Lancaster University Management School
Working Paper
1998/006

Some Recent Developments in Capital Market Theory: A Survey

R C Stapleton

The Department of Accounting and Finance
Lancaster University Management School
Lancaster LA1 4YX
UK

Some recent developments in capital market theory: a survey

Richard C. Stapleton
The Management School, Lancaster University

January 28, 1998

Abstract

This paper surveys some recent developments in the theory of capital markets. Particular emphasis is given to two strands of the literature. The first covers some recent and fundamental extensions to the theory of risk aversion and the demand for risky assets. These papers are concerned with the effect of non-hedgeable background risk on risk attitudes. The important implications for finance are for the size of the risk premium (the equity premium puzzle) and for the demand for and pricing of contingent claims. For example, background risk may help to explain the apparent over-pricing of options on equity indices.

The second topic is interest rate term structure models. Stochastic term structure models try to capture the possible future shapes of the term structure of interest rates. This is relevant for the pricing of contingent claims, in particular for the pricing of interest rate derivatives such as American-style swaptions. The paper will survey the most important recent models in the literature, each of which satisfies the fundamental no-arbitrage property. It will discuss the implications of the models for the pricing of both European-style and American-style options.

1 Capital Market Theory: Introduction

Capital market theory is about the pricing of assets, principally corporate assets. It holds a central place in the structure of finance theory. This is because finance studies the behaviour of corporations in issuing securities and investing in assets, and as a consequence studies the pricing of securities in capital markets. The cornerstones of capital market theory: the Capital Asset Pricing Model (CAPM), the Modigliani-Miller theorems (MM), and the Black-Scholes Option Pricing Model (OPM) were all in place twenty five years ago. Before I begin to discuss more recent developments in the field and current research activity, let me briefly outline these three paradigms and their relationship to each other.

Consider a firm whose securities, stock and bonds, are traded on the stock market. Suppose, for example that the current total value of the firm is 1,000 million, made up of an equity stock market capitalisation of 700 million, and bonds with a value of 300 million. The role of asset pricing models in finance is to explain why the asset value of the firm is 1,000, given it's exogenously given expected cash flow. In particular, the CAPM attempts to explain the required expected rate of return on the firm's assets, given the beta of the firm. The MM theorems on the other hand tell us how the total value of the firm will be affected if the firm recapitalised itself by issuing more debt and repaying some of its equity. Thirdly, since equity can be modelled as a call option on the firm's assets, the OPM can be used to explain why the equity of the firm is selling for 700 million given the volatility of the underlying assets.

Of course, much work has been done extending the three basic paradigms over the last few years. Particular progress has been made in the area of option pricing and especially in the pricing of derivatives whose payoffs depend on interest rates. The Black-Scholes model was initially applied to options on stocks, and then to options on foreign exchange and bonds. However, straightforward applications were found not to be appropriate in this latter case, since the assumptions of a constant volatility, lognormal process simply do not apply. At the same time the hedging of interest rate risk with options has become commonplace in practice. Hence a great deal of academic work has been undertaken, adapting the OPM to cope with the pricing of interest rate related contingent claims. In order to price more complex derivatives with path-dependent payoffs, good models of the underlying term structure of interest rates are required. The most important developments in this area will be surveyed in the section 3 of this paper.

The foundations of the CAPM are expected utility theory and the theory of state contingent pricing. The former is based on Von Neumann-Morgenstern utility, and the subsequent work of Pratt (1964) who defined various concepts of risk aversion and showed that, when faced with the choice of a risky and a riskless asset, more risk averse investors buy less of the risky asset. To an extent, the foundations of asset pricing have been broadened in recent years following the contribution of Harrison and Kreps (1979), who showed that a no-arbitrage economy implies the existence of an 'equivalent martingale measure' (EMM). The significance of the EMM lies in the fact that the price of any asset is the expectation of it's payoff under the EMM. This has allowed many of the utlity based theories in finance to be generalised to a no-arbitrage economy. However, it is perhaps the case that this has led more to a change in the appearance of many papers in finance than to a significant change in content. Today we tend to speak of the *pricing kernel* whereas in the past we would have used the utility function of the representative investor. We will see this use of the EMM when we review interest rate models in section 3.

