joint density function of the random variables (ξ_t, X_T) , and ρ_{ξ}, ρ_X are the respective marginal probabilities. By the independence of X_T and σ , we find that

$$\mathbb{Q}[(\xi_t < x) \cap (X_T = x_i)] = \sum_k^m \mathbb{Q}(\xi_t < x \mid X_T = x_i, \sigma = \sigma_k) \mathbb{Q}(X_T = x_i) \mathbb{Q}(\sigma = \sigma_k),$$

from which it follows that

$$\rho_{\xi X}(x, i) = \sum_k^m q_k p_i \frac{1}{\sqrt{2\pi t(T-t)/T}} \exp \left(-\frac{1}{2} \frac{(x - \sigma_k x_i t)^2}{t(T-t)/T} \right)$$

since conditional on $X_T = x_i$ and $\sigma = \sigma_k$, the random variable ξ_t is normally distributed with mean $\sigma_k x_i t$ and variance $t(T-t)/T$.

An alternative way of deriving the mutual information is to use the formula [13]

$$J(\xi_t, X_T) = H_0 - \mathbb{E}[H_t],$$

where the Shannon-Wiener entropy $\{H_t\}$ is defined by the expression

$$H_t = - \sum_{i=1}^n \pi_{it} \ln \pi_{it}.$$

2.11.1 Limiting property of mutual information

The entropy process $\{H_t\}_{0 \leq t < T}$ has the property that $\lim_{t \rightarrow T} H_t = 0$. This follows from the fact that the conditional probability process $\{\pi_{it}\}_{0 \leq t < T}$ has the limiting behaviour

$$\lim_{t \rightarrow T} \pi_{it}(\omega) = \mathbb{1}\{X_T(\omega) = x_i\}$$

for $i = 1, \dots, n$. To see this, suppose that for some $\omega \in \Omega$ we have $X_T(\omega) = x_a$ and $\sigma(\omega) = \sigma_b$ for some a, b . For this realisation the information process is

$$\xi_t = \sigma_b t x_a + \beta_{tT}.$$

Substituting this expression of ξ_t into the expression for π_{it} , and dividing the denominator and the numerator by the exponential factor containing x_a and σ_b , we deduce that

$$\pi_{at} = \frac{p_a \left(q_b + \sum_{k \neq b} q_k \exp \left[\frac{T}{T-t} \left(x_a (\sigma_k - \sigma_b) \beta_{tT} - \frac{1}{2} x_a^2 (\sigma_k - \sigma_b)^2 t \right) \right] \right)}{p_a q_b + \sum_{i \neq a} \sum_{k \neq b} p_i q_k \exp \left[\frac{T}{T-t} \left((x_i \sigma_k - x_a \sigma_b) \beta_{tT} - \frac{1}{2} (x_i^2 \sigma_j^2 - x_a^2 \sigma_b^2) t \right) \right]}.$$

As $t \rightarrow T$ all of the terms in the sums vanish. Therefore, $\lim_{t \rightarrow T} \pi_{at} = 1$ and furthermore, since

$$\sum_i \pi_{it} = 1$$

for all t , we must have

$$\lim_{t \rightarrow T} \pi_{it} = 0$$

for $i \neq a$. Finally, since

$$H_t = -\ln \prod_{i=1}^n \pi_{it}^{\pi_{it}},$$

we deduce that

$$\lim_{t \rightarrow T} H_t = 0.$$

In figure 2.6 we provide a graphical illustration of the behaviour of the mutual information. The idea here is to relate this with the behaviour of the average volatility that we observed in the previous section. We have used exactly the same parameter values to make the comparison more transparent. Hence, the mutual information for the BHM model with $\sigma = 0.9$ and $\sigma = 0.5$ form the upper and lower bounds, respectively, at early times in figure 2.6. The mutual information, or, the learning curves that lie within these bounds at early times are those under the uncertain information model, taking the values $\sigma_k = \{0.5, 0.9\}$ with probabilities $q_k = \{0, 1\}$, $q_k = \{0.1, 0.9\}$, $\dots$, $q_k = \{1, 0\}$, respectively, from top to bottom.

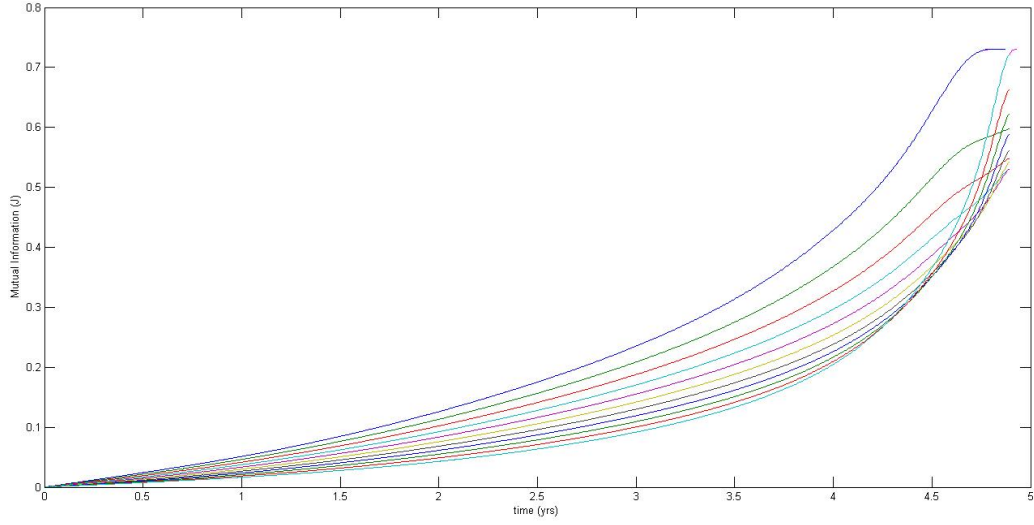


Figure 2.6: Mutual information under the random σ model. The vertical axis represents the mutual information and the horizontal axis represents the time elapsed. We let $\sigma_k = \{0.5, 0.9\}$ with probabilities, from top to bottom at early times, $q_k = \{0, 1\}$, $q_k = \{0.1, 0.9\}, \dots, q_k = \{1, 0\}$ respectively; $X_T = \{0, 1\}$, $p_i = \{0.2, 0.8\}$, and $T = 5$.

The results shown in the figure match exactly the findings of the mean volatility analysis. The gap in the plots close to $t = T$ is again due to numerical limitations. Observe the crossover pattern seen here, as we change the distribution of $\bar{\sigma}$. The processes with a low mutual information at early times cross the higher ones at later times because the volatility is greater towards the end so as to “catch up” with the learning rate. Further, there are still uncertainties left until the very last moment.

Chapter 3

Options on Credit-risky Bonds with Random Information Flows

In this chapter we examine the pricing of options on credit-risky bonds. In particular, we work out in semi-closed form, up to numerical root-finding, pricing formulae for European-style call options on a bond. The results are expressed in terms of the BHM implied volatility parameter, exhibiting familiar smile characteristics. What this means is that the elementary stochastic volatility model introduced in the previous chapter has not only the analytical tractability but also sufficient flexibility to fit implied volatility surfaces. We shall also determine the sensitivity of the option price with respect to the initial bond price. These results show that our random-information model entails a significant degree of practicality.

3.1 Option pricing formula

Following a procedure similar to the one provided in [17], we consider the valuation of a European call option on the asset described in the previous chapter, exercisable at a fixed time $t > 0$ on a defaultable bond maturing at time $T > t$ with strike price K . The value of the call option at time zero is

$$\begin{aligned} C_0 &= P_{0t} \mathbb{E}[(B_{tT} - K)^+] \\ &= P_{0t} \mathbb{E} \left[\left(\sum_i P_{tT} \pi_{it} x_i - K \right)^+ \right], \end{aligned} \tag{3.1}$$

where the conditional probability π_{it} is as given in (2.10). Let us write

$$\begin{aligned} p_{ikt} &= p_i q_k \exp \left[\frac{T}{T-t} \left(x_i \sigma_k \xi_t - \frac{1}{2} x_i^2 \sigma_k^2 t \right) \right], \\ \pi_{it} &= \frac{\sum_k p_{ikt}}{\Phi_t}, \\ \Phi_t &= \sum_i \sum_k p_{ikt}. \end{aligned} \quad (3.2)$$

Then equation (3.1) can be rewritten as

$$C_0 = P_{0t} \mathbb{E} \left[\frac{1}{\Phi_t} \left(\sum_i \sum_k (P_{tT} x_i - K) p_{ikt} \right)^+ \right]. \quad (3.3)$$

Our objective now is to determine this expectation.

3.2 Measure change technique

In order to derive the formula for the option price we shall use a change of measure technique. In particular, the factor Φ_t^{-1} appearing in (3.3) will be used to effect a measure change. First we show that $\{\Phi_t^{-1}\}$, as defined in (3.2), can be expressed in the form of a unit-initialised exponential martingale. We then show that the Novikov condition is satisfied, and thus $\{\Phi_t^{-1}\}$ can be used as a density martingale to change the measure. Let us fix u such that $t \leq u < T$, and define a process $\{\Phi_t\}_{0 \leq t \leq u}$ by (3.2). Recall the dynamics of p_{ikt} :

$$\begin{aligned} \frac{dp_{ikt}}{p_{ikt}} &= \left[\frac{T}{(T-t)^2} \left(\sigma_k x_i \xi_t - \frac{1}{2} \sigma_k^2 x_i^2 t \right) - \frac{1}{2} \frac{T}{T-t} \sigma_k^2 x_i^2 \right] dt \\ &\quad + \frac{T}{T-t} \sigma_k x_i d\xi_t + \frac{1}{2} \frac{T^2}{(T-t)^2} \sigma_k^2 x_i^2 (d\xi_t)^2 \\ &= \frac{T^2}{(T-t)^2} \sigma_k x_i \mathbb{E}[\bar{\sigma} X_T \mid \xi_t] dt + \frac{T}{T-t} \sigma_k x_i dW_t \end{aligned}$$

Summing over indices i and k , and making use of the relation

$$\frac{1}{\Phi_t} \sum_i \sum_k x_i \sigma_k p_{ikt} = \mathbb{E}[\bar{\sigma} X_T \mid \xi_t],$$

we obtain

$$d\Phi_t = \frac{T^2}{(T-t)^2} \mathbb{E}^2[\bar{\sigma} X_T | \xi_t] \Phi_t dt + \frac{T}{T-t} \mathbb{E}[\bar{\sigma} X_T | \xi_t] \Phi_t dW_t.$$

Applying Ito's quotient rule, we deduce that the dynamical equation satisfied by the random process $\{\Phi_t^{-1}\}$ is

$$d\Phi_t^{-1} = -\frac{T}{T-t} \mathbb{E}[\bar{\sigma} X_T | \xi_t] \Phi_t^{-1} dW_t.$$

The solution to this equation is given by

$$\Phi_t^{-1} = \exp \left(-\int_0^t \frac{T}{T-s} \mathbb{E}[\bar{\sigma} X_T | \xi_s] dW_s - \frac{1}{2} \int_0^t \frac{T^2}{(T-s)^2} \mathbb{E}^2[\bar{\sigma} X_T | \xi_s] ds \right). \quad (3.4)$$

Since $\mathbb{E}[\bar{\sigma} X_T | \xi_t]$ is bounded, we see that the Novikov condition

$$\mathbb{E} \left[\exp \left(-\frac{1}{2} \int_0^t \frac{T^2}{(T-s)^2} \mathbb{E}^2[\bar{\sigma} X_T | \xi_s] ds \right) \right] < \infty$$

is satisfied for any $t \leq u$. Hence $\{\Phi_t^{-1}\}_{0 \leq t \leq u}$ is a martingale. We also see from the normalisation condition that

$$\Phi_0^{-1} = \left(\sum_i \sum_k p_i q_k \right)^{-1} = 1, \quad (3.5)$$

and hence

$$\mathbb{E}[\Phi_t^{-1}] = 1.$$

Thus, the factor Φ_t^{-1} can be used to effect a change of measure. The option price becomes, under the new measure,

$$C_0 = P_{0T} \mathbb{E}^{\mathbb{B}_T} \left[\left(\sum_i \sum_k (P_{iT} - K) p_{ikt} \right)^+ \right]. \quad (3.6)$$

As in the original BHM model reviewed in chapter 1, we shall call this new measure the “bridge measure”, denoted $\mathbb{B}_T$, because under this measure the information process $\{\xi_t\}$ is a standard Brownian bridge. That is to say, the random variable ξ_t is Gaussian with mean 0 and variance $t(T-t)/T$ under the $\mathbb{B}_T$ measure. We note first some general remarks concerning this measure change. The Radon-Nikodym derivative is $d\mathbb{B}_T = \Phi_t^{-1} d\mathbb{Q}$; the random variables X_T and $\bar{\sigma}$ have the same probability laws with respect to $\mathbb{B}_T$ and $\mathbb{Q}$.

Furthermore, the conditional expectation of any integrable function $f(X_T, \bar{\sigma})$ under the bridge measure can be expressed as

$$\mathbb{E}^{\mathbb{B}_T}[f(X_T, \bar{\sigma}) \mid \mathcal{F}_t] = \frac{\mathbb{E}^{\mathbb{Q}}[\Phi_t^{-1} f(X_T, \bar{\sigma}) \mid \mathcal{F}_t]}{\mathbb{E}^{\mathbb{Q}}[\Phi_t^{-1} \mid \mathcal{F}_t]}.$$

In particular, the process $\{W_t^*\}_{0 \leq t \leq u}$ defined by

$$W_t^* = \int_0^t \frac{T}{T-s} \mathbb{E}[\bar{\sigma} X_T \mid \xi_s] ds + W_t \quad (3.7)$$

is a $\mathbb{B}_T$ -Brownian motion.

To verify that the process $\{\xi_t\}$ is a $\mathbb{B}_T$ -Brownian bridge, we begin by redefining the above $\{W_t\}$ process in the $\mathbb{Q}$ measure by the use of equation (2.13):

$$W_t = \xi_t + \int_0^t \frac{\xi_s}{T-s} ds - \int_0^t \frac{T}{T-s} \mathbb{E}[\bar{\sigma} X_T \mid \xi_s] ds.$$

Substituting this expression into equation (3.7) we get

$$W_t^* = \xi_t + \int_0^t \frac{\xi_s}{T-s} ds.$$

From this relation we obtain an SDE of the Ornstein-Uhlenbeck type:

$$d\xi_t = -\frac{\xi_t}{T-t} dt + dW_t^*.$$

The solution to the SDE above is given by

$$\xi_t = (T-t) \int_0^t \frac{1}{T-s} dW_s^*.$$

We see that under the $\mathbb{B}_T$ measure the process $\{W_s^*\}_{0 \leq s \leq t}$ is a Brownian motion with the Radon-Nikodym derivative given by (3.4), hence $\{\xi_s\}_{0 \leq s \leq t}$, defined above, is a standard Brownian bridge process, as claimed.

Returning to the problem of option pricing, let us begin by considering the case of a binary bond, whereby the cash flow takes the two possible values $\{x_0, x_1\}$. Then we have

$$C_0 = P_{0t} \mathbb{E}^{\mathbb{B}_T} \left[\left((P_{tT} x_1 - K) \sum_k p_{1kt} + (P_{tT} x_0 - K) \sum_k p_{0kt} \right)^+ \right]. \quad (3.8)$$

From equation (3.6), this expectation is nontrivial when $P_{tT}x_1 > K > P_{tT}x_0$. Otherwise, the option will be sure to expire either in the money or out of the money. In the nontrivial case, the option can expire either in the money or out of the money, depending on whether $\xi_t > \xi^*$ or $\xi_t < \xi^*$, where ξ^* is the unique critical value of ξ_t such that $B_tT = K$, or, equivalently, the unique solution to the equation

$$\frac{\sum_k q_k \exp(\frac{T}{T-t}(\sigma_k x_0 \xi^* - \frac{1}{2}\sigma_k^2 x_0^2 t))}{\sum_k q_k \exp(\frac{T}{T-t}(\sigma_k x_1 \xi^* - \frac{1}{2}\sigma_k^2 x_1^2 t))} = \frac{p_1(K - P_{tT}x_1)}{p_0(P_{tT}x_0 - K)}. \quad (3.9)$$

Note that in general ξ^* has no closed-form expression. However, the solution to (3.9) can be obtained by simple numerical root-finding methods. That the solution to (3.9) is unique (assuming that $\bar{\sigma}$ is a positive random variable so that $\sigma_k > 0$ for all k) can be seen from the fact that the bond price is an increasing function of ξ_t . To see this, let us define $B(t, \xi)$ by

$$B(t, \xi) = P_{tT} \frac{\sum_i \sum_{k=1}^m x_i p_i q_k \exp[\frac{T}{T-t}(x_i \sigma_k \xi - \frac{1}{2}x_i^2 \sigma_k^2 t)]}{\sum_{s=1}^n \sum_{r=1}^m p_s q_r \exp[\frac{T}{T-t}(x_s \sigma_r \xi - \frac{1}{2}x_s^2 \sigma_r^2 t)]}$$

so that the bond price is given by $B_{tT} = B(t, \xi_t)$. We then differentiate $B(t, \xi)$ in ξ to get

$$\begin{aligned} \left. \frac{\partial B(t, \xi)}{\partial \xi} \right|_{\xi=\xi_t} &= P_{tT} \frac{T}{T-t} \left(\frac{\sum_i \sum_k p_{ikt} x_i^2 \sigma_k}{\sum_i \sum_k p_{ikt}} - \frac{\sum_i \sum_k x_i p_{ikt} \sum_j \sum_l x_j \sigma_l p_{jlt}}{[\sum \sum p_{ikt}]^2} \right) \\ &= P_{tT} \frac{T}{T-t} (\mathbb{E}[X_T^2 \sigma \mid \xi_t] - \mathbb{E}[X_T \mid \xi_t] \mathbb{E}[X_T \sigma \mid \xi_t]) \\ &= P_{tT} \frac{T}{T-t} V_{tT}^\sigma \\ &\geq 0 \end{aligned}$$

and hence $B(t, \xi_t)$ is an increasing function of ξ_t . It follows that there exists a unique ξ_t^* that satisfies equation (3.9).

The problem of option pricing thus reduces to performing an elementary Gaussian integration. We now consider the case where X_T need not be a binary variable. The computation simplifies further if we introduce a standard normal variable Z according to

$$Z = \frac{\xi_t}{\sqrt{t(T-t)/T}}.$$

We write Z^* for the corresponding critical value. Then the option pricing formula is:

$$C_0 = P_{0t} \sum_k \sum_i q_k p_i (P_{tT}x_i - K) N(\sqrt{\tau} \sigma_k x_i - Z^*), \quad (3.10)$$

where $\tau = tT/(T - t)$ and $N(\cdot)$ denotes the standard cumulative normal density function.

Below are plots of call option prices against strike for various values of the option maturity t , for both the BHM and the random information models:

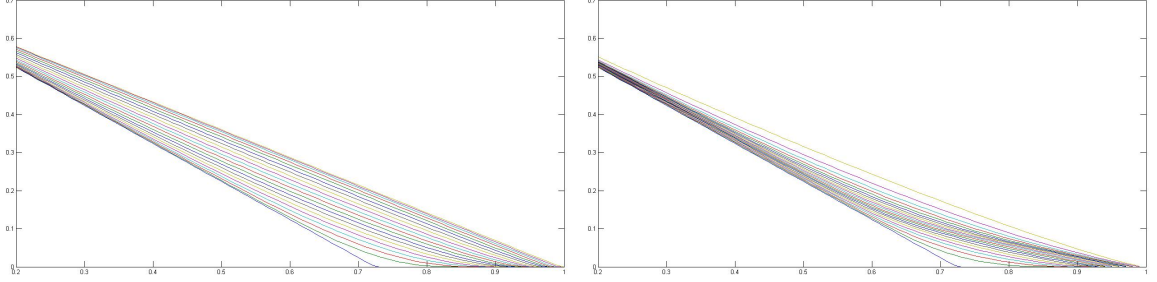


Figure 3.1: Call price against strike for different option maturities under the BHM (left) and uncertain information (right) models. We see that options with longer maturities are less expensive in the random information model, for any strike price. This is because the additional uncertainty in σ indeed reduces volatility in the initial stage of a long maturity option. As the options approach maturity the two models become the same.