I regard the border line between economics and finance in terms of the type of models employed as well as the subject matter. Economics tends to use equilibrium models whereas finance employs no-arbitrage techniques. The MM, OPM, and the APT (Arbitrage Pricing Theory) models, as well as much of the theory of International Finance, are prime examples of no-arbitrage theories. However, it should be remembered that these models are themselves embedded in some equilibrium economy. In most cases the propositions of finance rely on the fundamental work of economists. In this context it is interesting to note that some of the foundations of utility theory and risk-taking behaviour have been questioned by economists in recent work. The particular issue that I will emphasise in this survey is the effect of secondary, non-hedgeable risks on investors' attitude to the risk of marketed assets. These non-hedgeable risks are often termed *background* risks, and typical among them is labor income risk. It appears that the existence of background risk can have profound implications for risk-taking, and for the pricing and use of financial securities. The recent literature and its implications for finance theory is surveyed in section 2.

2 Background Risk and its Implications for Finance

How is the attitude towards taking a risk affected by the presence of a second, non-hedgeable, risk? For example, how is an investor's aversion towards the risk of an investment in the market portfolio affected by the risk of his or her labor income? Attempts to answer this rather fundamental question has led to an important extension of the theory of risk bearing, which has particular relevance to capital market theory. Finance theory is based on the ideas of Von Neumann-Morgenstern expected utility maximisation and on the Pratt-Arrow definition of risk aversion. The CAPM for example assumes that investors are risk averse in the Pratt-Arrow sense, and derives an equilibrium in which the beta of a company's stock determines its cost of capital. A set of early papers by Kihlstrom, Romer and Williams (1981) (KRW), Ross(1981), and by Nachman, (1982) look at the question of whether the original conclusions of Pratt (1964) are preserved when a second, non-hedgeable, *background risk* is present. KRW and Nachman introduce the idea of a derived utility function for the investor who faces a marketed risk in the presence of a second, non-hedgeable risk. They then show that an investor who is risk averse in the Pratt (1964) sense, remains risk averse in the presence of background risk. Ross, in a related paper, however argues that, when comparing risks, the Pratt (1964) concept of risk aversion requires strengthening. He shows that when two individuals each face two risks, one of which is insurable, the more risk averse individual, in the Pratt-Arrow sense, will not necessarily be prepared to pay a higher insurance premium than the less risk averse individual. The problem is that Ross' proposed condition on utility functions is so strong that few well-known utility functions satisfy it. In this survey, I will concentrate on a set of recent papers that have looked for an answer to the following question. When is it true that the addition of a background risk increases the derived risk aversion of an agent? Also, from the point of view of finance theory, if the derived risk aversion of agents facing background risk increases, how does this affect their demand for capital assets, and the pricing of those assets? We start, however, with a discussion of some examples of background risk and discuss their relevance in finance.

2.1 Examples of Background Risk

The most straightforward and important example concerns portfolio choice in the presence of labor income risk. Assume that the choice facing individuals is to invest a proportion of available wealth in risky stocks and the remainder in risk-free bonds. Consider the investment strategy of two university professors, one of whom has tenure, and one of whom does not. Suppose they have the same wealth and tastes. Which of the professors will invest more in the risky asset portfolio? In this example the labor income risk of the professor with no tenure is the background risk and the risky stocks are the marketed assets.

As a second example, consider the hedging decisions of an entrepreneur who faces an insurable foreign exchange or interest rate risk. Suppose that the entrepreneur is also subject to other production risks which are uninsurable. If these uninsurable risks increase, will the entrepreneur be prepared to pay more for insurance against the insurable risks? In this example the production risk that the entrepreneur faces is the background risk, whereas the foreign exchange risk is the marketed risk.

Now consider the decision facing a portfolio manager who chooses from a set of marketed claims, including stocks, bonds, and derivatives. Suppose that he or she maximises their personal utility, given a compensation package that includes two elements: the first depends on the performance of the portfolio, the second depends on its relative performance compared to the performance of other specified portfolio managers. The second element creates an uninsurable background risk for the manager, since he does not know what other managers will do. The question is, how will the portfolio choice of the manager be affected by the compensation package offered? In this case the risk of the portfolio is the marketed risk, and the risk taken by other managers is the background risk.