3.3 Sensitivity analysis

Having obtained the expression for the option price, it is natural to examine the option sensitivity against the initial price of the underlying, i.e. the option delta. In this connection it is worth remarking that in the context of credit risk modelling, the option delta is usually difficult to determine because the models commonly used are driven by jump processes. The advantage of the BHM approach, on the other hand, is that it allows for the use of a Brownian-based filtration to generate the dynamics of a credit-risky bond. This means that the option delta can be determined explicitly, just as in the Black-Scholes model. Here we shall work out the expression for the delta

$$\Delta = \frac{\partial C_0}{\partial B_{0T}}.$$

in our extended model. To determine the option delta, let us first consider the case of an option on a digital bond as a starting point. By definition we have the following relations:

$$B_{0T} = P_{0T}(p_0x_0 + p_1x_1), \quad \text{and} \quad p_0 + p_1 = 1.$$

Substituting these into equation (3.10) we get

$$C_0 = P_{0T} \left[\frac{1}{x_1 - x_0} \left(\frac{B_{0T}}{P_{0T}} - x_0 \right) \sum_k q_k (P_{tT} x_1 - K) N(\sqrt{\tau} \sigma_k x_1 - Z^*) \right. \\ \left. + \frac{1}{x_1 - x_0} \left(x_1 - \frac{B_{0T}}{P_{0T}} \right) \sum_k q_k (P_{tT} x_0 - K) N(\sqrt{\tau} \sigma_k x_0 - Z^*) \right].$$

Upon differentiation with respect to B_{0T} , we obtain:

$$\Delta = \frac{\sum_k q_k (P_{tT} x_1 - K) N(d_{1k}) - \sum_k q_k (P_{tT} x_0 - K) N(d_{0k})}{P_{tT} (x_1 - x_0)},$$

where

$$d_{ik} = \sqrt{\tau} \sigma_k x_i - Z^* \quad \text{for } i = 1, 2.$$

We note that, apart from the option delta, we can consider other sensitivity indicators commonly characterised by the various “Greeks”. In addition, in our example we can determine the sensitivity of the option price against the degree of uncertainty in the magnitude of the volatility $\bar{\sigma}$. In particular, in the binary case where we have $q_0 = q$ and $q_1 = 1 - q$. Thus, letting

$$\nu = \sqrt{q\sigma_0^2 + (1-q)\sigma_1^2 - (q\sigma_0 + (1-q)\sigma_1)^2}$$

denote the standard deviation of $\bar{\sigma}$ the sensitivity of the call price against the “informational volatility”, ν can easily be calculated by use of the chain rule.

3.4 Stochastic volatility

As remarked earlier, the random information model considered here is an elementary example of a stochastic volatility model for option pricing, since the price process depends on the distribution of $\bar{\sigma}$. The advantage of this model is that we do not have to explicitly define the structure, i.e. the probability law of the volatility, since this can be calibrated from options data. But exactly how do we calibrate the distribution of $\bar{\sigma}$ in order to fix the price process?

We shall proceed as follows: First, note that the original BHM model is a “constant” volatility model (where we take the liberty to call the parameter σ *the volatility*). Thus, option prices given by the original BHM cannot match market prices across different values of

strike K , just as in the Black-Scholes model. Therefore, if we regard the BHM model as the benchmark model, then we expect to observe a volatility smile across different strike prices, if we calibrate the BHM for at-the-money option price. What this means is that we can quote the prices of bond options in terms of the BHM implied volatility parameter, just as in the equity market where the Black-Scholes implied volatility is used to quote prices.

Alternatively, in the random information model we can calibrate the BHM implied volatility surface by means of adjusting the parameters $\{q_k\}$ and $\{\sigma_k\}$. Numerical studies indicate that with a small number of parameters, a relatively wide range of volatility surfaces can be fitted reasonably well. To illustrate the idea, here we have taken specific values $\sigma_k = \{0.3, 2.7\}$ and $q_k = \{0.5, 0.5\}$ for the model parameters to compute the option prices across different strikes and different maturities, and expressed these prices in terms of the BHM implied volatility. The results are shown in figure 3.1 and figure 3.2, representing a characteristic “smile” surface where there are stronger skews for shorter maturities.

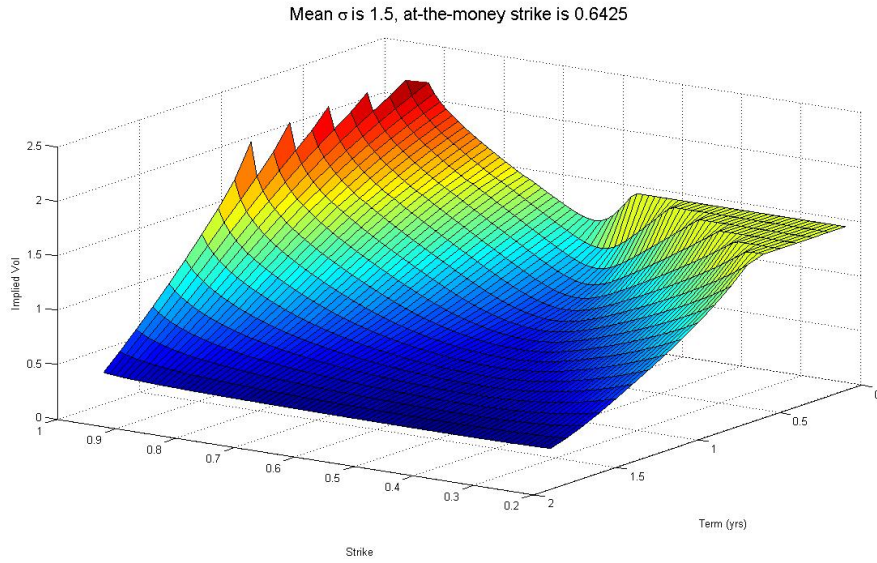


Figure 3.2: *The implied BHM volatility.* The implied volatility surface resulting from calibrating each strike and maturity of the random information model to the BHM model is shown. The parameters are set to be $X_T = \{0, 1\}$, $p_i = \{0.2, 0.8\}$, $\sigma_{\text{average}} = 1.5$, $\sigma_{\text{Rand}} = \{0.3, 2.7\}$, $q_k = \{0.5, 0.5\}$, $r = 0\%$ and $T \in [0, 2]$.

3.5 Remarks on the random information and the BHM mixture

We have noted that our random information model can be viewed as a mixture of BHM models. In the context of the Black-Scholes model, Renault and Touzi [48] or Brigo *et al.* [10], for instance, have carried out a similar analysis, where the Black-Scholes volatility pa-

parameter σ is taken to be a time-dependent random variable that is nevertheless independent of the underlying Brownian motion. By conditioning on the volatility path, the European call option is obtained as the expectation of the Black-Scholes call price with time-averaged volatility. In particular, when σ is time-independent, the result of [10, 48] for the extended Black-Scholes model is similar in nature to our result on the extended BHM model.

The randomisation of the Black-Scholes volatility parameter (the so-called log-normal mixture) in the literature is carried out essentially in an *ad hoc* manner, without any fundamental economic reason. Rather, it is justified on the practical grounds that it gives a better handling of calibration.

In contrast, in our model the randomisation arises from a more realistic analysis on the structure of the market information process; in particular, from the fundamental observation that the information flow rate is virtually always random and in many cases is not even observable. Hence, although the net effect is similar in both cases, our model is accompanied by a better justification, which in turn also gives a better justification for the lognormal mixture models. In particular, since the log-normal model can also be derived from the information-based framework [17, 42], we can now assign a deeper significance to the log-normal mixture models.

3.6 Information manipulation

We conclude by drawing attention to another interesting feature of the variable σ model. This concerns the notion of information manipulation, or, equivalently, a deliberate misrepresentation of information. The question that we are interested in here is the following: How does one model the manipulation of information in the information-based framework? One possible solution is given by the mis-specification of the information flow-rate σ .

The idea can be sketched as follows. Each market agent reveals information, expressed in the form of one of the information processes of (4.3). The impact of that agent's information on the market is then determined through formula (4.4). If an agent releases information that is based purely on speculation, then that information source is noise dominated, having a small value of the information flow rate. However, if that agent is trying to mislead the market, based on a reliable piece of information, then the nature of unreliability of that information is not the same as that arising from pure noise. Instead, this misinformation can be modelled by the fact that the agent provides an incorrect value for the information flow rate parameter. In this way, the market will estimate the fair price

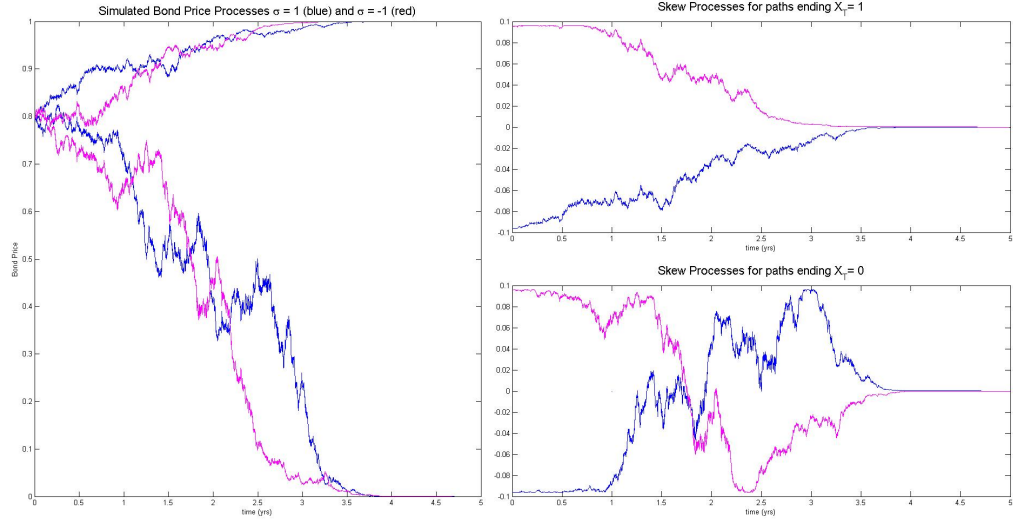


Figure 3.3: Left: Sample paths of the BHM, i.e. bond price against time elapsed, with $\sigma = 1$ (blue) and $\sigma = -1$ (red) using same Brownian bridges; and the corresponding skew processes associated with $X_T(\omega) = 1$ (Top right) and $X_T(\omega) = 1$ (Bottom right). Here we set $X_T = \{0, 1\}$, $p_i = \{0.2, 0.8\}$, $r = 0\%$ and $T = 5$.

of the asset by use of the pricing formula (4.5), but based on the incorrect value for the information flow rate parameter. As a consequence, the market price will be misled.

As an illustration of this behaviour, in figure 3.3 we show sample paths for the defaultable digital bond price process—in one case where the bond did not default, whereas in the other case where the bond defaulted. In each case, two sample paths are given; one corresponding to where the true market price ought to be if there is no misleading information, and one corresponding to where the realised price is, owing to the existence of a deliberate price manipulation.

In order to make the effect of price manipulation visually pronounced, here we have taken a slightly extreme case in which the true value of the information flow rate is $\sigma = +1$ (blue), whereas the “conjugate” price process is generated by the false belief that the flow rate is given by $\sigma = -1$ (red). We find in these examples the existence of a kind of anti-correlation between the “true” and “false” prices around their conditional means. This is illustrate more clearly in the skewness plot, also shown in figure 3.3.

One might enquire in which way a mis-specification of the information flow rate σ is realised in practice. In this connection it is worth remarking that there is an extended literature on price manipulation, often in the context of insider trading. One typical way of spreading a false information is by taking a trading position in a strategic manner (cf. [2, 20]).

For example, suppose that an informed trader has the information that the price of an asset

is likely to drop in near future. In this case, taking a short position amounts to effectively revealing the content of that information. Hence, by momentarily taking a long position before taking a short position, an informed trader can mislead the market. One can think of such a deliberate manoeuvre being represented abstractly in the form of one of the information processes in (4.3) taking an “incorrect” value of σ .

We see therefore that the information-based framework allows for a range of flexible extensions to model various scenarios that might occur in a given financial market. It would be of considerable interest, in particular, to develop further the information-based approach to model various form of price manipulations.

Chapter 4

Theory of Information Pricing

In the previous chapters we have shown that under an uncertain information-flow model we can nevertheless derive asset price process, as well as associated derivatives, in closed form. The way in which we have extended the original BHM model is motivated to some extent by “realities” in financial markets concerning the fuzziness of the information flow rate. If there is a high-quality information-extraction “device” (like a search engine), however, the value of the information flow rate can be enhanced by means of an automated speedy access to a large number of data sources.

This naturally leads to the question as regards the valuation of information itself! If an information provider (but not an agent in financial markets) can extract information that is useful for financial institutions, in which way should the provider fix its price? In this chapter we shall be regarding information as a kind of “perishable” tradable asset (since the lifetime of a useful information is typically short), and use an equilibrium argument as a basis for pricing information itself. In particular, we shall be expressing information in terms of its natural unit (the bit) and determine its price. We find what might be a surprising result: that the cost of information is not necessarily monotonically increasing in the quantity of information being supplied.

4.1 Introduction

The role and importance of information and its costing has long been recognised in the economics literature [31]. Nowadays, people often obtain information through internet search engines; but the very concept that one should search price information and utilise it to reduce expenditure, in the context of commodity markets, has been noted for at least

half a century [52], if not considerably longer.

Nevertheless, the use of information has never been as important in the past. Indeed, access to internet search engines—whose objectives are in information provision—constitutes an integrated part of the lives of hundreds of millions. Internet retailers likewise not only sell products but also provide a range of information via their recommendation engines, helping consumers find what they like but along the way also yielding additional profits. The quality of a recommendation engine, however, is difficult to assess if one cannot assign a value to the information it provides.

Rapid developments in communication technology have their upsides as well as downsides. On the one hand, access to targeted information has been made easy (for example, when one knows the relevant web address from which information can be obtained), while on the other hand a generic search for specific information is becoming harder due to the flooding of irrelevant information that constitutes noise. This is a phenomenon envisaged by Norbert Wiener, who used the second law of thermodynamics to explain the fact that in the long run one cannot overcome the impact of the noise that overwhelms useful information, and that this effect is in part due to the advancements of communication technology [55]. Wiener also recognised that this inevitable flow can locally be reversed by means of innovations. It is for this reason that internet search engines and other information providers are constantly on alert of innovative ideas for information extraction and provision. Evidently, enhancement of information extraction and provision, and implementation of innovation, cannot be realised without cost. As a consequence, access to valuable information will likewise entail some costs. The question then is: how do we price information?

There is a form of information that is regarded as particularly valuable—the kind of information that allows one to anticipate, at least with some statistical significance, what is to come, that is, to anticipate future trends. For example, suppose that a company has just released a new product. A search engine can access various internet sites and analyse texts (e.g., blog articles and tweets) to determine how satisfied consumers are with the product. If customer satisfaction is high, the search engine can anticipate that future sales will increase and hence the company will perform well. Of course, neither the data source nor the analysis can be perfect; nevertheless, if the methodology of information extraction is adequate, then the prediction of the search engine can be statistically significant. After a while, however, sales figures are released, by which time the value of this information diminishes.

The above example illustrates the idea of information diffusion. The search engine extracts information from in principle publicly available sources. Hence one need not be a large

search engine to obtain the same information at the same time. However, in practice most people do not have the resources to analyse in real time millions of articles available on the internet. Nevertheless, after a sufficient amount of time (which can be short or long), important information will be appreciated by a broad mass (and thus reflected in the share price, in the above example), diminishing the worth of information as providing predictive insights. At this stage, information has truly reached the public domain, but this is necessarily sometime after its initial appearance somewhere in “public”.

From the viewpoint of a search or recommendation engine, and given their resources, capabilities, and objectives, therefore, it makes sense to extract useful information before the public as a whole has had chance to do so, and sell it off to a third party who is in the position to utilise this information. We can thus think of information as a tradable asset, albeit that it will have an expiry date. In order to price information, however, we must cast these intuitive ideas in terms of precise mathematical language.

4.2 Heuristic illustration

To illustrate how information can be valued, let us consider a simple example. Suppose that we are interested in the price tomorrow of a stock, whose current price is s_0 . Our formulation of information pricing is applicable to a broader class of situations involving anticipation of uncertain future events. Nevertheless, we find that the example concerning information for predicting the future market very convenient for providing an intuitive understanding of the theory (cf. [12]). To keep the discussion simple, let us assume a single-period setup so that the price s_0 today will become the price s_1 tomorrow. The value of s_1 is of course unknown today, but we assume that it satisfies a known probability law $p(x)$. That is, we regard s_1 as a random variable defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, where $\mathbb{P}$ denotes the physical probability measure. A search engine now performs text mining to extract information relevant to determining the value of s_1 . Evidently, no search engine is capable of identifying the value of s_1 ; at best, it gathers noisy information in the form of *signal plus noise*. In the simplest setup, s_1 itself constitutes the signal, but a search engine can only extract the value of

$$\xi = s_1 + \epsilon, \tag{4.1}$$

where ϵ represents noise, independent of s_1 , with known density $f(x)$. In other words, a search engine can sample the value of ξ , which contains new information that is not already encoded in the valuation today of the stock, but not sufficient to determine the

value tomorrow of the stock. With the knowledge of ξ the prior density $p(x)$ of s_1 can thus be updated to the posterior density

$$\pi(x|\xi) = \frac{p(x)f(\xi - x)}{\int dx p(x)f(\xi - x)} \quad (4.2)$$

in accordance with the Bayes formula. Writing $\pi_\xi(x) = \pi(x|\xi)$ for the posterior density we define the concept of information provision as follows: It is an operation that supplies the updated probability law:

$$p(x) \rightarrow \pi_\xi(x). \quad (4.3)$$

Stated more precisely, information provision is an updating of probability law in the form of (4.3) where $\pi_\xi(x)$ is necessarily adapted to a larger filtration (i.e. information set or knowledge) than that of $p(x)$. Having defined the notion of information provision, we proceed to consider the pricing issues. Let us first illustrate the idea, and then try to formalise it. In the above example, an uninformed investor will use the knowledge $p(x)$ to determine the optimal asset allocation, which may contain a certain amount of stock, given the initial budget W_0 . An informed investor is one who has purchased the knowledge $\pi_\xi(x)$ at a cost c and has used this knowledge to determine the optimal asset allocation, given a smaller initial budget $W_0 - c$. For example, if the p -likelihood of the stock price going up is high while the π -likelihood of this event is low, then the informed investor will purchase a smaller amount of the stock so as to circumvent likely loss. The problem then is to identify a fair value of the cost c . We would like to use a version of utility indifference argument to determine the cost. However, as we shall indicate below, for information pricing, the standard utility argument has to be augmented in a subtle manner so as to determine the fair price. Let us illustrate this by means of a simple example.

4.2.1 Optimal portfolio choice for an uninformed investor

Consider a portfolio consisting of a single risky asset s_0 and a risk-free money market (bank) account. Let θ be the amount allocated to the money market account, and ϕ be the unit of stock invested. Then starting with the initial wealth W_0 we have the allocation

$$W_0 = \theta + \phi s_0. \quad (4.4)$$

Let δ be the discount factor over the time period so that a unit cash invested in the money market account will yield δ^{-1} at the end of the period. Then the terminal wealth becomes

$$W_1 = \theta\delta^{-1} + \phi s_1.$$

The uninformed investor will determine the optimal asset allocation strategy (θ^*, ϕ^*) such that the expected utility $\mathbb{E}_p[U(W_1)]$, with respect to a given choice of the utility function $U(x)$, of the terminal wealth, is maximised. Assume that s_1 is a binary random variable taking values in the set $\{0, 1\}$ with probabilities $\{p, 1-p\}$, and the investor has an exponential utility

$$U(x) = 1 - \exp(-\alpha x) \quad \alpha \in \mathbb{R}^+.$$

The utility describes the “happiness” or “satisfaction” of the investor in relation to monetary gain. The computation of the expected utility is straightforward:

$$\mathbb{E}_p[U(W_1)] = 1 - e^{-\alpha\delta^{-1}(W_0 - \phi s_0)} (p + (1-p)e^{-\alpha\phi}),$$

where we have eliminated θ by use of (4.4). Maximising this over ϕ and writing

$$\bar{s}_0 = s_0\delta^{-1}$$

for the forwarded stock price, we obtain the following optimal asset allocation strategies:

$$\phi^* = \frac{1}{\alpha} \ln \left[\frac{(1-p)(1-\bar{s}_0)}{p\bar{s}_0} \right], \quad \theta^* = W_0 - \phi^* s_0. \quad (4.5)$$

The optimal expected utility of terminal wealth is thus

$$u(W_0) = 1 - e^{-\alpha\delta^{-1}(W_0 - \phi^* s_0)} (p + (1-p)e^{-\alpha\phi^*}).$$

Note that the optimisation has been performed without constraints. If short selling of either asset is forbidden, (θ^*, ϕ^*) are either $(0, W_0/s_0)$ or $(W_0, 0)$ depending on whether $\phi^* \geq W_0/s_0$ or $\phi^* < 0$; otherwise, (θ^*, ϕ^*) are given by (4.5). Equivalently, these conditions can be expressed in term of p in the sense that ϕ^* is given by (4.5) if

$$\frac{1 - \bar{s}_0}{\bar{s}_0 \exp(W_0\alpha/s_0) + 1 - \bar{s}_0} \leq p \leq 1 - \bar{s}_0.$$

Otherwise, $\phi^* = 0$ or W_0/s_0 .