What are the main implications of the analysis of background risk for finance? Perhaps the most dramatic effect is on the pricing of risky assets and the size of the market risk premium. If background risk is faced by investors in equities, then their derived risk aversion will exceed the risk aversion of their utility functions. This could help to explain the equity premium puzzle, as suggested by Weil (1992). Second, if background risk is significant this could affect the demand for insurance on those risks that are marketed. In particular, the existence of background risk could explain, partly at least, the demand for and pricing of options. This line of enquiry is pursued by Franke, Stapleton and Subrahmanyam (FSS) (1997a),(1997b). Thirdly, if background risk is significant for portfolio managers and if put options on the market portfolio are overpriced, then this could explain the underperformance of professional fund managers as well as the puzzling herd behaviour that has been noted. In sum, the extensive existence of incomplete markets for background risks has the potential for solving at least three major puzzles in finance: the equity premium puzzle, the demand for options (a redundant asset in many equilibrium models), and the herd-like behaviour and seeming underperformance of portfolio managers.

2.2 The Effect of Background Risk on Risk Taking

Kimball (1990) considers the following multiperiod problem. How much will a consumer save out of period 1 income, given that period 2 income is uncertain. This problem is similar, but somewhat simpler to the background risk problem discussed above, with period 2 income taking on the role of the background risk. In this problem, a prudent individual would save more out of period 1 income, the more risky was the period 2 income. In finance, we are familiar with the concept of risk aversion, having applied Pratt's concept of absolute risk aversion to the portfolio problem. A more risk averse individual, i.e. one with greater absolute risk aversion, invests less in a risk asset than a less risk averse individual. Similarly, we can define a concept of absolute prudence, such that an individual with greater absolute prudence saves more in the above multiperiod problem. If $u(w)$ is the utility function, define $u'(w)$, $u''(w)$, and $u'''(w)$ as the first three derivatives of the function. The Pratt-Arrow measure of absolute risk aversion is $a(w)$ in

$$a(w) = \frac{-u''(w)}{u'(w)} \tag{1}$$

The definition of absolute prudence is given, analogously, as $p(w)$ in

$$p(w) = \frac{-u'''(w)}{u''(w)} \tag{2}$$

Most utility functions commonly used in finance have the property of declining absolute risk aversion, or more formally $a'(w) \leq 0$. (For example, the CPRA utility function, $u(w) = w^{\gamma}, \gamma < 1$ has $a(w) = (1-\gamma)/w$.) Also, most have the property of declining absolute prudence, $p'(w) \leq 0$. (For example, in the case of CPRA, $p(w) = (2-\gamma)/w$). It turns out that the combination of declining absolute risk aversion and declining absolute prudence is a critical determinant of an individual's response to background risk.

Kimball,(1993)

Kimball's paper is one of a set of recent contributions, each of which looks at the impact of an independently distributed background risk, on the risk-taking behaviour of an individual agent. Other relevant papers are Gollier and Pratt (GP) (1996), Pratt and Zeckhauser, (PZ) (1987), Eeckhoudt, Gollier, and Schlesinger (EGS) (1996), and Franke, Stapleton and Subrahmanyam (FSS)(1997a). The difference between these papers is that each asks a slightly different question. All ask the question, what is the effect of a background risk on the derived utility function. However, the exact nature of the background risk differs in each case. To be precise, suppose that wealth $w = x + y$ where x is the payoff on a portfolio of investments and y is an independent background risk. Then we can write the derived utility of x as $\nu(x)$ in