4.2.2 Optimal portfolio choice for informed investor

We now consider the action of an informed investor in this simple example. We assume that there is an information provider who has the ability to sample

$$\xi = s_1 + \epsilon,$$

where ϵ is a noise term independent of s_1 , which for simplicity of exposition we assume to take values $\{0, 1\}$ with probabilities $\{q, 1 - q\}$. In this example the value of ξ can be 0, 1, or 2, and the *a posteriori* probability π_ξ that $s_1 = 0$ is given by

$$\begin{aligned} \pi_0 &= 1, \\ \pi_1 &= \frac{p(1 - q)}{p(1 - q) + q(1 - p)}, \\ \pi_2 &= 0. \end{aligned} \tag{4.6}$$

Since the information provider has to have invested capital in developing a system that is capable of sampling ξ (for example, in the form of a large number of computer servers), and since it has to continue investments for system maintenance, it is only reasonable to expect a payment for the provision of this information. Before we proceed to value this cost, let us first examine the strategy of an informed investor.

Let c be the cost for the purchase of information. Then the initial budget condition for the informed investor becomes

$$W_0 = \vartheta + \varphi s_0 + c, \tag{4.7}$$

where (ϑ, φ) denote portfolio positions of the informed investor. An application of the optimisation procedure described above then yields:

$$\varphi^* = \frac{1}{\alpha} \ln \left[\frac{(1 - \pi_\xi)(1 - \bar{s}_0)}{\pi_\xi \bar{s}_0} \right], \quad \vartheta^* = W_0 - c - \varphi^* s_0, \tag{4.8}$$

where $\pi_\xi = \mathbb{P}(s_1 = 0 | \xi)$. In this case, the utility of the optimal terminal wealth conditional on the value of ξ is given by

$$u_c(W_0) = 1 - e^{-\alpha \delta^{-1}(W_0 - c - \varphi^* s_0)} (\pi_\xi + (1 - \pi_\xi) e^{-\alpha \varphi^*}). \tag{4.9}$$

Similarly to the previous example, these results are based on unconstrained optimisation.

If short selling is not permitted, for instance, then these results hold if

$$0 \leq \varphi^* \leq (W_0 - c)/s_0.$$

Otherwise we must take appropriate boundary values.

4.3 Utility indifference pricing of information

We are now in a position to discuss the valuation of information in the above examples. To this end we would like to employ an argument based on utility indifference, similar to the one presented, for example, by Grossman and Stiglitz [28]. The argument intuitively goes as follows. If possession of extra information at no cost provides on average a better payoff, then every investor would seek to be informed. Consequently, there should be a positive value assigned to this information. If on the other hand the price is too high, then no one would want to be informed. Hence there exists an equilibrium level for the cost of information at which investors are indifferent, as far as their preferences are concerned.

Naively, one might then identify the utility indifferent price of information by solving the following equation for c :

$$u(W_0) = u_c(W_0) \tag{4.10}$$

so that the expected utility of the *a posteriori* optimal strategy after paying for the information equals the utility of the *a priori* optimal strategy. The intuition behind this relation is as follows: Provided that the cost c is sufficiently small, the possession of additional relevant knowledge, in the form of a strictly larger filtration, necessarily increases expected utility. Since utility is a decreasing function of the cost c , there must be a positive number c^* such that (4.10) is satisfied for $c = c^*$.

In the context of information pricing, however, this intuitive argument fails to determine the correct value of c^* . On further reflection, the reason for this becomes apparent: ignorance can make people happy. While it is true that with the additional knowledge an investor will on average perform better (cf. [13]), an uninformed investor can be more optimistic about the future outlook.

Take, for instance, an extreme case where the *a priori* probability that the stock price moving up by a significant amount is high. An uninformed investor will thus put the majority of the initial wealth into purchasing this stock. Suppose, however, that the *a posteriori* proba-

bility indicates that the stock price is more likely to move down. An informed investor will thus only invest a small fraction of the initial wealth into this stock to avoid a large loss. In this case, a likely scenario is that an informed investor knows that there is little prospect of making a large profit (i.e. small expected utility), but manages to prevent a loss, while an uninformed investor has a very high hope (i.e. large expected utility), but ending up with a loss. According to the pricing formula (4.10), therefore, the information cost becomes negative in such a scenario.

Grossman and Stiglitz [28] (see also [40, 45, 53]), on the other hand, propose the use of the following identity to fix the cost:

$$\mathbb{E}_p[U(W_p^*)] = \mathbb{E}_p[U(W_{\pi,c}^*)]. \quad (4.11)$$

Here, $W_p^* = \theta^* \delta^{-1} + \phi^* s_1$ denotes the random variable associated with the terminal wealth based on the implementation of the *a priori* optimal strategy (4.5); whereas the second wealth $W_{\pi,c}^* = \vartheta^* \delta^{-1} + \varphi^* s_1$ denotes the random variable associated with the terminal wealth based on the implementation of the *a posteriori* optimal strategy (4.8), having paid c for the information. This is likewise the pricing formula proposed by Amendinger *et al.* [3] (see also [5]). In the present example, (4.11) leads to the pricing formula

$$c_{GS}^* = \frac{\delta}{\alpha} \ln \left[\frac{1}{q^{\bar{s}_0} (1-q)^{1-\bar{s}_0}} \right]. \quad (4.12)$$

The shortcoming of the Grossman-Stiglitz pricing formula (4.11), however, is that the cost is fixed irrespective of its quality. For example, if the realised value of ξ is either 0 or 2, and if there is no cap on short selling, then an investor can purchase this information for a modest cost to make infinite profit, leading to an arbitrage because starting with zero initial wealth $W_0 = 0$ one can construct a portfolio such that $\mathbb{E}[W_1] = \infty$ under any measure equivalent to $\mathbb{P}$. In contrast, in our theory we would like the cost of information be infinite, if this information were to provide infinite benefit. (See also remarks in Appendix A).

4.4 Information dependent price

With this in mind, we deduce that the “correct” pricing formula should be given by solving

$$\mathbb{E}_\pi[U(W_p^*)] = \mathbb{E}_\pi[U(W_{\pi,c}^*)] \quad (4.13)$$

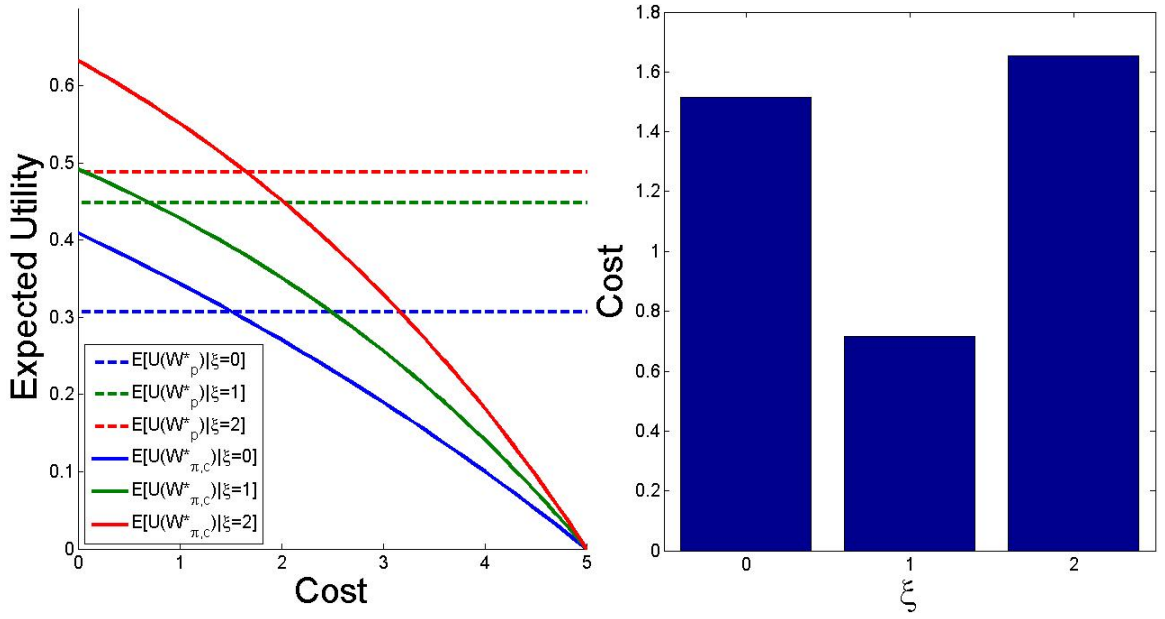


Figure 4.1: Left: Expected utilities $\mathbb{E}_\pi[U(W_p^*)]$ for $\xi = 0, 1, 2$ as a function of the cost c are shown here as three horizontal lines. For each value of ξ we also plot $\mathbb{E}_\pi[U(W_{\pi,c}^*)]$, when short selling is strictly prohibited. When c is sufficiently small, the latter utilities are higher, but as the cost increases, they decrease. At $c = c^*$ the two utility functions for each ξ agree and determine the unique equilibrium price for which the investor is indifferent. Here we have chosen parameters to be $\alpha = 0.1$, $s_0 = 0.5$, $W_0 = 5$, $p = 0.4$, $q = 0.7$, and $\delta = 0.95$. The numerical values for the cost for each ξ are plotted as a histogram on the right side. In this example, we have $c^*(0) = 1.51$, $c^*(1) = 0.71$, and $c^*(2) = 1.65$. The average cost is $\bar{c} = 1.10$.

for c . In other words, we consider with hindsight what would have happened to the uninformed strategy, and compare this with the informed strategy. In this way, a consistent price for information can be deduced. We propose this to be the basis with which information can be priced using a utility indifference argument. In the case of the above example, (4.13) amounts to solving

$$e^{-\alpha\delta^{-1}(W_0-\phi^*s_0)} (\pi_\xi + (1 - \pi_\xi)e^{-\alpha\phi^*}) = e^{-\alpha\delta^{-1}(W_0-c-\varphi^*s_0)} (\pi_\xi + (1 - \pi_\xi)e^{-\alpha\varphi^*})$$

for c . This gives

$$c^* = \frac{\delta}{\alpha} \ln \left[\frac{\pi_\xi e^{\alpha\bar{s}_0\phi^*} + (1 - \pi_\xi)e^{\alpha\phi^*(\bar{s}_0-1)}}{\pi_\xi e^{\alpha\bar{s}_0\varphi^*} + (1 - \pi_\xi)e^{\alpha\varphi^*(\bar{s}_0-1)}} \right],$$

which makes it apparent that when $\phi^* = \varphi^*$, i.e. when ξ provides no additional information (which can happen in the present example if for instance $p = q = \frac{1}{2}$ and $\xi = 1$), or when the additional information provided by ξ results in no change of the strategy, the cost is

identically zero. Note that the dependence of π_ξ on ξ implies that the value of c^* also depends on ξ . Thus, for instance if $\xi = 0$ or $\xi = 2$, and if unlimited short selling is allowed, then $c^* = \infty$ as desired. In many cases the dependence on ξ is natural, because different values of ξ embody different information contents.

The plot for cost versus information content in figure 4.1 shows the indifference price of information for different values of ξ . It is given by the value of c where the two curves of uninformed and informed expected utilities meet. If the cost were higher then this would be favourable to the information provider, while if the cost were lower this would be more favourable to the information seeker. The lowest cost is at $\xi = 1$, where the corresponding pairs (S_T, ϵ) take the values $\{0, 1\}$ or $\{1, 0\}$. It is where the uncertainties are the highest as compared to other values of ξ . Indeed, this is the only case of uncertainty in our example.

In contrast to the informational cost formula proposed by Grossman and Stiglitz, the one that we proposed is that of no arbitrage. This can be shown by the fact that the wealth process, $W_{\pi,c}^*$, is a martingale (see Appendix A for detail calculation). In particular, the unconditional expectation of the terminal wealth is a constant under the $\mathbb{P}$ measure:

$$\mathbb{E}[W_{\pi,c}^*] = B(\varphi_1^*) - A(\varphi_1^*)/\alpha + \phi^*((1-p)(1-q)(1-\bar{s}_0) - pq\bar{s}_0).$$

where $A(\varphi_1^*)$, $B(\varphi_1^*)$ are constants defined in Appendix A. Hence one can find an equivalent martingale measure $\mathbb{Q}$ such that

$$\mathbb{E}^{\mathbb{Q}}[W_{\pi,c}^*] = W_0,$$

and this implies no arbitrage.

4.5 Information independent price

While it is useful to be able to express the cost $c^*(\xi)$ as a function of the “content” ξ of the information, there are likewise cases where it is desirable to associate a fixed price for information. This applies to the provision of generic information; whereas for the purchase of a specialised information, it might be unreasonable to expect a flat rate for the information, irrespective of its content or quality. We shall therefore consider both cases.

If short selling is prohibited, then even with the knowledge $\xi = 0$ or $\xi = 2$ an investor can only make a finite profit. Hence in this case the cost should also be strictly bounded. In figure 4.1 on the left panel we plot $\mathbb{E}_\pi[U(W_p^*)]$ for three values of ξ as functions of c (clearly,

$\mathbb{E}_\pi[U(W_p^*)]$ for each value of ξ is constant in c), and compared these with $\mathbb{E}_\pi[U(W_{\pi,c}^*)]$ as functions of c , when short selling is forbidden. For each value of ξ the intersection of the two functions determines the cost $c^*(\xi)$, shown as a histogram also in figure 4.1. For the parameter values chosen in this example, the “upside” information (when $\xi = 2$) is worth a little more than the “downside” information (when $\xi = 0$). In this way we are able to assign prices for information of specific quality. If we are interested in assigning a flat rate $\bar{c}^*$, then we average the individual cost according to

$$\bar{c}^* = \sum_k \mathbb{P}(\xi = \xi_k) c^*(\xi_k).$$

In the present example we have

$$\begin{aligned} \mathbb{P}(\xi = 0) &= pq, \\ \mathbb{P}(\xi = 1) &= p(1 - q) + (1 - p)q, \quad \text{and} \\ \mathbb{P}(\xi = 2) &= (1 - p)(1 - q), \end{aligned}$$

which can be used to determine $\bar{c}^*$. Since the *a priori* expectation of our pricing formula (4.13) gives the Grossman-Stiglitz formula (4.11), we see that solutions to (4.11) and (4.13) correspond, respectively, to the concepts of an *annealed* average and a *quenched* average often considered in the spin-glass theory. In particular, by an application of Jensen’s inequality, the Grossman-Stiglitz cost provides a lower bound for our cost:

$$c_{GS}^* \leq \bar{c}^*.$$

4.6 Separation of density functions

The informed investor has purchased extra information, putting him in a better position when choosing his assets. One way of quantifying the “gap” between the uninformed and the informed information is to express their separation in terms of the Bhattacharyya spherical distance between p and π . The Bhattacharyya spherical distance between two probabilities $\{p_i\}$ and $\{q_i\}$ for $i = \{1, \dots, N\}$ is given by

$$D_B(p|q) = \cos^{-1} \sum_{i=1}^N \sqrt{p_i} \sqrt{q_i}.$$

Intuitively, the idea of the Bhattacharyya spherical distance can be understood by regarding

$\{\sqrt{p_i}\}$ and $\{\sqrt{q_i}\}$ as representing normalised vectors in the Hilbert space $\mathcal{H} = l^2$. Then $0 \leq D_B \leq \pi/2$ can be seen to represent the angular separation between the two vectors in the Hilbert space $\mathcal{H}$.

In figure 4.2 we plot the Bhattacharyya spherical distance $D_B(p, \pi_\xi)$ for three different realised values of ξ , as a function of the prior probability p . From these distance plots we can deduce how much the posterior probability deviates from the prior probability in different circumstances. In particular, values of $D_B(p, \pi_\xi)$ are large when the additional information comes as a significant surprise.

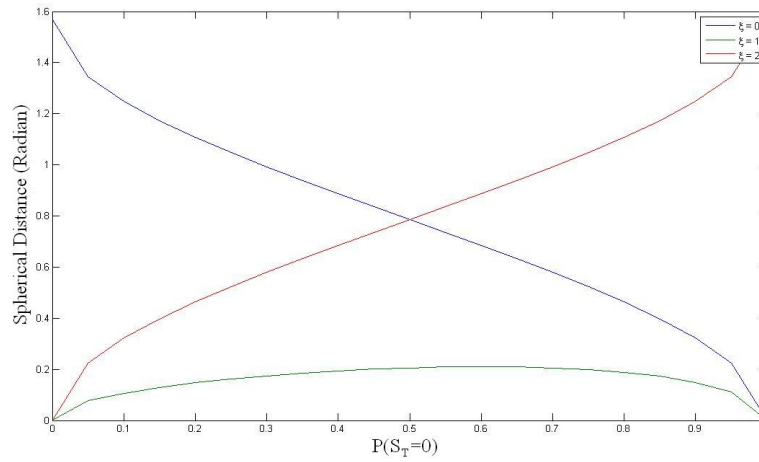


Figure 4.2: The Bhattacharyya spherical distance between p and π for difference values of ξ (the left top curve corresponds to $\xi = 0$, left middle to $\xi = 2$, and left bottom to $\xi = 1$). The horizontal axis represents different values of p .

4.7 Utility indifference information pricing with continuous payoff

Having introduced our framework for information pricing, our next objective is to express the cost in terms of the quantity of information so that we are able to quote the price in the form, for example: “\$X for Y bits of upside information”. This is highly desirable because the representation of the cost as a function of ξ need not be very practical. Our definition of information as a quantity that generates the transformation (4.3) allows us to proceed by the consideration of relative entropy between the *a priori* and the *a posteriori* probabilities. The examples considered above, however, are too restrictive, because the *a priori* and the *a posteriori* probabilities are not absolutely continuous with respect to each other, and thus relative entropy cannot be defined. Indeed, the idea that an information provider can ascertain the future value s_1 at time 0, which would be the case if the sampled

value of ξ is either 0 or 2, is somewhat artificial and unrealistic. We therefore consider another simple example in which this issue does not arise.

We assume the setup as before, except that the random variable s_1 is assumed normally distributed with mean μ and variance σ^2 :

$$\xi = s_1 + \epsilon, \quad \text{where} \quad s_1 \sim N(\mu, \sigma), \quad \epsilon \sim N(0, \sigma_\epsilon).$$

Thus, the model is similar to the Bachelier model, rather than a geometric Gaussian model, in that the asset price can take negative values. In the case of an uninformed investor, a straightforward Gaussian integration then gives the expected utility of terminal wealth as:

$$\mathbb{E}_p[U(W_1)] = 1 - \exp \left[-\alpha \left(\delta^{-1}(W_0 - \phi s_0) + \phi \mu - \frac{1}{2} \alpha \phi^2 \sigma^2 \right) \right].$$

It follows that the optimal allocation strategy is determined by

$$\phi^* = \frac{\mu - \bar{s}_0}{\alpha \sigma^2}, \quad \theta^* = W_0 - \frac{(\mu - \bar{s}_0) s_0}{\alpha \sigma^2}.$$

Similarly to the previous example, if there is a restriction on short selling, then these results hold provided that $0 \leq \phi^* \leq W_0/s_0$; otherwise, $\phi^* = 0$ or $\phi^* = W_0/s_0$.

An informed investor is able to purchase the sampled value of ξ given by (4.1) above, except that the noise variable ϵ is now assumed normally distributed with mean zero and variance σ_ϵ^2 . In this case, a standard result in Gaussian filtering theory [54] (in fact, a straightforward application of the Bayes formula) shows that the *a posteriori* probability law for s_1 given ξ is also normally distributed:

$$s_1 | \xi \sim N \left(\frac{\mu \sigma_\epsilon^2 + \xi \sigma^2}{\sigma_\epsilon^2 + \sigma^2}, \sqrt{\frac{\sigma^2 \sigma_\epsilon^2}{\sigma^2 + \sigma_\epsilon^2}} \right) \equiv N(\mu_\xi, \sigma_\xi).$$

The optimisation problem thus reduces to maximising

$$u_c(W_0) = 1 - \exp \left(-\alpha \left(\delta^{-1}(W_0 - c - \varphi s_0) + \varphi \mu_\xi - \frac{1}{2} \alpha \varphi^2 \sigma_\xi^2 \right) \right).$$

The result is

$$\varphi^* = \frac{\mu_\xi - \bar{s}_0}{\alpha \sigma_\xi^2}, \quad \vartheta^* = W_0 - c - \frac{(\mu_\xi - \bar{s}_0) s_0}{\alpha \sigma_\xi^2}. \quad (4.14)$$

As before, if short selling is not permitted, we have the conditions $0 \leq \varphi^* \leq (W_0 - c)/s_0$

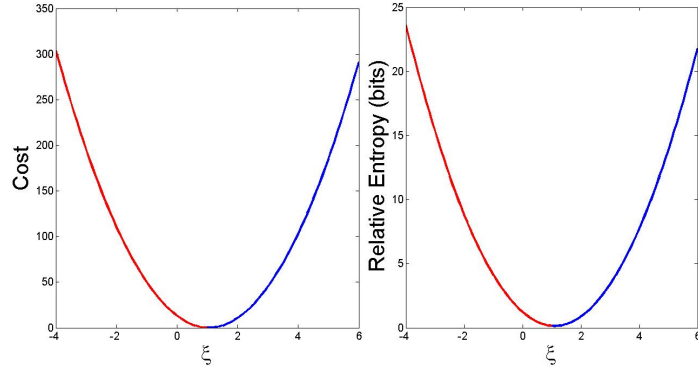


Figure 4.3: Left: Cost $c^*(\xi)$ of information as a function of the realised value of ξ . Right: Relative entropy $I(\xi) = I(\pi_\xi|p)$ as a function of the realised value of ξ . The parameter values used here are: $s_0 = 1$, $\delta = 0.95$, $\alpha = 0.1$, $W_0 = 10$, $\mu = 1.1$, $\sigma^2 = 0.2$, $\mu_\epsilon = 0$, and $\sigma_\epsilon^2 = 0.2$. The red curves correspond to pieces of downside information (i.e., values of ξ that indicate decrease in s_1 as compared to the *a priori* expectation) and the blue curves correspond to upside information.

for which (4.14) is valid; otherwise we must take boundary values.

Let us evaluate the cost of information by the principle (4.13). This amounts to finding the value for c that solves

$$\begin{aligned} & \exp \left[-\alpha \left((W_0 - \phi^* s_0) \delta^{-1} + \phi^* \mu_\xi - \frac{1}{2} \alpha \phi^{*2} \sigma_\xi^2 \right) \right] \\ &= \exp \left[-\alpha \left((W_0 - c - \varphi^* s_0) \delta^{-1} + \varphi^* \mu_\xi - \frac{1}{2} \alpha \varphi^{*2} \sigma_\xi^2 \right) \right], \end{aligned}$$

with the solution

$$c^* = \frac{1}{2} \alpha \delta \sigma_\xi^2 (\phi^* - \varphi^*)^2. \quad (4.15)$$

This provides the cost $c^*(\xi)$ as a function of the sampled value of ξ . That is, before the purchase is made, a client of the information provider can agree to the price structure (4.15) so that depending on which value of ξ the information provider produces, the client will pay the cost appropriate for that information. If the use of a flat rate $\bar{c}^*$ for the cost is desirable, then we average $c^*(\xi)$ over ξ under the suitable Gaussian measure for ξ . This is given by the expression

$$\bar{c}^* = \frac{1}{2\alpha} \delta \sigma_\xi^2 \left[(\mu - \bar{s}_0)^2 (\sigma^{-2} - \sigma_\xi^{-2})^2 + \frac{\sigma^2}{\sigma_\xi^2 \sigma_\epsilon^2} \right].$$

It is interesting to draw attention to the expression (4.15) we have obtained for the cost. Up to a proportionality factor, the cost is given by the quadratic difference between the *a priori*

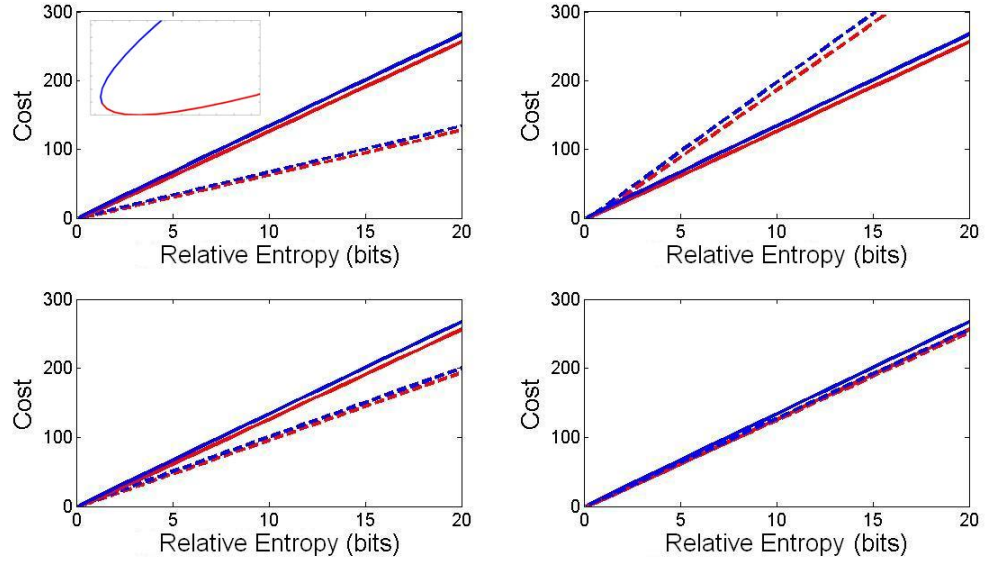


Figure 4.4: The cost $c^*(I)$ as a function of the quantity of information I , when there is no constraint for short selling. The solid lines in all four plots (they are identical) correspond to the parameter values used in figure 4.3. While keeping other parameters fixed, the following changes in parameter values are used, represented by dashed lines in the plots. Top left: α has been increased to $\alpha = 0.2$; top right: σ^2 has been increased to $\sigma^2 = 0.4$; bottom left: σ_ϵ^2 has been increased to $\sigma_\epsilon^2 = 0.4$; and bottom right: δ has been decreased to $\delta = 0.92$. The red curves correspond to downside information, and the blue curves correspond to upside information.

and the *a posteriori* portfolio positions. This indicates that the cost is indeed determined by the “usefulness” of information. We shall come to discuss further on this point below.

As indicated above, we wish to associate the cost with the quantity of information. For this purpose, we consider the relative entropy

$$\begin{aligned}
 I(\pi_\xi|p) &= \int_{-\infty}^{\infty} \pi_\xi(s) \ln \left(\frac{\pi_\xi(s)}{p(s)} \right) ds \\
 &= \frac{1}{2} \left[\ln \left(\frac{\sigma^2}{\sigma_\xi^2} \right) + \frac{\sigma_\xi^2}{\sigma^2} - 1 \right] + \frac{1}{2} \frac{(\mu_\xi - \mu)^2}{\sigma^2}
 \end{aligned} \tag{4.16}$$

between the prior and the posterior densities. It should be evident from the decomposition $\xi = s_1 + \epsilon$ that there is a critical value $\xi^\dagger$, given by the *a priori* expectation of s_1 , such that the *a posteriori* density is least informative, and such that the information content increases as the value of ξ increases or decreases away from this critical level. In other words, relative entropy is monotonically decreasing in ξ if $\xi < \xi^\dagger$, and monotonically increasing in ξ if $\xi \geq \xi^\dagger$. One might therefore expect that the cost should also exhibit the same monotonic

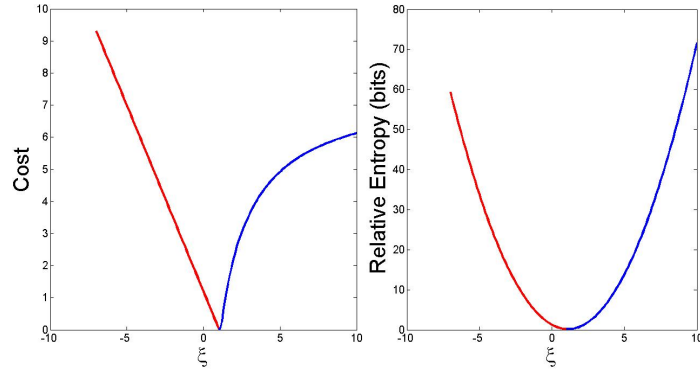


Figure 4.5: Left: Cost $c^*(\xi)$ of information as a function of the realised value of ξ when short selling is strictly prohibited. Right: Relative entropy $I(\xi) = I(\pi_\xi|p)$ as a function of the realised value of ξ . The cost $c^*(\xi)$ measures the usefulness of information, while relative entropy $I(\xi)$ measures the quantity of information. The parameter values used here are: $s_0 = 1$, $\delta = 0.95$, $\alpha = 0.1$, $W_0 = 10$, $\mu = 1.1$, $\sigma^2 = 0.2$, $\mu_\epsilon = 0$, and $\sigma_\epsilon^2 = 0.2$. The red curves correspond to pieces of downside information (i.e. values of ξ that indicate decrease in s_1 as compared to the *a priori* expectation) and the blue curves correspond to upside information.

dependence on $|\xi - \xi^\dagger|$. Surprisingly, however, this is not the case. As indicated above, the pricing formula (4.15) in the Gaussian model shows that if the additional information does not alter portfolio positions, then an investor is not going to pay for that information. As a consequence, minimum information can be more costly (i.e. useful) than a larger quantity of information.

In figure 4.3 we plot the cost function $c^*(\xi)$ where there are no constraints for short selling, and the relative entropy $I(\xi)$. By inverting these two functions we can determine the cost $c^*(I)$ as a function of the information content, as shown in figure 4.4 for a range of parameter values. Note that in the numerical plots we have converted the logarithmic basis so that relative entropy I is measured in the conventional binary units of “bits” rather than the Shannon units used in (4.16). Hence the horizontal axis represents information content expressed in binary units.

4.7.1 Unconstrained scenario

We observe that costs of upside information in these examples are always slightly higher than those of downside information, and that in each case the cost is approximately a linear function of the quantity of information. The general trend that we have found is as follows.

The cost in the unconstrained case is a decreasing function of the risk aversion parameter

α in the utility function. This is because for a high value of α the investor “saturates” the upper bound of the utility function sooner, as compared to a lower value of α ; hence the investor reaches a “happy” or “satisfied” level, and any further monetary gain will only add little value. In other word, a high α investor is an investor of a risk-averse type.

The cost is an increasing function of the signal uncertainty σ . This is natural, because if one has a very noisy signal to start with, then a better-quality additional information becomes more valuable. However, if one is *a priori* clear about the realised value of s_1 then additional news seems less attractive.

The cost is a decreasing function of the noise uncertainty σ_ϵ . That is, the stronger the noise contained in the information is, the less valuable the information becomes. This is entirely expected, since the reliability of the information is clearly decreasing in the width σ_ϵ of the noise.

Finally, the cost is a slightly increasing function of the discount factor δ , which again seems natural because the growth of the money market account is given by δ^{-1} .

4.7.2 Constrained scenario

The cost structure changes somewhat if we restrict short selling. This is because the upside gain from exploiting additional information is restricted in this case. In particular, if short selling (including cash borrowing) is prohibited, then the cost is bounded above by the initial wealth W_0 .

We plot in figure 4.5 the cost $c^*(\xi)$ as a function of ξ in the case for which short selling is strictly forbidden. The result reveals some interesting insights: a client of an information provider is willing to pay significantly higher costs for large downside information, as compared to upside information, and the increase in $c^*(\xi)$ for upside information is satiated relatively early.

This feature of course is closely related to the model setup chosen here: Since s_1 is assumed Gaussian, purchase of the stock by any amount can result in an unbounded loss. Therefore, investors are willing to spend the totality W_0 of the initial budget as a protection against large losses, leaving behind no asset for any investment. Conversely, investors are inclined not to spend significant portions of the initial budget on upside information, because if there are opportunities to profit from investments, they would not want to lose these opportunities by spending all the budgets. This is a remarkable feature inherited from the utility function, in which the gradient decreases as we go up the x axis.

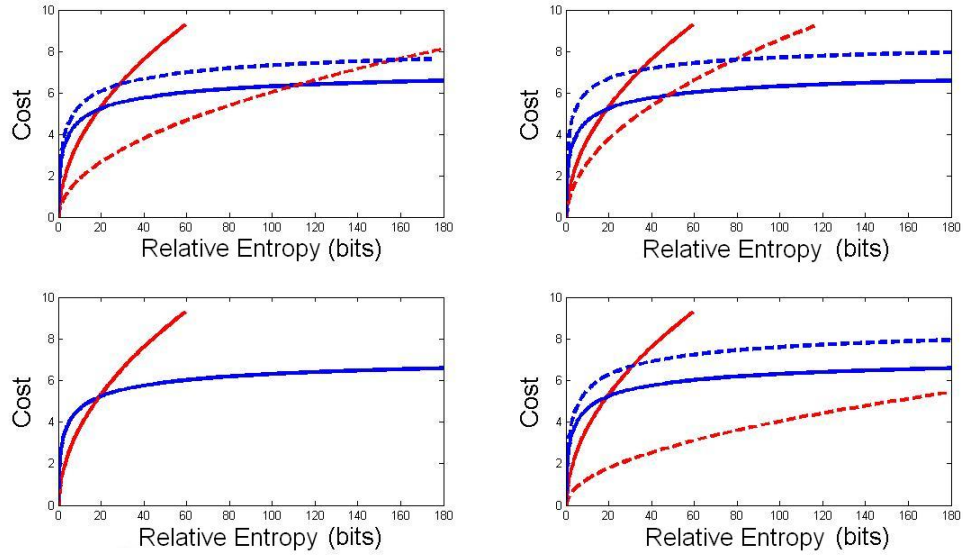


Figure 4.6: The cost $c^*(I)$ as a function of the quantity of information I , when short selling is strictly prohibited. The solid lines in all four plots (they are identical) correspond to the parameter values used in figure 4.5. While keeping other parameters fixed, the following changes in parameter values are used, represented by dashed lines in the plots. Top left: α has been increased to $\alpha = 0.2$; top right: σ^2 has been increased to $\sigma^2 = 0.4$; bottom left: σ_ϵ^2 has been increased to $\sigma_\epsilon^2 = 0.4$; and bottom right: δ has been decreased to $\delta = 0.92$. The red curves correspond to downside information, and the blue curves correspond to upside information.

Although the cost structure changes when there are constraints, we can again invert these relations to quote unique prices for upside and downside information. The results are shown in figure 4.6 for a range of parameter values. We observe that when there are constraints the price of information exhibits somewhat nontrivial behaviour as we change the model parameters.

One noticeable behaviour seen across all constrained prices is that downside information has sharper gradients than upside news. This reflects the fact that as we move down the utility curve from any particular point we have sharper changes in its gradient. Therefore, investors with such utility are always willing to pay a higher cost for negative news than positive ones of the same magnitude, when measured in bits.

Generally, the cost for downside information decreases and upside information increases as we increase the risk aversion parameter α . The reason for this is because, with a high α value the utility of the investor would already be at a higher values, as compared to a low α investor, referenced at the same fixed wealth. As wealth decreases, the high α investor would remain content initially, hence not yet ready to pay a high cost for the

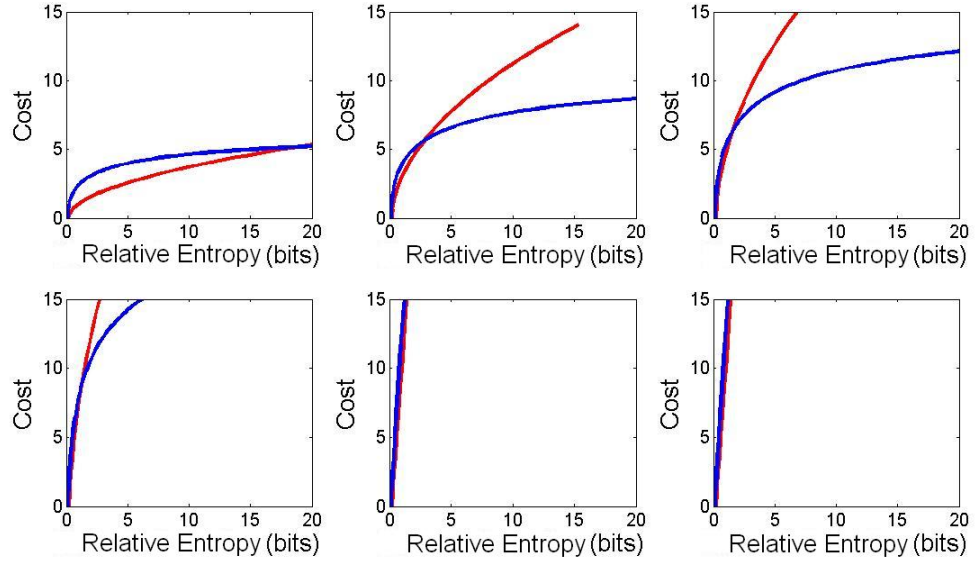


Figure 4.7: Interpolation of the information price $c^*(I)$ from the strictly constrained case to the unconstrained case. Top row, from left: (i) borrowing is strictly forbidden; (ii) borrowing is capped at $W_0/2$; (iii) borrowing is capped at W_0 . Bottom row, from left: (i) borrowing is capped at $2W_0$; (ii) borrowing is capped at $5W_0$; (iii) unlimited borrowing permitted.

downside information. Whereas for a low α investor, she may enter the “unhappy” level faster, and thus is willing to pay a higher cost to compensate the potential loss. For the upside information, a high α investor has the steeper gradient, hence is willing to pay more. The effect does not appear on the downside because we restrained from short selling, so a negative information does not help the investor to increase utility significantly, but only to limit the loss.

As the signal uncertainty σ increases, the cost for downside information becomes less expensive while upside news become more expensive. That upside news become more expensive is because the investor is willing to pay more for a clearer signal to profit more. For downside news, on the other hand, the worth of information seems to diminish rapidly as the signal uncertainty is increased. These two features are not symmetric because in this constrained case where no short selling is permitted, the investor cannot profit from downside news, but can only limit losses, whereas the investor can profit from upside news.

The bottom left plot in figure 4.6 has varying levels of noise variance, this includes the cases of $\sigma_\epsilon = \{0.1, 0.2, 0.4\}$. The graphs are identical in all of the cases. The idea is that the noisier the piece of information is, the less one is willing to pay and vice versa. This is because an information with large noise gives less information as measured by the relative

entropy, and this in turn gives less bits to be paid for. If we know that a piece of information is noisy than this does not affect our decision, only that we know it is not very informative and we pay less.

The two cases examined here, namely, the strictly constrained and totally unconstrained cases, can be interpolated by gradually relaxing the borrowing cap, as shown in figure 4.7. These intermediate cases are often more realistic because of leveraging. In the case of the example considered here we observe that by raising the cap of borrowing to five times the total wealth the net effect is already equivalent to allowing for unlimited borrowing. The effect of leveraging is also of interest in showing in which way the upside and downside information costs intertwine.

4.8 Analysis of expected loss at maximum information cost

One interesting observation that we have made in the case where there are trading constraints on short selling is that for a sufficiently strong indication that the price may drop significantly, investors are willing to spend the totality of their initial wealth to purchase such information so as to avoid large losses. In this case, no upside gain will be made; instead, all the asset is consumed to prevent large losses. A natural question that arises then is: what is the amount of expected loss, as a percentage of the initial wealth, such that investors are willing to trade their entire asset in exchange for the information so that no asset is left for further investment? To address this interesting question, let us denote by ξ^* the critical value of information ξ such that $c^*(\xi^*) = W_0$. Then the conditional expected loss is given by

$$W_0 - \mathbb{E}[W_T | \xi = \xi^*] = W_0(1 - \exp(rT)) + \phi^*(\mu_{\xi^*} - \bar{s}_0),$$

where ϕ^* is the optimal risky asset allocation with no short selling. The standard deviation of the conditional loss is similarly given by

$$(\text{Var}[W_T | \xi = \xi^*])^{1/2} = \phi^* \sigma_{\xi^*}.$$

We have calculated the conditional expected loss and the associated standard deviation for a range of parameter values, expressed in percentage of the initial wealth W_0 . The result is given in Table 4.1 below.

	α			$\alpha = 0.1$	$s_0 = 10,$
	0.1	0.3	0.5	$W_0 = 100$	$W_0 = 100$
Expected loss (%)	99.67	99.85	99.89	99.94	99.67
Standard deviation of loss (%)	7.7	2.58	1.54	0.77	2.43

Table 4.1 A table showing the expected loss in percentage of the initial wealth for a range of parameter values. The parameter values being used here are $r = 5\%$, $T = 1$, $\alpha = 0.1$, $\sigma = 0.2$, $\mu = 1.1$, $\sigma_\epsilon = 0.2$, and $\mu_\epsilon = 0$. The values of the initial wealth are $W_0 = 10$ and $W_0 = 100$, as indicated in the table, and the value of the initial asset is $s_0 = 1$, unless otherwise specified.

Since the terminal value of the risky asset is normally distributed in the present example, an investor purchasing the asset is in effect entering into a contract that entails unbounded gain, but also unbounded loss. This is precisely why when short selling is not permitted an investor would rather spend the totality of the initial wealth to purchase the information so that investment opportunities are lost, provided that the *a posteriori* expected loss is greater than a threshold value. The results of Table 4.1 show that this threshold is about 100% of the initial wealth, irrespective of the model parameters. It is interesting that the the expected loss turns out to be more or less exactly the initial wealth.

4.9 Cost of information in continuous time

It is of interest to extend the single-period models considered previously to continuous-time models within the context of utility indifference pricing theory. In this section we shall generalise the method we presented in the previous section to a dynamical, continuous time setting. The scenario of receiving an extra noisy information falls under the context of “filtration enlargement” in the literature. We shall follow existing literature of filtration enlargement techniques, specifically, Amendinger *et al.* [3], and employ the standard portfolio optimisation procedure using the concept of convex duality [44, 46] to work out the cost of information through the augmented utility indifferent price framework.

4.9.1 Utility optimisation

We begin by reviewing the theory of utility-based portfolio pricing in incomplete markets using the convex duality method. This method has been widely studied and developed in the literature; to name a few, references [18, 19, 37, 39, 44, 46] and many others since the 1980s. This section follows, in particular, the setting of [44, 46]. In what follows, we

present some results on information cost in continuous time; background material regarding filtration enlargement and utility optimisation are presented in Appendices B and C. We refer readers to these appendices and references for detailed arguments.

We shall not assume market completeness. Incomplete markets are usually associated with the notion of perfect hedgeability, i.e. whether a contingent claim can be replicated perfectly. We consider the incomplete market in the sense of insufficient information regarding the dynamics of the asset. This would lead to models of stochastic volatility and uncertain drift, because there are extra randomness or uncertainty in the asset. Consequently, the market is inefficient, in the sense that information flow is not perfect and hence not all information available in the market are taken into account by every participant. Extra information can nevertheless be acquired by extra costs.

In this section, we assume the setup, as described in the BHM information framework [17], whereby the asset price process is adapted to the filtration generated by a market information process $\{\xi_t\}$, and that the drift and volatility are deterministic given ξ_t . Hence the background filtration is given by $\mathcal{F}_t = \sigma(\xi_s : 0 \leq s \leq t)$. The time set is denoted by $\mathbb{T} = [0, T]$.

There are $n + 1$ assets, one money market account S_t^0 and n risky assets $(S^1, \dots, S^n)$. Assume that the dynamics of the assets take the form:

$$dS_t^i = \mu_t^i S_t^i dt + \sigma_t^i S_t^i dW_t, \quad (4.17)$$

where $\{W_t\}$ is a d -dimensional Brownian motion on $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, \mathbb{P})$ and $d \geq n$. To simplify the notation we shall write $dS = dS^i$, and similarly for $\{\mu_t\}$ and $\{\sigma_t\}$. The processes $\{\mu_t\}$, $\{\sigma_t\}$ are progressively measurable processes in $\mathbb{R}^n$ and $\mathbb{R}^{n \times d}$ with the property that

$$\int_0^T |\mu_t| dt + \int_0^T |\sigma_t|^2 dt < \infty$$

almost surely. We assume that for all $t \in [0, T]$, and the matrix σ_t is of rank n . Thus the matrix $\sigma_t \sigma_t^T$ of dimension $n \times n$ is invertible. The money market account with deterministic interest rate $\{r_t\}$ is given by

$$S_t^0 = \exp \left(\int_0^t r_s ds \right).$$

For the requirement of no arbitrage, we assume that the set $\mathcal{M}_e$ of equivalent martingale

measures associated with the asset is nonempty:

$$\mathcal{M}_e(S) \neq \emptyset.$$

If the set $\mathcal{M}_e(S)$ contains only one element then evidently the market is complete. (Further details about the no arbitrage condition can be found in Appendix C.1.)

Now a self-financing portfolio is defined by a pair $\{w, \phi_t\}$ where $w \in \mathbb{R}$ is the initial wealth and $\phi_t = \{\phi_t^i\}_{i=1,\dots,d}$ is a predictable and S -integrable vector of processes specifying the number of shares of each of the risky assets held in the portfolio. The value of the self-financing portfolio at time t is given by:

$$W_t := w + \int_0^t \phi_u dS_u =: w + (\phi \cdot S)_t,$$

where $\int \phi dS \equiv \sum^n \int \phi^i dS^i$. Knowing the “control process” $\{\phi_t\}$ is similar to knowing the wealth process $\{W_t\}$. If $\{S_t\}$ is a local-martingale then $\{W_t\}$ is known to be a sigma-martingale.

As in the previous section, we consider a utility function of the form

$$U(x) = -e^{-\alpha x}, \quad (4.18)$$

where α is the risk aversion rate; large values of α correspond to high risk aversion. The exponential utility function is strictly concave and satisfies the Inada conditions:

$$U'(0) := \lim_{x \downarrow 0} U'(x) = \infty, \quad \text{and} \quad U'(\infty) := \lim_{x \rightarrow \infty} U'(x) = 0.$$

Our aim now is to solve the optimisation problem

$$u(w) := \sup_{\phi \in \mathcal{A}} \mathbb{E}[U(W_T)] = \sup_{\phi \in \mathcal{A}} \mathbb{E}U \left(w + \int_0^T \phi_t d\bar{S}_t \right), \quad (4.19)$$

where

$$\bar{S}_t = S_t \exp\left(-\int_0^t r_s ds\right)$$

is the discounted price and the admissible set of strategies is defined as

$$\mathcal{A} = \{\phi : (\phi \cdot S) \text{ is a } \mathbb{Q}\text{-martingale } \forall \mathbb{Q} \in \mathcal{M}_e\}.$$

We aim to use convex optimisation techniques to solve the problem described by equation (4.19). We shall only summarise here important steps in solving the primal problem (4.19); further details are provided in appendix C.2. By Fenchel-Legendre transform, we obtain the conjugate relation

$$V(y) := \sup_{x \in \mathbb{R}^+} [U(x) - xy], \quad y > 0,$$

and

$$U(x) = \inf_{y > 0} [V(y) + xy], \quad x > 0.$$

In other word, the solution to the primal problem can be deduced by solving the dual problem:

$$v(y) := \inf_{\mathbb{Q} \in \mathcal{M}_e} \mathbb{E}V[(yZ_T)], \quad (4.20)$$

where Z_T is the Radon-Nikodym derivative of $\mathbb{Q}$ with respect to $\mathbb{P}$. In this formulation, the optimal asset allocation strategy is given by (Appendix C):

$$\phi_t^* \bar{S}_t = \frac{1}{\alpha} \sigma_t (\sigma_t \sigma_t^T)^{-1} \sigma_t (\sigma_t \sigma_t^T)^{-1} (\mu_t - r_t \mathbf{1}). \quad (4.21)$$

The primal value function can be derived at once to be

$$u(w) = -\exp(-\alpha w - \mathcal{H}(\mathbb{Q}^E, \mathbb{P})).$$

The so-called maximum entropy has the following form:

$$\mathcal{H}(\mathbb{Q}^E, \mathbb{P}) = \frac{1}{2} \int_0^T |\lambda_t|^2 dt.$$

4.9.2 Utility optimisation with information

The effect of new information ξ is reflected through the enlargement of filtration, the new probability space is now $(\Omega, \mathcal{F} \vee \mathcal{G}, \mathbb{G} = \{\mathcal{G}_t\}, \mathbb{P})$ where $\mathcal{G} = \sigma(\xi)$ and $\mathcal{G}_t = \mathcal{F}_t \vee \mathcal{G}$. Following the line of arguments presented in Amendinger *et al.* we shall employ the martingale preserving measure $\tilde{\mathbb{Q}}$ such that $\tilde{\mathbb{Q}} = \mathbb{P}$ on $\mathcal{G}$ and $\tilde{\mathbb{Q}} = \mathbb{Q}$ on $\mathcal{F}_T$, where $Z_T^G = Z_T^F / p^G$ is its Radon-Nikodym derivative with respect to $\mathbb{P}$. The key property of the measure $\tilde{\mathbb{Q}}$ is that $\mathcal{F}_t$ -martingale remains martingale under $\mathcal{F}_t \vee \mathcal{G}$. Further details about the martingale

preserving measure are given in Appendix B.

For an investor to obtain information $\{\xi_t\}$ adapted to a larger filtration, it will naturally entail a cost c . The new portfolio and value function are defined as

$$u^G(w - c) := \sup_{\varphi \in A^G} \mathbb{E}^{\mathbb{P}} \left[U \left(w + \int_0^T \varphi_t d\bar{S}_t - c \right) \middle| \mathcal{G}_t \right].$$

The admissible set of strategies A^G in this case is defined by a set of integrable strategies such that

$$A^G = \{ \varphi : (\varphi, \bar{S})_t \text{ is a } \tilde{\mathbb{Q}}\text{-martingale under } \mathbb{G} \}.$$

For informed investors, we have the new definition of equivalent martingale measures:

$$\mathcal{M}_e^G = \{ \mathbb{Q} \sim \mathbb{P} \text{ on } (\Omega, \mathcal{G}_T) : \text{discounted price of } S \text{ is a } \tilde{\mathbb{Q}}\text{-local martingale} \}.$$

We can follow the same argument as for the uninformed case and derive the relative entropy for the informed investor. If we denote by $\mathbb{E}^G[\cdot]$ the risk-neutral expectation under $\mathbb{G}$, and using the properties that this defines the minimal martingale preserving measure $\tilde{\mathbb{Q}}^E$:

$$\begin{aligned} \mathcal{H}^G(\tilde{\mathbb{Q}}^E, \mathbb{P}) &= \mathbb{E} \left[\frac{d\tilde{\mathbb{Q}}^E}{d\mathbb{P}} \ln \frac{d\tilde{\mathbb{Q}}^E}{d\mathbb{P}} \middle| \mathcal{G}_t \right] \\ &= \mathbb{E}^G \left[\ln Z_T^G \middle| \mathcal{G}_t \right] \\ &= \mathbb{E}^G \left[\ln \frac{Z_T^G}{Z_T} + \ln Z_T \middle| \mathcal{G}_t \right] \\ &= \mathbb{E}^G \left[\ln \frac{Z_T^G}{Z_T} \middle| \mathcal{G}_t \right] + \mathbb{E}[Z_T \ln Z_T]. \end{aligned}$$

The minimal change-of-measure martingale is given by

$$Z_T^G = \frac{d\tilde{\mathbb{Q}}^E}{d\mathbb{P}} = \frac{1}{p^G} \exp \left(- \int_0^T \lambda_t dW_t - \frac{1}{2} \int_0^T |\lambda_t|^2 dt \right).$$

This gives a different initial wealth

$$w = W_0 = \frac{1}{\alpha} \left(\mathbb{E} \left[\ln \frac{\alpha p^G}{y} \middle| \mathcal{G}_t \right] - \frac{1}{2} \int_0^T |\lambda_t|^2 dt \right).$$

Though the running strategy is unchanged if we have a deterministic λ_t , which is given by

equation (4.21). The dual variable y is given by

$$y^* = \alpha \exp \left(-\alpha w - \mathcal{H}^G(\widetilde{\mathbb{Q}}^E, \mathbb{P}) \right),$$

which gives the value function for the informed investor

$$u^G(w - c) = -\exp \left(-\alpha(w - c) - \mathcal{H}(\mathbb{Q}^E, \mathbb{P}) \right) \mathbb{E}^G \left[\exp \left(-\mathbb{E}^G \left[\ln \frac{Z_T^G}{Z_T} \middle| \mathcal{G}_t \right] \right) \middle| \mathcal{G}_t \right],$$

where $\widetilde{\mathbb{Q}}^E \in \mathcal{M}_e^G$ is the minimal entropic martingale preserving measure for the informed.

For an uninformed investor taking the posterior expectation, the value function is defined by

$$\begin{aligned} u^G(w) &:= \sup_{\phi \in A} \mathbb{E} [U(W_T) | \mathcal{G}_t] \\ &= -e^{-\alpha w} \mathbb{E} [\exp(-\alpha(w + 1/\alpha [\mathcal{H}(\mathbb{Q}^E, \mathbb{P}) - \mathbb{E}^Q[\ln Z_T | \mathcal{F}_t]])) | \mathcal{G}_t] \\ &= -e^{-\alpha w} \mathbb{E} [Z_T \exp(-\mathcal{H}(\mathbb{Q}^E, \mathbb{P})) | \mathcal{G}_t] \\ &= -e^{-\alpha w - \mathcal{H}(\mathbb{Q}^E, \mathbb{P})}, \end{aligned}$$

due to the martingale preserving property.

We follow the argument of indifferent pricing as shown in equation (4.13) :

$$u^G(w^*) = u^G(w - c) \quad (4.22)$$

where w^* is the optimal initial wealth under $\mathbb{F}$. The cost c can be deduced at once

$$c^* = -\frac{1}{\alpha} \ln \mathbb{E}^G \left[\exp \left(-\mathbb{E}^G \left[\ln \frac{Z_T^G}{Z_T} \middle| \mathcal{G}_t \right] \right) \middle| \mathcal{G}_t \right]. \quad (4.23)$$

Remark: Since we are taking conditional expectation on the relative entropy that depends on $\mathcal{G}_t$, the cost will as a consequence depend on $\mathcal{G}_t$, that is, on the outcome of the random variable ξ_t .

4.10 Cost for extra noisy information

We are interested in analysing the gain of information from the update of information. Specifically, following from chapter 2, an aggregate of information processes can be repre-

sented in a form of one single effective information process. Hence we assume that

$$\mathcal{G}_t = \mathcal{F}_t \vee (\mathcal{H}_t = \sigma(\eta_s, 0 \leq s \leq t)).$$

The additional information ξ_t is given in the form of signal plus noise

$$\eta_t = \nu Y_T t + \beta_{tT} \quad (4.24)$$

where $\{\beta_{tT}\}$ is a Brownian bridge; ν is a random variable denoting the information flow rate and Y_T a random variable denoting the future cash flow, independent of β_{tT} and ν . Due to proposition B.1.5 and the Markovian property of this information process, the relative entropy can be written as

$$\mathbb{E} \left[Z_T^G \ln \frac{Z_T}{Z_T^G} \middle| \mathcal{H}_t \right] = \mathbb{E}[Z_T^G \ln p^G | \eta_t]. \quad (4.25)$$

Furthermore, the *a posteriori* distribution given $\mathcal{G}_t$ of Y_T is given by:

$$p^G = \frac{1}{\sqrt{2\pi t(T-t)/T}} \exp \left(-\frac{1}{2} \frac{(\eta_t - \nu Y_T t)^2}{t(T-t)/T} \right).$$

The entropy function in equation (4.25) becomes:

$$\begin{aligned} \mathbb{E}[Z_T^G \ln p^G | \xi] &= \frac{1}{2} \left(\ln \frac{T}{2\pi t(T-t)} - \frac{\mathbb{E}^G[(\eta_t - \nu Y_T t)^2 | \eta_t]}{t(T-t)/T} \right) \\ &= \frac{1}{2} \left(\ln \frac{T}{2\pi t(T-t)} - \frac{\eta_t^2 + \int (vyt)^2 \pi(y | \eta_t) dy - 2\eta_t \int vyt \pi(y | \eta_t) dy}{t(T-t)/T} \right). \end{aligned} \quad (4.26)$$

In the first equality, we have used the martingale preserving property that

$$\mathbb{E}^G[(\eta_t - \nu Y_T t)^2 | \eta_t] = \mathbb{E}^Q[(x - \nu Y_T t)^2] |_{x=\eta_t}.$$

By employing the Bayesian method, as shown in chapters 2, one can derive the *a posteriori* distribution $\pi(x | \eta_t)$ explicitly:

$$\pi(y | \eta_t) = \int \pi(y, v | \eta_t) dv = \frac{\int p(y) p(v) \exp \left[\frac{T}{T-t} (yv\eta_t - \frac{1}{2} y^2 v^2 t) \right] dv}{\int \int p(y) p(v) \exp \left[\frac{T}{T-t} (yv\eta_t - \frac{1}{2} y^2 v^2 t) \right] dv dy}.$$

Hence if we exponentiate (4.26) and take the conditional expectation under the enlarged filtration, we obtain

$$\begin{aligned}
& \mathbb{E}^{\mathbb{G}}[e^{\mathbb{E}[Z_T^G \ln p^G | \eta]} | \eta_t] \\
&= \mathbb{E}^{\mathbb{G}} \left[\sqrt{\frac{T}{2\pi t(T-t)}} \exp \left(-\frac{1}{2} \frac{(\eta_t - \int v y t \pi_\eta dy)^2}{t(T-t)/T} \right) \middle| \eta_t \right] \cdot \\
&\quad \exp \left(-\frac{Tt}{2(T-t)} \text{Var}(\nu Y_T | \eta_t) \right) \\
&= \sqrt{\frac{T}{2\pi t(T-t)}} \exp \left(-\frac{1}{2} \frac{(\eta_t - \int v y t \pi_\eta dy)^2}{t(T-t)/T} \right) \cdot \\
&\quad \exp \left(-\frac{Tt}{2(T-t)} \text{Var}(\nu Y_T | \eta_t) \right). \tag{4.27}
\end{aligned}$$

Hence the cost depends on the *a posteriori* mean and variance of the cash flow, denoted by $\mu_{\nu Y_T | \eta}$ and $\sigma_{\nu Y_T | \eta}^2$ respectively:

$$c_t^* = \frac{1}{\alpha} \left[\ln \left(\frac{2\pi t(T-t)}{T} \right) + \frac{T(\eta_t - \mu_{\nu Y_T | \eta})^2}{2t(T-t)} + \frac{Tt\sigma_{\nu Y_T | \xi_t}^2}{2(T-t)} \right]. \tag{4.28}$$

We observe similar behaviour in this function as in the one-period setting in the previous section. In particular, the cost is an increasing function of the “degree of surprise”, i.e. the *a posteriori* variance $\sigma_{\nu Y_T | \xi_t}^2$.

4.11 Discussion

In summary, we have modified the conventional approach to information pricing so as to obtain a pricing structure for information so that the pre-agreed price of information depends on the quality of information provided. This is reasonable because it is difficult for an information provider in reality to ensure the quality of information, on account of noise. One of the consequences is that the behaviour of the information cost is very different for the constrained and unconstrained scenarios. The main difference between the two scenarios is that when short selling is not allowed in the constrained case, downside information can only alert investors concerning possible losses (i.e. one cannot make profit out of downward price movements since short selling is forbidden). On the other hand, in the unconstrained scenario the investor can use downside information to create profit, which in turn generates a similar cost-information profile to upside information. It should be added that these characteristic behaviours are invariant under the overall scaling of the

initial wealth and the asset prices.

The results revealed here correspond to, at least to some extent, situations one might encounter in real-life contexts, even though the model setup is quite simple. In particular, it is interesting to draw attention, in the case where short selling is not allowed, to the fairly universal observation that (based on our model setup) when the expected loss of the investor amounts to approximately 100% of the initial wealth, investors are willing to opt out of any risky investment by spending the total initial wealth to purchase that information, so as to prevent the possibility of loss beyond the initial wealth.

Finally, we conclude this section by remarking that the existence of information asymmetries is not only fundamental to modern microeconomic theory [41] but also constitutes the basis for the existence of information providers, the roles of which are becoming increasingly more important in modern society (internet search engines, recommendations, etc.). Our main objective here has been to demonstrate in which way valuable information can be priced in a rational manner. Although the models used here to illustrate our pricing theory are simple, they nevertheless provide a useful conceptual guideline for evaluating information in a variety of contexts.

Conclusion

In this thesis, we have considered a variety of mathematical aspects of the role of informational inhomogeneities in financial markets. In particular, we have identified one of the likely sources for the stochasticity of asset-price volatilities as arising from uncertainties in information flow rates. We have been able to obtain closed-form expressions for asset price processes and associated derivative prices in the simplest modelling setups—in a context which is nevertheless sufficiently rich to be enable a fit to a wide range of volatility surfaces.

We have also addressed the issue of the assignment of a value or price to a given piece of information. Since there is no liquid market for information *per se*, we have in the present work adopted a utility indifference argument as a reasonable starting point to tackle this problem. Our point of departure, however, is to use the *a posteriori* measure, rather than the *a priori* measure used in the literature for this purpose, in the determination of the fair price of information. As a result, we have been able to obtain quality-dependent valuation formulae for information.

The problem of information valuation has been analysed in this thesis in the context of a financial market in which an information provider has access to a source of noisy or imperfect information concerning future returns, a source that is in particular adapted to a filtration *larger* than that accessible to general market participants. That this kind of setup is indeed reasonably realistic has been argued and reported in recent empirical studies—as described, for example, in [12] and [8]. The problem of information valuation is, nevertheless, not confined exclusively to financial markets as such. The “value” of an internet search engine, for instance, lies precisely in its ability to consistently provide information that users do not already have and do not otherwise have ready access to. Internet retailers likewise provide information in the form of recommendations and targeted advertisements, which in turn also lead to profits. It would be of interest to extend and apply our theory in broader settings such as these, since in the future the role of the value of information will undoubtedly be an issue of increasing importance.

Appendix A

Risk Neutral Measure for the Cost of Information

A.1 Presence of arbitrage in the Grossman-Stiglitz formula

The purpose of this section is to demonstrate, by using a somewhat extreme example, that according to the Grossman-Stiglitz formula for the cost of information it is possible to start with zero initial wealth such that with positive probability one can gain infinite return. Using the arguments presented in the paper by Grossman and Stiglitz [28], one finds that the terminal wealth of an investor with zero initial wealth is given by

$$W_1^* = -c^* \delta^{-1} + \varphi^*(s_1 - \bar{s}_0),$$

where we recall that φ^* is the *a posteriori* optimal strategy. If this quantity has finite expectation under $\mathbb{P}$, then clearly under any equivalent measure it will also have finite expectation. In what follows, however, we show that the $\mathbb{P}$ -expectation of W_1^* does not exist, i.e. it is infinity.

First we note that the posterior probability of s_1 , π_ξ can take the following values:

$$\begin{array}{ll} \pi_0 = 1 & \text{with probability } pq \\ \pi_1 = \frac{p(1-q)}{p(1-q) + q(1-p)} & \text{with probability } p(1-q) + (1-p)q \\ \pi_2 = 0 & \text{with probability } (1-p)(1-q). \end{array}$$

It is straightforward to show from (4.8) that the optimised strategy φ_ξ^* of the informed

investor takes the following values:

$$\varphi_0^* = -\infty \quad (\text{A.1})$$

$$\varphi_1^* = \frac{1}{\alpha} \ln \left[\frac{q(1-p)(1-\bar{s}_0)}{p(1-q)\bar{s}_0} \right] \quad (\text{A.2})$$

$$\varphi_2^* = +\infty. \quad (\text{A.3})$$

Therefore, we have the triplet $(\pi_\xi, s_1, \varphi^*)$ of random variables that can take the following values with the specified joint probability:

$$(\pi_\xi, s_1, \varphi^*) = \begin{cases} (\pi_0, 0, \varphi_0^*) & \text{with probability } pq \\ (\pi_1, 0, \varphi_1^*) & \text{with probability } p(1-q) \\ (\pi_1, 1, \varphi_1^*) & \text{with probability } q(1-p) \\ (\pi_2, 1, \varphi_2^*) & \text{with probability } (1-p)(1-q). \end{cases}$$

This gives

$$\begin{aligned} \mathbb{E}[\varphi^*(s_1 - \bar{s}_0)] &= pq[-\bar{s}_0\varphi_0^*] + (p-pq)[- \bar{s}_0\varphi_1^*] \\ &\quad (q-qp)[(1-\bar{s}_0)\varphi_1^*] + (1-p-q+pq)[(1-\bar{s}_0)\varphi_2^*] \\ &= \varphi_0^*[-pq\bar{s}_0] + \varphi_1^*[\bar{s}_0(2pq-q-p) + pq+q] \\ &\quad \varphi_2^*[(1-\bar{s}_0)(1-q)(1-p)]. \end{aligned}$$

Using equations (A.1) to (A.3) we see that the right side is infinity, hence $\mathbb{E}[W_1^*] = \infty$. This means that there exists no risk-neutral measure such that W_1^* has vanishing expectation value.

A.2 Finiteness of the terminal cost in discrete setting

Our objective here is to demonstrate, in the example just considered, that, in contrast to the Grossman-Stiglitz formula, according to our pricing formula the $\mathbb{P}$ -expectation of the terminal wealth W_1^* is finite, and hence admits the possibility of the existence of the risk-neutral measure. In our case, the cost c^* is itself a random variable dependent on ξ . Hence the expected terminal wealth is

$$\mathbb{E}[W_1^*] = -\mathbb{E}[c^*]\delta^{-1} + \mathbb{E}[\varphi^*(s_1 - \bar{s}_0)].$$

For the case with a discrete random variable $s_1 \in \{0, 1\}$ with probabilities $\{p, 1 - p\}$, we have already established that $\mathbb{E}[\varphi^*(s_1 - \bar{s}_0)] = \infty$, and this leads to the conclusion that the Grossman-Stiglitz model of information cost is not free of arbitrage. Our attention is now on the expectation

$$\mathbb{E}[c^*]\delta^{-1} = \frac{1}{\alpha} \mathbb{E} \left[\ln \frac{\pi_\xi e^{\alpha \bar{s}_0 \phi^*} + (1 - \pi_\xi) e^{\alpha(\bar{s}_0 - 1)\phi^*}}{\pi_\xi e^{\alpha \bar{s}_0 \varphi_\xi^*} + (1 - \pi_\xi) e^{\alpha(\bar{s}_0 - 1)\varphi_\xi^*}} \right].$$

We have the following relations:

$$0 \leq \bar{s}_0 \leq 1, \quad \varphi_\xi^* = \frac{1}{\alpha} \ln \left[\frac{(1 - \pi_\xi)(1 - \bar{s}_0)}{\pi_\xi \bar{s}_0} \right].$$

It follows that

$$\begin{aligned} \varphi_0^* &= -\infty \\ \varphi_1^* &= \frac{1}{\alpha} \ln \left[\frac{q(1 - p)(1 - \bar{s}_0)}{p(1 - q)\bar{s}_0} \right] \\ \varphi_2^* &= \infty, \end{aligned}$$

and that

$$\pi_0 = 1 \tag{A.4}$$

$$\pi_1 = \frac{p(1 - q)}{p(1 - q) + q(1 - p)} \tag{A.5}$$

$$\pi_2 = 0. \tag{A.6}$$

As before, we have

$$\begin{aligned} \mathbb{P}(\xi = 0) &= pq, \\ \mathbb{P}(\xi = 1) &= p(1 - q) + (1 - p)q, \quad \text{and} \\ \mathbb{P}(\xi = 2) &= (1 - p)(1 - q). \end{aligned}$$

Hence,

$$\begin{aligned}\mathbb{E}[c^*]\alpha &= pq \ln \left[\frac{\pi_0 e^{\alpha \bar{s}_0 \phi^*} + (1 - \pi_0) e^{\alpha(\bar{s}_0 - 1)\phi^*}}{\pi_0 e^{\alpha \bar{s}_0 \varphi_0^*} + (1 - \pi_0) e^{\alpha(\bar{s}_0 - 1)\varphi_0^*}} \right] \\ &+ (p(1 - q) + q(1 - p)) \ln \left[\frac{\pi_1 e^{\alpha \bar{s}_0 \phi^*} + (1 - \pi_1) e^{\alpha(\bar{s}_0 - 1)\phi^*}}{\pi_1 e^{\alpha \bar{s}_0 \varphi_1^*} + (1 - \pi_1) e^{\alpha(\bar{s}_0 - 1)\varphi_1^*}} \right] \\ &+ (1 - p)(1 - q) \ln \left[\frac{\pi_2 e^{\alpha \bar{s}_0 \phi^*} + (1 - \pi_2) e^{\alpha(\bar{s}_0 - 1)\phi^*}}{\pi_2 e^{\alpha \bar{s}_0 \varphi_2^*} + (1 - \pi_2) e^{\alpha(\bar{s}_0 - 1)\varphi_2^*}} \right],\end{aligned}$$

and using the equalities for π_0, π_2 (A.4) and (A.6) we get

$$\mathbb{E}[c^*]\alpha = pq\alpha\bar{s}_0(\phi^* - \varphi_0^*) \quad (\text{A.7})$$

$$+ (p(1 - q) + q(1 - p)) \ln \left[\frac{\pi_1 e^{\alpha \bar{s}_0 \phi^*} + (1 - \pi_1) e^{\alpha(\bar{s}_0 - 1)\phi^*}}{\pi_1 e^{\alpha \bar{s}_0 \varphi_1^*} + (1 - \pi_1) e^{\alpha(\bar{s}_0 - 1)\varphi_1^*}} \right] \quad (\text{A.8})$$

$$+ (1 - p)(1 - q)\alpha(\bar{s}_0 - 1)(\phi^* - \varphi_2^*). \quad (\text{A.9})$$

Let us denote the second term in the right side of (A.9) by $A(\varphi_1^*)$, which is a finite quantity.

We now have

$$\begin{aligned}\mathbb{E}[W_1^*] &= -\mathbb{E}[c^*]\delta^{-1} + \mathbb{E}[\varphi^*(s_1 - \bar{s}_0)] \\ &= -pq\bar{s}_0(\phi^* - \varphi_0^*) - A(\varphi_1^*)/\alpha - (1 - p)(1 - q)(\bar{s}_0 - 1)(\phi^* - \varphi_2^*) \\ &\quad - pq\bar{s}_0\varphi_0^* + B(\varphi_1^*) - (1 - p)(1 - q)(\bar{s}_0 - 1)\varphi_2^* \\ &= B(\varphi_1^*) - A(\varphi_1^*)/\alpha + \phi^*((1 - p)(1 - q)(1 - \bar{s}_0) - pq\bar{s}_0),\end{aligned}$$

where $B(\varphi_1^*) = \varphi_1^*[\bar{s}_0(2pq - p - q) + q - pq]$ is also finite. It follows that there exists at least one risk-neutral measure such that the wealth “process” W is a martingale.

To look for a risk neutral measure μ such that $\mathbb{E}^\mu[W_1^*] = 0$ let us start from the most elementary case. Let s_u and s_d denote the upper and lower values of the discrete random variable s_1 :

$$\mu s_d + (1 - \mu)s_u = \bar{s}_0.$$

Immediately, we can derive the risk-neutral measure for this simple case:

$$\mu = \frac{\bar{s}_0 - s_u}{s_d - s_u}.$$

This means s_1 is a μ -martingale. Hence in the simplest case when we do not have an extra

information and $W_0 = 0$:

$$\begin{aligned}\mathbb{E}^\mu[W_1^*] &= \mathbb{E}^\mu[\varphi^*(s_1 - \bar{s}_0)] \\ &= 0.\end{aligned}$$

Now, if we were given extra information as set out previously with a cost c , the risk neutral measure ν is the set of probabilities $\{r_i\}_{i=1,\dots,4}$ such that the following relation must hold:

$$\mathbb{E}^\nu[W_{1,c}^*] = 0$$

We now perform explicit calculations

$$\begin{aligned}\mathbb{E}^\nu[W_{1,c}^*] &= -\mathbb{E}^\nu[c^*]\delta^{-1} + \mathbb{E}^\nu[\varphi^*(s_1 - \bar{s}_0)] \\ &= \mathbb{E}^\nu \left[\ln(e^{\alpha(s_1 - \bar{s}_0)\varphi^*})^{1/\alpha} - \ln \left(\frac{\pi_\xi e^{\alpha\bar{s}_0\phi^*} + (1 - \pi_\xi)e^{\alpha(\bar{s}_0 - 1)\phi^*}}{\pi_\xi e^{\alpha\bar{s}_0\varphi_\xi^*} + (1 - \pi_\xi)e^{\alpha(\bar{s}_0 - 1)\varphi_\xi^*}} \right)^{1/\alpha} \right] \\ &= \frac{1}{\alpha} \mathbb{E}^\nu \left[\ln \left(\frac{\pi_\xi e^{\alpha s_1 \varphi^*} + (1 - \pi_\xi)e^{\alpha(s_1 - 1)\varphi^*}}{\pi_\xi + (1 - \pi_\xi)e^{-\alpha(s_1 - 1)\phi^*}} \right) \right] - \bar{s}_0 \phi^* \\ &= \frac{1}{\alpha} \left(r_1 \ln[1] + r_4 \ln[e^{\alpha\phi^*}] + r_2 \ln \left[\frac{\pi_1 + (1 - \pi_1)e^{-\alpha\varphi_1^*}}{\pi_1 + (1 - \pi_1)e^{-\alpha\phi^*}} \right] \right. \\ &\quad \left. + r_3 \ln \left[\frac{\pi_1 e^{\alpha\varphi_1^*} + (1 - \pi_1)}{\pi_1 + (1 - \pi_1)e^{-\alpha\phi^*}} \right] - \alpha\bar{s}_0\phi^* \right),\end{aligned}$$

where $\{r_i\}$ represents the joint probabilities of the random variables ξ and s_1 :

$$(\pi_\xi, s_1, \varphi^*) = \begin{cases} (\pi_0, 0, \varphi_0^*) & \text{with probability } r_1 \\ (\pi_1, 0, \varphi_1^*) & \text{with probability } r_2 \\ (\pi_1, 1, \varphi_1^*) & \text{with probability } r_3 \\ (\pi_2, 1, \varphi_2^*) & \text{with probability } r_4. \end{cases}$$

Continuing with the calculation by substituting values for π_1, φ_1^* and ϕ^* we have

$$\begin{aligned}
\alpha \mathbb{E}^\nu[W_{1,c}^*] &= (r_4 - \bar{s}_0) \ln \frac{(1-p)(1-\bar{s}_0)}{p\bar{s}_0} + r_2 \ln \frac{p(1-q) + p(1-q)\frac{\bar{s}_0}{1-\bar{s}_0}}{p(1-q) + pq\frac{\bar{s}_0}{1-\bar{s}_0}} \\
&\quad + r_3 \ln \frac{q(1-p)\frac{1-\bar{s}_0}{\bar{s}_0} + q(1-p)}{p(1-q) + pq\frac{\bar{s}_0}{1-\bar{s}_0}} \\
&= \ln \left[\left(\frac{(1-p)(1-\bar{s}_0)}{p\bar{s}_0} \right)^{(r_4-\bar{s}_0)} \right. \\
&\quad \times \left. \left(\frac{1-q}{1-\bar{s}_0} \frac{1}{(1-q) + q\frac{\bar{s}_0}{1-\bar{s}_0}} \right)^{r_2} \left(\frac{q(1-p)}{p\bar{s}_0} \frac{1}{(1-q) + q\frac{\bar{s}_0}{1-\bar{s}_0}} \right)^{r_3} \right].
\end{aligned}$$

As stated above, we require this expression to vanish for a suitable choice of $\{r_i\}$. Exponentiating, the problem reduces to finding a solution that satisfies the following relation:

$$\begin{aligned}
&\left(\frac{(1-p)(1-\bar{s}_0)}{p\bar{s}_0} \right)^{(r_4-\bar{s}_0)} \left(\frac{1-q}{1-\bar{s}_0} \frac{1}{(1-q) + q\frac{\bar{s}_0}{1-\bar{s}_0}} \right)^{r_2} \\
&\times \left(\frac{q(1-p)}{p\bar{s}_0} \frac{1}{(1-q) + q\frac{\bar{s}_0}{1-\bar{s}_0}} \right)^{r_3} = 1.
\end{aligned}$$

This is equivalent to

$$\begin{aligned}
&\left(\frac{(1-p)}{p\bar{s}_0} \right)^{(r_4+r_3-\bar{s}_0)} \left(\frac{1}{(1-q) + q\frac{\bar{s}_0}{1-\bar{s}_0}} \right)^{(r_2+r_3)} \\
&\times (1-\bar{s}_0)^{(r_4-r_2-\bar{s}_0)} \cdot (1-q)^{r_2} \cdot q^{r_3} = 1.
\end{aligned} \tag{A.10}$$

The factors cannot be further reduced. To demonstrate the existence of the solution, we simply construct one such that all exponents vanish:

$$\underline{r_2 = r_3 = 0},$$

which implies

$$\underline{r_4 = \bar{s}_0}.$$

Using $r_1 + r_4 = 1$ we have

$$\underline{r_1 = 1 - \bar{s}_0}.$$

Hence the risk-neutral measure ν is given by the probabilities $\{r_i\}$ as defined.

A.3 Gaussian case

In this section we demonstrate a similar no arbitrage property in our information cost model for the case when the terminal value of the risky asset is Gaussian distributed. From equation (4.15) the augmented utility indifference information cost is given by

$$c^* = \frac{1}{2} \alpha \delta \sigma_\xi^2 (\phi^* - \varphi^*)^2,$$

where the *a priori* and *a posteriori* trading strategies are, respectively,

$$\phi^* = \frac{\mu - \bar{s}_0}{\alpha \sigma^2},$$

and

$$\varphi^* = \frac{\mu_\xi - \bar{s}_0}{\alpha \sigma_\xi^2}.$$

We want to work out

$$\bar{c} = \int_{-\infty}^{\infty} c^*(\xi) d\mathbb{P}(\xi) \equiv \mathbb{E}^\xi[c^*]. \quad (\text{A.11})$$

and show that this indeed is finite, hence without arbitrage. This reduces to working out

$$\mathbb{E}^\xi[c^*] = \frac{1}{2} \alpha \delta \sigma_\xi^2 \cdot \mathbb{E}^\xi [(\phi^* - \varphi^*)^2]. \quad (\text{A.12})$$

Since ϕ is the *a priori* strategy, it does not depend on the extra information ξ . The main problem is to work out $\mathbb{E}^\xi[\varphi^*]$. As before, $\xi = S_T + \epsilon$, where both (S_T, ϵ) Gaussian random variables as defined in chapter 4, is a Gaussian random variable $\xi \sim N(\mu, \sigma^2 + \sigma_\epsilon^2)$. Also from chapter 4, we have

$$\mu_\xi = \frac{\mu \sigma_\epsilon^2 + \xi \sigma^2}{\sigma_\epsilon^2 + \sigma^2},$$

therefore, taking the conditional expectation we get

$$\begin{aligned} \mathbb{E}^\xi[\mu_\xi] &= \frac{\mu \sigma_\epsilon^2 + \mu \sigma^2}{\sigma_\epsilon^2 + \sigma^2} \\ &= \mu, \end{aligned}$$

where $\mathbb{E}^\xi$ denotes the expectation taken under the $\mathbb{P}(\xi)$ measure. Furthermore,

$$\begin{aligned}
 \mathbb{E}^\xi[\mu_\xi^2] &= \mathbb{E} \left[\frac{(\mu\sigma_\epsilon^2 + \xi\sigma^2)^2}{(\sigma_\epsilon^2 + \sigma^2)^2} \right] \\
 &= \mathbb{E} \left[\frac{\xi^2\sigma^4 + \mu^2\sigma_\epsilon^4 + 2\mu\xi\sigma^2\sigma^2\epsilon}{(\sigma_\epsilon^2 + \sigma^2)^2} \right] \\
 &= \frac{(\sigma^2 + \sigma_\epsilon^2 + \mu^2)\sigma^4 + \mu^2\sigma_\epsilon^4 + 2\mu^2\sigma^2\sigma^2\epsilon}{(\sigma_\epsilon^2 + \sigma^2)^2} \\
 &= \frac{\sigma^4(\sigma^2 + \sigma_\epsilon^2) + \mu^2(\sigma_\epsilon^2 + \sigma^2)^2}{(\sigma_\epsilon^2 + \sigma^2)^2} \\
 &= \frac{\sigma^4}{\sigma^2 + \sigma_\epsilon^2} + \mu^2.
 \end{aligned}$$

Consequently, we deduce that

$$\text{Var}^\xi[\mu_\xi] = \frac{\sigma^4}{\sigma^2 + \sigma_\epsilon^2}.$$

Therefore, the *a posteriori* expectation of the informed strategy is given by

$$\begin{aligned}
 \mathbb{E}^\xi[\varphi^*] &= \frac{\mathbb{E}^\xi[\mu_\xi] - \bar{s}_0}{\alpha\sigma_\xi^2} \\
 &= \frac{\mu - \bar{s}_0}{\alpha\sigma_\xi^2}.
 \end{aligned}$$

And the conditional expectation of the squared of the strategy is

$$\begin{aligned}
 \mathbb{E}^\xi[(\varphi^*)^2] &= \mathbb{E}^* \left[\frac{(\mu_\xi - \bar{s}_0)^2}{\alpha^2\sigma_\xi^4} \right] \\
 &= \frac{\mathbb{E}^*[\mu_\xi^2] + \bar{s}_0^2 - 2\bar{s}_0\mathbb{E}^*[\mu_\xi]}{\alpha^2\sigma_\xi^4} \\
 &= \frac{\frac{\sigma^4}{\sigma^2 + \sigma_\epsilon^2} + \mu^2 + \bar{s}_0^2 - 2\bar{s}_0\mu}{\alpha^2\sigma_\xi^4} \\
 &= \frac{1}{\alpha^2} \left(\frac{\sigma^2 + \sigma_\epsilon^2}{\sigma_\epsilon^4} + \frac{(\mu - \bar{s}_0)^2}{\sigma_\xi^4} \right).
 \end{aligned}$$

We are now in position to solve equation (A.12)

$$\begin{aligned}
 \mathbb{E}^\xi[c^*] &= \frac{1}{2}\alpha\delta\sigma_\xi^2 \cdot \mathbb{E}^\xi[(\phi - \varphi)^2] \\
 &= \frac{1}{2}\alpha\delta\sigma_\xi^2 \cdot (\phi^2 + \mathbb{E}^\xi[\varphi^2] - 2\phi\mathbb{E}^\xi[\varphi]) \\
 &= \frac{1}{2\alpha}\delta\sigma_\xi^2 \cdot \left(\frac{(\mu - \bar{s}_0)^2}{\sigma^4} + \frac{\sigma^2 + \sigma_\epsilon^2}{\sigma_\epsilon^4} + \frac{(\mu - \bar{s}_0)^2}{\sigma_\xi^4} - 2\frac{(\mu - \bar{s}_0)}{\alpha\sigma^2} \frac{(\mu - \bar{s}_0)}{\alpha\sigma_\xi^2} \right) \\
 &= \frac{1}{2\alpha}\delta\sigma_\xi^2 \cdot \left((\mu - \bar{s}_0)^2 \left(\frac{1}{\sigma^2} - \frac{1}{\sigma_\xi^2} \right)^2 + \frac{\sigma^2}{\sigma_\xi^2\sigma_\epsilon^2} \right).
 \end{aligned}$$

The average cost is a constant depending only on ξ . One can similarly, as in the discrete case, find a risk-neutral distribution such that $\bar{c} = 0$. Hence this suggests that there exists a risk neutral measure and therefore our information cost model is without arbitrage.

Appendix B

Filtration Enlargement

In the following, we shall state existing results on martingale theory and filtration enlargement. For detailed proofs we refer the reader to [3, 5, 21, 35, 46].

Let $S = (\{S_t\}_{0 \leq t \leq T})$ be a vector of locally bounded semi-martingales on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ equipped with a filtration $\mathbb{F} = \{\mathcal{F}_t\}_{0 \leq t \leq T}$ satisfying the usual conditions¹ and $\mathbb{T} = [0, T]$, $0 \leq T < \infty$. The probability measure $\mathbb{P}$ is understood to be the real measure, and the filtration $\{\mathcal{F}_t\}$ is generated by the càdlàg process S .

Definition B.0.1. A stochastic process X is said to be *càdlàg* if for each $\omega \in \Omega$, the path $X(\omega)$ is right-continuous and admits a left limit.

Definition B.0.2 (Semimartingale). A semimartingale is a càdlàg adapted process X having a decomposition in the form:

$$X = X_0 + M + A$$

where M is a càdlàg local martingale, and A is a adapted process with finite variation.

Definition B.0.3. The filtration $\mathbb{G} = \{\mathcal{G}_t\}_{t \in [0, T]}$ is called the initially enlarged filtration of $\mathbb{F}$, and is defined by

$$\mathcal{G}_t := \mathcal{F}_t \vee \mathcal{G}, \quad t \in [0, T]$$

where $\mathcal{G} = \sigma(\xi) \in \mathcal{F}$ is generated by some random variable ξ taking values in a general measurable space $(X, \mathcal{X})$. Without loss of generality, we choose $(X, \mathcal{X}) := (\Omega, \mathcal{G})$.

¹Right continuous ($\mathcal{F}_{t+} = \bigcap_{s \geq t} \mathcal{F}_s = \mathcal{F}_t$) and complete ($\mathcal{F}_0$ contains the negligible sets of $\mathcal{F}_T$).

We have now two filtrations, a background filtration $\mathbb{F}$ which is generated by the semimartingale S , and an initially enlarged filtration $\mathbb{G}$ given by some extra information ξ . We want to draw relations between the two filtrations, in particular, the semimartingale property relationship.

Theorem B.0.4 (Jacod). Suppose the following condition is satisfied: For every t , a regular conditional distribution of G given $\mathcal{F}_t$ exists and is $\mathbb{P}$ -almost surely equivalent to the law of G , i.e. for all $B \in \mathcal{X}$

$$\mathbb{P}[G \in B | \mathcal{F}_t](\omega) \sim \mathbb{P}[G \in B], \quad \text{for } \mathbb{P} - \text{almost anywhere } \omega \in \Omega.$$

Then every $(\mathbb{P}, \mathbb{F})$ -semimartingale remains a $(\mathbb{P}, \mathbb{G})$ -semimartingale on the interval $[0, T]$.

Consequently, we need not to distinguish between semimartingale adapted to $\mathbb{F}$ and $\mathbb{G}$. The condition in the previous theorem implies the existence of an $\mathcal{X} \otimes \mathcal{F}_t$ -measurable function $p : X \times \Omega \rightarrow (0, \infty)$ such that for P -almost any ω

$$\mathbb{P}[G \in B | \mathcal{F}_t](\omega) = \int_B p(x, \omega) \mathbb{P}[G \in dx]. \quad (\text{B.1})$$

We denote $p^G(\omega) := p(G(\omega), \omega)$ and $p^x(\omega) := p(x, \omega)$ for $\omega \in \Omega$. Note that each p^x is $\mathcal{F}_t$ -measurable and p^G is $\mathcal{G}_t$ -measurable.

Amendinger *et al.* [3] show that this condition is equivalent to the following:

Condition B.0.5 (Amendinger *et al.* [3]). There exists a probability measure $R \sim \mathbb{P}$ such that $\mathcal{F}_t$ and $\mathcal{G}$ are R -independent.

The R measure is called the decoupling measure and ensures the existence of the martingale preserving probability measure. We shall assume such measure exist.

Lemma B.0.6 (Amendinger *et al.* [3]). Suppose condition B.0.5 is satisfied. Then $\mathbb{G}$ satisfies under $\mathbb{P}$ the usual conditions of right-continuity and completeness.

B.1 Martingale preserving probability measure

Following the setting in [3], an immediate consequence of the decoupling measure is that we can introduce separate measures for different filtrations:

Lemma B.1.1 (Amendinger *et al.* [3]). Let P_1 and P_2 be probability measures on $\mathcal{F}_t$ and $\mathcal{G}$ respectively which are both equivalent to P . If condition B.0.5 holds, there is a **unique**

probability measure $\mu \sim P$ on $\mathcal{F}_t \vee \mathcal{G}$ such that $\mu = P_1$ on $\mathcal{F}_t$, $\mu = P_2$ on $\mathcal{G}$, and $\mathcal{F}_t$ and $\mathcal{G}$ are μ -independent. This is denoted by $P_{dec}(P_1, P_2) := \mu$.

Proof. For fixed t , $\mathcal{F}_t$ and $\mathcal{G}$ are R -independent. Define

$$\frac{d\mu}{dR} = \frac{dP_1}{dR} \Big|_{\mathcal{F}_t} \frac{dP_2}{dR} \Big|_{\mathcal{G}}.$$

Then $\mathbb{E}^\mu[\mathbf{1}_A \mathbf{1}_B] = P_1[A]P_2[B]$ for $A \in \mathcal{F}_t$ and $B \in \mathcal{G}$, using the R -independence of $\mathcal{F}_t$ and $\mathcal{G}$. $\square$

Definition B.1.2. We say $\mathbb{Q}$ is a risk neutral measure of S if $\mathbb{Q} \in \mathcal{M}_e(S)$, where

$$\mathcal{M}_e(S) = \{\mathbb{Q} \sim \mathbb{P} \text{ on } (\Omega, \mathcal{F}_T) : \text{discounted price of } S \text{ is a } \mathbb{Q}\text{-local martingale}\}, \quad (\text{B.2})$$

and the associated Radon-Nikodym density is denoted by

$$Z_T = \frac{d\mathbb{Q}}{d\mathbb{P}}.$$

For all $t \in [0, T]$, the restriction of $\mathbb{Q}$ to $\mathcal{F}_t$ is absolutely continuous with respect to the restriction of $\mathbb{P}$ to $\mathcal{F}_t$. Hence the process $\{Z_t\}_{t \in [0, T]}$ is a positive martingale.

Following from lemma B.1.1, we give a specific name for the decoupling measure that concerns with the risk neutral measure

Definition B.1.3 (Amendinger *et al.* [3]). Let $\mathbb{Q} \sim \mathbb{P}$ and let $\mathbb{Q}|_{\mathcal{F}_t}$ and $\mathbb{P}|_{\mathcal{G}}$ denote the restrictions on $\mathcal{F}_t$ and $\mathcal{G}$ respectively. The measure $\tilde{\mathbb{Q}} := P_{dec}(\mathbb{Q}|_{\mathcal{F}_t}, \mathbb{P}|_{\mathcal{G}})$ is called the *martingale preserving probability measure* (corresponding to $\mathbb{Q}$).

The measure $\tilde{\mathbb{Q}}$ is unique on $\mathcal{F}_t \vee \mathcal{G}$ and possesses the properties described in lemma B.1.1. Specifically, $(\mathbb{Q}, \mathbb{F})$ -martingale remains $(\tilde{\mathbb{Q}}, \mathbb{F})$ -martingale. Furthermore, the martingale property is preserved under $\tilde{\mathbb{Q}}$ in $\mathbb{G}$. Summarizing all the useful ingredients:

Corollary B.1.4 (Amendinger *et al.* [3]). Suppose condition B.0.5 is satisfied. Then,

1. We have $\mathcal{M}_e(S, \mathbb{Q}, \mathbb{F}) = \mathcal{M}_e(S, \tilde{\mathbb{Q}}, \mathbb{F}) \subseteq \mathcal{M}_e(S, \tilde{\mathbb{Q}}, \mathbb{G})$ and every semimartingale with respect to $(\mathbb{Q}, \mathbb{F})$ is also a semimartingale with respect to $(\mathbb{Q}, \mathbb{G})$.
2. Any $\mathbb{F}$ -adapted process L has the same distribution under $\tilde{\mathbb{Q}}$ and $\mathbb{Q}$. If L has in addition $(\mathbb{Q}, \mathbb{F})$ independent increments, i.e. $L_t - L_s$ is $\mathbb{Q}$ -independent of $\mathcal{F}_s$ for time $0 \leq s \leq t \leq T$, then L has also $(\tilde{\mathbb{Q}}, \mathbb{G})$ -independent increments, and the semimartingale characteristics of L are the same for $(\mathbb{Q}, \mathbb{F})$ and $(\tilde{\mathbb{Q}}, \mathbb{G})$.

3. Let S be a multidimensional $(\mathbb{P}, \mathbb{F})$ -semimartingale. Then an $\mathbb{F}$ -predictable process H is S -integrable with respect to $\mathbb{F}$ if and only if H is S -integrable with respect to $\mathbb{G}$, and the stochastic integrals of H with respect to S coincide for both filtration, i.e.

$$\int_{\mathbb{T}} H dS^{\mathbb{F}} = \int_{\mathbb{T}} H dS^{\mathbb{G}}.$$

Proposition B.1.5 (Amendinger *et al.* [3]). If the condition in theorem B.0.4 is satisfied then $\tilde{\mathbb{Q}} = P_{dec}(\mathbb{Q}, \mathbb{P})$ exists and is given by

$$\frac{d\tilde{\mathbb{Q}}}{d\mathbb{P}} = \frac{Z_T}{p^G}. \quad (\text{B.3})$$

i.e. for $A \in \mathcal{F}_t$ and $B \in \mathcal{X}$

$$\tilde{\mathbb{Q}}[A \cap \{G \in B\}] = \tilde{\mathbb{Q}}[A] \tilde{\mathbb{Q}}[\{G \in B\}] = \mathbb{Q}[A] \mathbb{P}[G \in B]$$

The proof can be found in [3, 26]. The idea is that Z_T/p^G is like Z_T a $\mathbb{P}$ -density, strictly positive. There exists a measure $\mu \sim \mathbb{P}$ defined by $d\mu := (Z_T/p^G)d\mathbb{P} = (1/p^G)d\mathbb{Q}$, together with the decoupling property we get

$$\begin{aligned} \mu[A \cap \{G \in B\}] &= \mathbb{E} \left[\frac{Z_T}{p^G} \cdot \mathbf{1}_{\{A \cap G \in B\}} \right] \\ &= \mathbb{E}^{\mathbb{Q}} \left[\mathbf{1}_A \mathbb{E}^{\mathbb{Q}} \left[\mathbf{1}_{G \in B} \frac{1}{p^G} \middle| \mathcal{F}_t \right] \right] \\ &= \int_A \mathbb{E}^{\mathbb{Q}} \left[\mathbf{1}_{G \in B} \frac{1}{p^G} \middle| \mathcal{F}_t \right] (\omega) \mathbb{Q}(d\omega). \end{aligned} \quad (\text{B.4})$$

By equation (B.1) we have

$$\mathbb{E}^{\mathbb{Q}} \left[\mathbf{1}_{G \in B} \frac{1}{p^G} \middle| \mathcal{F}_t \right] = \int_B \frac{1}{p(x, \omega)} p(x, \omega) P[G \in dx] = \mathbb{P}[G \in B]$$

Hence, equation (B.4) becomes

$$\mu[A \cap \{G \in B\}] = \mathbb{Q}[A] \mathbb{P}[G \in B].$$

And we take $\mu = \tilde{\mathbb{Q}}$ to represent such measure.

The martingale preserving measure allows us to determine a measure on the enlarged filtration $\mathbb{G}$ and hence standard theories of utility maximisation can be applied.

B.2 Strong predictable representation property

The following proposition from [3] shows that under condition B.0.5 the martingale representation property of S with respect to the filtration $\mathbb{F}$ and measure $\mathbb{Q}$ implies the same property with respect to the enlarged filtration $\mathbb{G}$ and the corresponding martingale preserving measure $\tilde{\mathbb{Q}}$.

Proposition B.2.1 (Amendinger *et al.* [3]). Suppose the general filtration $\mathbb{H} = (\mathcal{H}_t)_{t \in \mathbb{T}}$ satisfies the usual conditions and there is a probability measure $\mathbb{Q}^{\mathbb{H}} \sim \mathbb{P}$ such that $S \in \mathcal{M}_e(S, \mathbb{Q}^{\mathbb{H}}, \mathbb{H})$. Denote

$$\Gamma^{\mathbb{H}} := \left\{ \mathbb{Q} \sim \mathbb{Q}^{\mathbb{H}} \left| \frac{d\mathbb{Q}}{d\mathbb{Q}^{\mathbb{H}}} \text{ is } \mathcal{H}_T\text{-measurable, } \mathbb{Q} = \mathbb{Q}^{\mathbb{H}} \text{ on } \mathcal{H}_0, S \in \mathcal{M}_e(\mathbb{Q}, \mathbb{H}) \right. \right\}.$$

Then the following statements are equivalent:

1. $\Gamma^{\mathbb{H}} = \{\mathbb{Q}^{\mathbb{H}}\}$.
2. The set $\mathcal{M}_e(S, \mathbb{Q}^{\mathbb{H}}, \mathbb{H})$ of local $(\mathbb{Q}^{\mathbb{H}}, \mathbb{H})$ -martingales is equal to the set

$$\{\theta \cdot S \mid \theta \text{ is } S\text{-integrable with respect to } (\mathbb{Q}^{\mathbb{H}}, \mathbb{H}) \text{ in the sense of local martingales}\} \quad (\text{B.5})$$

of stochastic integrals with respects to S .

We say that S has the *strong predictable representation property* (P.R.P) with respect to $(\mathbb{Q}^{\mathbb{H}}, \mathbb{H})$ if one of these statements is valid.

Theorem B.2.2 (Amendinger *et al.* [3]). Suppose condition B.0.5 is satisfied and S has the strong predictable representation property with respect to $(\mathbb{Q}^{\mathbb{F}}, \mathbb{F})$ for some $\mathbb{Q}^{\mathbb{F}} \sim \mathbb{P}$. Let $\mathbb{Q}^{\mathbb{G}} = \tilde{\mathbb{Q}}^{\mathbb{F}}$ denote the martingale preserving probability measure corresponding to $\mathbb{Q}^{\mathbb{F}}$. Then S has the strong predictable representation property with respect to $(\mathbb{Q}^{\mathbb{G}}, \mathbb{G})$. For short $\Gamma^{\mathbb{F}} = \{\mathbb{Q}^{\mathbb{F}}\}$ implies $\Gamma^{\mathbb{G}} = \{\mathbb{Q}^{\mathbb{G}}\}$, or

$$(\mathbb{Q}^{\mathbb{F}}, \mathbb{F})\text{-P.R.P. implies } (\tilde{\mathbb{Q}}^{\mathbb{F}}, \mathbb{G})\text{-P.R.P.} \quad (\text{B.6})$$

The argument proves that the uniqueness of the equivalent $\mathbb{F}$ -martingale measure for S implies the uniqueness of the equivalent $\mathbb{G}$ -martingale measure in the sense of Proposition B.2.1.

Appendix C

Utility Optimisation in Incomplete Market via Convex Duality

C.1 No arbitrage condition

The notion of an incomplete market implies that there is no unique risk-neutral measure selected by the market as such. Let us define the set of equivalent martingale measures $\mathcal{M}_e(S)$ to be the set of equivalent measures that makes S a local martingale:

$$\mathcal{M}_e(S) = \{\mathbb{Q} \sim \mathbb{P} \text{ on } (\Omega, \mathcal{F}_T) : \text{discounted price of } S \text{ is a } \mathbb{Q}\text{-local martingale}\}. \quad (\text{C.1})$$

The existence of an equivalent local martingale measure is a sufficient and necessary condition for no arbitrage. We assume

$$\mathcal{M}_e(S) \neq \emptyset. \quad (\text{C.2})$$

The idea is that there exists at least one risk-neutral measure so that the asset price is a martingale. For any $\mathbb{Q} \in \mathcal{M}_e(S)$ and $H \in \mathcal{A}(S)$, the admissible set of controls, the lower-bounded integral $\int \alpha dS$ (bounded by $-w$) is a local martingale under $\mathbb{Q}$, thus a supermartingale. i.e.

$$\mathbb{E}^{\mathbb{Q}}[W_T] = \mathbb{E}^{\mathbb{Q}}[\int_0^T H_u dS_u] \leq 0.$$

Therefore equation (C.2) implies

$$\nexists \quad H \in \mathcal{A}(S), \quad \int_0^T H_t dS_t \geq 0 \quad \text{a.s. and } \mathbb{P} \left[\int_0^T H_t dS_t > 0 \right] > 0.$$

In other words, we cannot find an admissible portfolio strategy, which allows us, starting from a null capital, to reach almost surely at T a nonnegative wealth, with a non-zero probability of being strictly positive.

In the context of pricing kernel [14], this is equivalent to assuming the existence of a pricing kernel π_t s.t.

$$M_t = \pi_t S_t$$

is a martingale. This implies there exists a Radon-Nikodym derivative $\rho_t = P_t \pi_t = \frac{d\mathbb{Q}}{d\mathbb{P}}$, P_t a riskless bank account, such that we have $\mathbb{E}^{\mathbb{Q}}[P_t^{-1} S_t]$ a martingale.

C.2 Dual formulation

Define the conjugate function of $U(x)$ to be the Fenchel-Legendre transform:

$$V(y) := \sup_{x \in \mathbb{R}^+} [U(x) - xy], \quad y > 0.$$

We restrict $\text{dom}(V) = \{y > 0 : V(y) < \infty\}$. Note that the optimisation is now performed over a set of variables y instead of set of processes H . The conjugate function V is decreasing, differentiable with respect to y , strictly convex on $(0, \infty)$ with $\hat{U}(0) = U(\infty)$ and $\hat{U}' = -(U')^{-1} = -I$.

We define $I : (0, \infty) \rightarrow (0, \infty)$ the inverse function of U' on $(0, \infty)$

$$I = (U')^{-1},$$

which is strictly decreasing, and satisfies $I(0) = \infty$ and $I(\infty) = 0$.

Differentiating V with respect to y , we find the maximum is attained at $x^* = I(y) > 0$ i.e.

$$V(y) = U(I(y)) - yI(y), \quad y > 0$$

and we have the conjugate relation

$$U(x) = \inf_{y>0} [V(y) + xy], \quad x > 0$$

with an infimum attained at $y^* = U'(x)$.

For any $\mathbb{Q} \in \mathcal{M}_e$ define the change of measure martingale $Z = (Z_t)_{0 \leq t \leq T}$ by

$$Z_t := \left. \frac{d\mathbb{Q}}{d\mathbb{P}} \right|_{\mathcal{F}_t}.$$

The Radon-Nikodym derivative characterizes the the measure $\mathbb{Q} \sim \mathbb{P}$. Define the dual value function $v(y)$ by

$$v(y) := \inf_{\mathbb{Q} \in \mathcal{M}_e} \mathbb{E}V(yZ_T). \quad (\text{C.3})$$

We shall assume that this function is finite valued and that the solution of the primal problem can be obtained by solving the dual problem.

We assume there exists an unique dual optimiser attaining the infimum (C.3). Denote the dual optimiser by Z_T^* , the optimal terminal wealth by X_T^* and the optimal trading strategy by H_T^* .

Since we have S a $\mathbb{Q}$ -local-martingale bounded below, it is therefore a $\mathbb{Q}$ -supermartingale. We emphasize the following characteristic for X_T :

$$\mathbb{E}[Z_T X_T] \leq x, \quad \forall Z_T \in \mathcal{M}_e.$$

Hence

$$\begin{aligned} \mathbb{E}[U(X_T)] &\leq \mathbb{E}[V(yZ_T) + yZ_T X_T] \\ &\leq \mathbb{E}[V(yZ_T)] + xy \end{aligned}$$

Then $\forall x > 0$

$$u(x) = \sup_{X_T \in \mathfrak{X}_T^+} \mathbb{E}[U(X_T)] \leq \inf_{y>0, Z_T \in \mathcal{M}_e} \{\mathbb{E}[V(yZ_T)] + xy\} = \inf_{y>0} [v(y) + xy]. \quad (\text{C.4})$$

Hence we have the conjugate relationship:

$$v(y) = \sup_{x \in \mathbb{R}^+} [u(x) - xy], \quad y > 0.$$

The solution to the primal maximize problem $u(x)$ is

$$X_T^* = I(y^* Z_T^*) \quad i.e. \quad U'(X_T^*) = y^* Z_T^* \quad (C.5)$$

where $\hat{y} > 0$ is the solution to

$$\mathbb{E}[Z_T^* I(y Z_T^*)] = \mathbb{E}^{\mathbb{Q}^*}[I(y Z_T^*)] = x.$$

And the following properties holds true:

$$u'(x) = \mathbb{E}U'(X_T^*), \quad v'(y) = \mathbb{E}V'(y Z_T^*).$$

Proposition C.2.1 (Amendinger *et al.* [3]). Suppose the no-arbitrage condition is satisfied and there exists an $\mathcal{H}_0$ -measurable random variable $y^{\mathbb{H}}(x) : \Omega \rightarrow (0, \infty)$ with

$$\mathbb{E}^{\mathbb{H}}[I(y^{\mathbb{H}} Z_T^{\mathbb{H}*}) | \mathcal{H}_0] = x$$

where $\mathbb{H} = (\mathcal{H}_t)_{t \in \mathbb{T}} \in \{\mathbb{F}, \mathbb{G}\}$ and $\mathbb{E}^{\mathbb{H}}[\cdot]$ denotes the risk-neutral filtration in $\mathbb{H}$, and such that the process $X_t^{*\mathbb{H}} = \mathbb{E}^{\mathbb{H}}[I(y^{\mathbb{H}} Z_T^{\mathbb{H}*}) | \mathcal{H}_t]$, $t \in [0, T]$ satisfies $U^-(X_T^{*\mathbb{H}}) \in L^1(P)$. Then $X_T^{*\mathbb{H}}$ is the solution to the optimisation problem (4.19).

C.3 Exponential utility optimisation

Having acquired all the tools, we are in position to solve the optimisation problem (4.19). We have for the utility function of the form, $U(x) = -\exp(-\alpha x)$. Consider a stock price process follows a dynamics as described by equation (4.17), and define the progressively measurable process valued in $\mathbb{R}$:

$$\lambda_t = \sigma_t(\sigma_t \sigma_t^T)^{-1}(\mu_t - r_t \mathbf{1}), \quad 0 \leq t \leq T.$$

We assume that λ_t is bounded. The process λ_t is the risk-adjusted return of the asset, commonly known as the Sharpe ratio.

Under the Brownian filtration framework, we can give explicit description for the set of martingale probability measures $\mathcal{M}_e(S)$. Let us consider the set

$$K(\sigma) = \{\nu \in L_{loc}^2(W) : \sigma \nu = 0, \text{ on } [0, T] \times \Omega, \quad dt \otimes d\mathbb{P} \text{ a.e.}\}.$$

In words, the set $K(\sigma)$ contains Gaussian distributed random variables such that they are orthogonal to the volatility of the stock price process, almost everywhere. And ν chosen such that it is orthogonal to λ and satisfies the Novikov condition, i.e.

$$\mathbb{E} \left[\exp \left(\frac{1}{2} \int_0^T |\lambda_t|^2 + |\nu_t|^2 dt \right) \right] < \infty.$$

For any $\nu \in K(\sigma)$, the Radon Nikodym derivative is given by

$$Z_t^\nu = \exp \left(- \int_0^t (\lambda_u + \nu_u) dW_u - \frac{1}{2} \int_0^t |\lambda_u + \nu_u|^2 du \right), \quad 0 \leq t \leq T.$$

By Girsanov's theorem, the process

$$W_t^\nu = W_t + \int_0^t (\lambda_u + \nu_u) du$$

is a Brownian motion under $\mathbb{Q}$, where $\mathbb{Q} \sim \mathbb{P}$ with martingale density process Z_t^ν . The asset process under the $\mathbb{Q}$ dynamics

$$\begin{aligned} dS_t &= S_t (\mu_t dt + \sigma_t dW_t) \\ &= S_t r_t dt + S_t \sigma_t dW^\nu \end{aligned}$$

therefore the discounted price takes the following form:

$$\begin{aligned} d(\bar{S}_t) &= \bar{S}_t [(\mu_t - r_t) dt + \sigma_t dW] \\ &= \bar{S}_t \sigma_t dW^\nu \end{aligned} \tag{C.6}$$

is a $\mathbb{Q}$ -supermartingale.

We can derive some dual functions that will be useful for our calculation. Differentiating the utility function in equation (4.18) and then take the functional inverse we get

$$I(y) = (U'(y))^{-1} = -\frac{1}{\alpha} \ln \frac{y}{\alpha}.$$

The conjugate function with the exponential utility is therefore

$$V(y) = U(I(y)) - yI(y) = \frac{y}{\alpha} \left(\ln \frac{y}{\alpha} - 1 \right).$$

The dual optimiser can be obtained by minimising the dual value function (C.3):

$$\begin{aligned}
\mathbb{E}[V(yZ_T^\nu)] &= \mathbb{E} \left[\frac{yZ_T^\nu}{\alpha} \left(\ln \frac{yZ_T^\nu}{\alpha} - 1 \right) \right] \\
&= \mathbb{E}^\mathbb{Q} \left[V(y) + \frac{y}{\alpha} \left(- \int \lambda + \nu dW_t - \frac{1}{2} \int |\lambda + \nu|^2 dt \right) \right] \\
&= \mathbb{E}^\mathbb{Q} \left[V(y) + \frac{y}{\alpha} \left(- \int \lambda + \nu dW_t^\nu + \frac{1}{2} \int |\lambda + \nu|^2 dt \right) \right] \\
&= \mathbb{E}^\mathbb{Q} \left[V(y) + \frac{y}{2\alpha} \int |\lambda + \nu|^2 dt \right].
\end{aligned}$$

The last line is due to that W^ν is a $\mathbb{Q}$ -Brownian motion. The solution to the dual optimisation problem $v(y)$ is attained for $\nu = 0$, in particular independent of y . Denote the dual optimiser by Z_T^0 . Equation (C.5) implies that

$$\begin{aligned}
X_t^* &= \mathbb{E}[X_T^* | \mathcal{F}_t] = \mathbb{E} [Z_T^0 I(yZ_T^0) | \mathcal{F}_t] \\
&= \mathbb{E}^\mathbb{Q} \left[-\frac{1}{\alpha} \ln \frac{yZ_T^0}{\alpha} \middle| \mathcal{F}_t \right] \\
&= \mathbb{E}^\mathbb{Q} \left[\frac{1}{\alpha} \left(\ln \frac{\alpha}{y} + \int_0^T \lambda_t dW_t + \frac{1}{2} \int_0^T |\lambda_t|^2 dt \right) \middle| \mathcal{F}_t \right] \\
&= \frac{1}{\alpha} \left(\ln \frac{\alpha}{y} + \int_0^t \lambda_u dW_u^0 - \frac{1}{2} \int_0^T |\lambda_t|^2 dt \right),
\end{aligned}$$

where W_t^0 is a Brownian motion under the optimal measure $\mathbb{Q}$. The dual variable is determined by the initial condition $X_0 = x$, hence

$$x = \frac{1}{\alpha} \left(\ln \frac{\alpha}{y} - \frac{1}{2} \int_0^T |\lambda_t|^2 dt \right).$$

We get that the dynamic of the portfolio is given by

$$dX_t^* = \alpha^{-1} \lambda_t dW_t^0.$$

Comparing with the definition of the portfolio process

$$dX_t = H_t d\bar{S}_t = H_t \sigma_t \bar{S}_t dW_t^0,$$

we find the optimal asset allocation strategy

$$H_t^* \bar{S}_t = \frac{1}{\alpha} \sigma_t (\sigma_t \sigma_t^T)^{-1} \lambda_t = \frac{1}{\alpha} \sigma_t (\sigma_t \sigma_t^T)^{-1} \sigma_t (\sigma_t \sigma_t^T)^{-1} (\mu_t - r_t \mathbf{1}).$$

We remark that the result is identical to the one period setting as shown in Grossman and Stiglitz [28] if μ_t represents the return over the risk free level. The allocation strategy H^* is risk-adjusted, which means an asset with large variance will be penalized and consequently we should hold less of it. The asset price S_t acts as a normalising factor, so we will spend the same proportion of wealth if two stocks have different prices but same drift and volatility. In this case, we are allowed to short selling, because we can have a negative drift. To impose the condition of no short selling, we need to define the convex cone for the dual variable y , and solve the dual problem for $y = y(v)$ by applying the HJB equation. We will not go into detail of this in this section.

We can solve the primal problem (4.19) through the bidual relationship. Recall the dual value function (C.3)

$$\begin{aligned} v(y) &= \inf_{\mathbb{Q} \in \mathcal{M}} \mathbb{E}[V(yZ_T^\nu)] \\ &= \inf_{\mathbb{Q} \in \mathcal{M}} \mathbb{E} \left[\frac{yZ_T^\nu}{\alpha} \left(\ln \frac{yZ_T^\nu}{\alpha} - 1 \right) \right] \\ &= V(y) + \frac{y}{\alpha} \inf_{\mathbb{Q} \in \mathcal{M}} \mathbb{E}[Z_T^\nu \ln Z_T^\nu] \end{aligned}$$

Define the relative entropy between measures $\mathbb{Q}$ and $\mathbb{P}$:

$$\mathcal{H}(\mathbb{Q}, \mathbb{P}) = \mathbb{E} \left[\frac{d\mathbb{Q}}{d\mathbb{P}} \ln \frac{d\mathbb{Q}}{d\mathbb{P}} \right],$$

and $\mathbb{Q}^E = \arg \min_{\mathbb{Q} \in \mathcal{M}} \mathcal{H}(\mathbb{Q}, \mathbb{P})$ the minimal entropy measure, where the minimal measure $\mathbb{Q}|_{\mathcal{F}_T}$ is described by the change of measure martingale Z_T^0 . Therefore,

$$v(y) = V(y) + \frac{y}{\alpha} \mathcal{H}(\mathbb{Q}^E, \mathbb{P}).$$

The conjugate relation (C.4) gives

$$u(x) = \inf_{y \in \mathbb{R}^+} \left[V(y) + \frac{y}{\alpha} \mathcal{H}(\mathbb{Q}^E, \mathbb{P}) + xy \right].$$

The infimum of the dual variable y^* is attained at

$$y^* = \alpha \exp(-\alpha x - \mathcal{H}(\mathbb{Q}^E, \mathbb{P})).$$

Alternatively, in a non-diffusive manner, proposition C.2.1 yields

$$\begin{aligned}
 X_t^* &= \mathbb{E}^{\mathbb{Q}}[I(y^* Z_T^0) | \mathcal{F}_t] \\
 &= \mathbb{E}^{\mathbb{Q}} \left[-\frac{1}{\alpha} \ln [Z_T^0 \cdot \exp(-\alpha x - \mathcal{H}(\mathbb{Q}^E, \mathbb{P}))] \middle| \mathcal{F}_t \right] \\
 &= \mathbb{E}^{\mathbb{Q}} \left[-\frac{1}{\alpha} (\ln [Z_T^0] + \ln [\exp(-\alpha x - \mathcal{H}(\mathbb{Q}^E, \mathbb{P}))]) \middle| \mathcal{F}_t \right] \\
 &= x + \frac{1}{\alpha} (\mathcal{H}(\mathbb{Q}^E, \mathbb{P}) - \mathbb{E}^{\mathbb{Q}}[\ln Z_T^0 | \mathcal{F}_t]) .
 \end{aligned}$$

Using the identity:

$$\exp(\mathbb{E}^{\mathbb{Q}}[\ln Z_T^0 | \mathcal{F}_t]) = Z_t^0,$$

the primal value function is given by

$$\begin{aligned}
 u(x) &= \mathbb{E}[U(X_T^*)] \\
 &= \mathbb{E}[Z_T^0 \exp(-\alpha x - \mathcal{H}(\mathbb{Q}^E, \mathbb{P}))] \\
 &= -\exp(-\alpha x - \mathcal{H}(\mathbb{Q}^E, \mathbb{P})) .
 \end{aligned}$$

This gives us the value function of the portfolio optimisation problem for the uninformed investor. Indeed one could explicitly write out the Entropy function:

$$\mathcal{H}(\mathbb{Q}^E, \mathbb{P}) = \frac{1}{2} \int_0^T |\lambda_t|^2 dt.$$

References

- [1] G.A. Akerlof. The market for ‘lemons’: Quality uncertainty and the market mechanism. *Quarterly Journal of Economics*, 84(3):488–500, 1970.
- [2] F. Allen and D. Gale. Stock-price manipulation. *Review of Financial Studies*, 5:503–529, 1992.
- [3] J. Amendinger, D. Becherer, and M. Schweizer. A monetary value for initial information in portfolio optimization. *Finance and Stochastics*, 7:29–46, 2003.
- [4] T. Anderson, T. Bollerslev, F.X. Diebold, and C. Vega. Real-time price discovery in stock, bond and foreign exchange markets. *Journal of International Economics*, Vol 73:251–277, 2007.
- [5] S. Ankirchner, S. Dereich, and P. Imkeller. The Shannon information of filtrations and the additional logarithmic utility of insiders. *The Annals of Probability*, 34:743–778, 2006.
- [6] T. Bjork. *Arbitrage Theory in Continuous Time, third edition*. OUP Oxford, 2009.
- [7] F. Black and M. Scholes. The pricing of options and corporate liabilities. *Journal of Political Economy*, 81:637–654, 1973.
- [8] J. Bollen, H. Mao, and X. Zeng. Twitter mood predicts the stock market. *Journal of Computational Science*, 2:1–8, 2011.
- [9] D. Brigo, F. Mercurio, and F. Rapisarda. Lognormal-mixture dynamics and calibration to market volatility smiles. *Working paper*.
- [10] D. Brigo, F. Mercurio, and F. Rapisarda. Smile at the uncertainty. *Risk Magazine*, May:97–101, 2004.
- [11] D.C. Brody. Information and finance. London Graduate School in Mathematical Finance Lecture Notes, 2008.

- [12] D.C. Brody, J. Brody, B.K. Meister, and M.F. Parry. Outsider trading. 2010.
- [13] D.C. Brody, M.H.A. Davis, R.L. Friedman, and L.P. Hughston. Informed traders. *Proceedings of Royal Society A*, Vol. 465, No. 2104:1103–1122, 2009.
- [14] D.C. Brody and L.P. Hughston. Chaos and coherence: a new framework for interest-rate modelling. *Proceedings of Royal Society A*, 460:85–110, 2004.
- [15] D.C. Brody, L.P. Hughston, and A. Macrina. Beyond hazard rates: a new framework for credit-risk modelling. *Advances in Mathematical Finance: Festschrift Volume in Honour of Dilip Madan*, Birkhauser, 2007.
- [16] D.C. Brody, L.P. Hughston, and A. Macrina. Dam rain and cumulative gain. *Proceedings of Royal Society A*, 464 (2095):1801–1822, 2008.
- [17] D.C. Brody, L.P. Hughston, and A. Macrina. Information-based asset pricing. *International Journal of Theoretical and Applied Finance*, 11:107–142, 2008.
- [18] M.H.A. Davis. Option pricing in incomplete markets. in *Mathematics of Derivative Securities*, 1997. Cambridge University Press.
- [19] F. Delbaen and W. Schachermayer. A general version of the fundamental theorem of asset pricing. *Mathematische Annalen*, 300:463–520, 1994.
- [20] D. Easley and M. O’Hara. Price, trade, size, and information in securities markets. *Journal of Financial Economics*, 19:69–90, 1987.
- [21] A. Es-Saghouani. Enlargement of filtration and insider trading. Master’s thesis, Korteweg-de Vries Institute for Mathematics, Netherlands, 2006.
- [22] R.A. Fisher. Theory of statistical estimation. *Proceedings of the Cambridge Philosophical Society*, 22:700–725, 1925.
- [23] M. Fujisaki, G. Kallianpur, and Kunita H. Stochastic differential equations for the nonlinear filtering problem. *Osaka Journal of Mathematics*, 9:19–40, 1972.
- [24] I.M. Gel’fand and A.M. Yaglom. Calculation of the amount of information about a random function contained in another such function. *Uspekhi Matematicheskikh Nauk*, 12:3–52, 1957.
- [25] J. Gleick. *The Information: A History, a Theory, a Flood*. Fourth Estate, Mar 2011.

- [26] A. Grorud and M. Pontier. Insider trading in a continuous time market model. *International Journal of Theoretical and Applied Finance*, 1:331–347, 1998.
- [27] S.J. Grossman. *The Informational Role of Prices*. The MIT Press, 1989.
- [28] S.J. Grossman and J.E. Stiglitz. On the impossibility of informationally efficient markets. *The American Economic Review*, Vol. 70, No. 3:393–408, Jun., 1980.
- [29] J.M. Harrison and D.M. Kreps. Martingales and arbitrage in multiperiod securities markets. *Journal of Economic Theory*, 20(3):381–408, 1979.
- [30] J.M. Harrison and S.R. Pliska. Martingales and stochastic integrals in the theory of continuous trading. *Stochastic Processes and Their Applications*, 11:215–260, 1981.
- [31] F.A. Hayek. The use of knowledge in society. *The American Economics Review*, Vol. XXXV No. 4:519–530, September, 1945.
- [32] S.L. Heston. A closed-form solution for options with stochastic volatility with applications to bond and currency options. *The Review of Financial Studies*, Vol 6 No. 2:327–343, 1993.
- [33] J. Hull and A. White. The pricing of options on assets with stochastic volatilities. *Journal of Finance*, Vol. 42, Issue 2:281–300, 1987.
- [34] P. Jackel. Stochastic volatility models: Past, present and future. Presentation slides.
- [35] J. Jacod. *Calcul Stochastique et Problmes de Martingales. Lecture Notes in Mathematics*. Springer-Verlag, 1979.
- [36] G. Kallianpur and C. Striebel. Estimation of stochastic systems: Arbitrary system process with additive white noise observation errors. *The Annals of Mathematical Statistics*, Vol. 39, No. 3:785–801, 1968.
- [37] I. Karatzas and S.E. Shreve. *Methods of Mathematical Finance*. Springer, 1998.
- [38] A.Y. Khintchin. The concept of entropy in the theory of probability. *Uspekhi Matematicheskikh Nauk*, 8:3–20, 1953.
- [39] D. Kramkov and W. Schachermayer. The asymptotic elasticity of utility functions and optimal investment in incomplete markets. *The Annals of Applied Probability*, 9:90–950, 1999.

- [40] A.S. Kyle. Informed speculation with imperfect competition. *The Review of Economic Studies*, 56:317–355, 1989.
- [41] K. Lofgren, T. Persson, and J.W. Weibull. Markets with asymmetric information: The contributions of george akerlof, michael spence and joseph stiglitz. *The Scandinavian Journal of Economics*, Vol. 104, No. 2:pp. 195–211, Jun., 2002.
- [42] A. Macrina. *An Information-based framework for asset pricing: X-factor theory and its application*. PhD thesis, King’s College London, 2006.
- [43] F. Mercurio. A multi-stage uncertain-volatility model. 2002.
- [44] M. Monoyios. Utility indifference pricing with market incompleteness. *Nova Science Publishers*, 2008.
- [45] J. Peress. Information vs. entry costs: What explains U.S. stock market evolution? *The Journal of Financial and Quantitative Analysis*, 40:563–594, 2005.
- [46] H. Pham. *Continuous-time Stochastic Control and Applications with Financial Applications*. Series Stochastic Modeling and Applied Probability, Vol 61, Springer, 2009.
- [47] C.R. Rao. Information and accuracy attainable in the estimation of statistical parameters. *Bulletin of the Calcutta Mathematical Society*, 37:81–91, 1945.
- [48] E. Renault and N. Touzi. Option hedging and implied volatilities in a stochastic volatility model. *Mathematical Finance*, Vol. 6 No. 3:279–302, July 1996.
- [49] A.N. Shiryaev. *Essentials of stochastic finance Facts, Models, Theory*. World Scientific, 1999.
- [50] A.M. Spence. Job market signaling. *Quarterly Journal of Economics*, 87(3):355–374, 1973.
- [51] E.M. Stein and J.C. Stein. Stock price distributions with stochastic volatility: An analytic approach. *The Review of Financial Studies*, Vol. 4, No. 4:727–752, 1991.
- [52] G.J. Stigler. The economics of information. *The Journal of Political Economy*, Vol LXIX, No. 3:pp 213–225, June 1961.
- [53] R.E. Verrecchia. Information acquisition in a noisy rational expectations economy. *Econometrica*, 50:1415–1430, 1982.

-
- [54] N. Wiener. *Cybernetics or Control and Communication in the Animal and the Machine*. Cambridge, Mass.: The Technology Press, 1948.
- [55] N. Wiener. *The Human Use of Human Beings*. Boston: Houghton Mifflin Co., 1954.
- [56] W.M. Wonham. Some applications of stochastic differential equations to optimal nonlinear filtering. *Journal SIAM Control*, Vol. 2, No. 3:347–369, 1965.