$$\nu(x) = E_y[u(w)] \tag{3}$$

Kimball along with GP, PZ, EGS, and FSS look at the effect of the background risk y on the function $\nu(x)$ and compare its properties with those of $u(x)$. The simplest question to ask (but not to answer) is the following; when will the addition of a zero-mean, independent background risk make $\nu(x)$ more risk averse than $u(x)$? Kimball shows that 'standard risk aversion', which is equivalent to $a'(w) \leq 0$ and $p'(w) \leq 0$, is a sufficient condition. It is important to note that many of these papers refer to 'unfair' background risks rather than simply zero-mean background risks (see for example PZ, GP and EGS). There appears to be little economic logic, however, in including as risks, negative incomes, that occur with certainty. In fact GP make a convincing case for considering zero-mean (GP, p1110) as opposed to unfair risks. It is convenient to assume unfair risks, however, because this rules out utility functions with increasing absolute risk aversion. Kimball actually considers the wider set of risks which raise expected marginal utility. He then shows that for this set of background risks, the derived risk aversion increases, if and only if, agents are standard risk averse. The set of risks considered is important as can be seen from the analysis of EGS, who define such a broad set that they require utility functions with properties not satisfied by most commonly used functions such as the HARA class. The paper by GP, p1117-1119 gives a clear summary of the different sets of background risks that have been considered, and the consequential restrictions on utility functions which lead to more risk averse behaviour towards the marketed risk.

Standard risk aversion is a reasonable restriction. It implies that reduced wealth (or increased background risk) increases both risk aversion and prudence. If utility is standard, the addition of zero-mean background risk increases derived risk aversion, and reduces the demand for a marketed risky asset. FSS (1997a) show also that, when we consider the demand for contingent claims on a risky asset, the slope of the function relating x to the aggregate marketed claim X falls if agents are standard risk averse. This is a rather more general manifestation of more risk averse behaviour. Also, standard risk aversion is assumed by Weil (1992) in his analysis of the equity premium puzzle. It appears then that although there are wider sets of utility functions that will guarantee 'more risk averse behaviour' in the presence of background risk, it is reasonable to proceed using the intuitively plausible assumption that utility is standard.

2.3 The Implications of Background Risk for Financial Theory

Weil,(1992)

The first and perhaps the most important application of background risk in finance is the subject of Weil's paper. Weil provides a simple model of asset prices in which he determines both the risk-free rate and the expected return on equities (on the one traded risky asset). In Weil's model, the market is incomplete due to the existence of non-insurable labor income risk, and agents have utility functions which exhibit declining absolute risk aversion and prudence. He shows first of all (Proposition 1), that the risk free rate will be lower in the background risk economy, compared to the no-background risk economy, due to the presence of precautionary saving. Weil's main result is then presented in his Proposition 3. This is a straightforward implication of Kimball's (1993) analysis of the effect of background risk on derived risk aversion. If agents are standard risk averse, the equity premium, which is the difference between the risky asset's expected return and the risk free rate, is higher in the background risk economy.

Weil's paper is noteworthy also from a methodological standpoint. He compares two economies, one with and one without background risk, by introducing a calibrator. The calibrator, an economist for example, is assumed to be unaware of the existence of background risk. Weil shows that such a calibrator will *overpredict* the risk-free rate and *underpredict* the risk premium. In order to match the observed prices with a model of equilibrium, the calibrator needs to over estimate the risk aversion of the representative investor. Using a similar argument, FSS (1997b) show that a calibrator will find all put and call options overpriced in a background risk economy.

Weil establishes his Proposition 3 in a two-stage argument. First he shows that the absolute risk aversion of the derived utility function (with background risk) exceeds that of the original utility function. This also follows, given standard risk aversion, from Kimball (1993). Next, he shows that this leads to the higher risk premium. He then provides an important collorary. If utility is of the HARA class, with declining absolute risk aversion, then utility is standard. In conclusion, Weil gives a numerical example, using CPRA, where the calibrator underestimates the risk premium by a factor of two.

Eeckhoudt and Kimball (EK) (1992)

EK apply the ideas of Kimball (1993) on standard risk aversion and the effect on risk taking to optimal insurance behaviour. The question is as follows; how will an agent react to the presence of background risk, if it is possible to buy insurance against another risk? They determine the optimal coinsurance rate, which in the insurance literature is the term used for the proportion of the risky asset covered by forward contracts. Since this problem is formally the same as choosing the proportion of investment in a single asset, when faced with the choice of a risk-free and a risky asset, it is not surprising that standard risk aversion is sufficient for the coinsurance rate to be lower in the background risk economy. EK then consider two extensions of this basic analysis. First, they show that the effect of background risk on the demand for insurance is re-inforced when there is a positive correlation between the marketed risk and the background risk. Finally, they deal with insurance in the conventional, finance sense, where the payout on the insurance is akin to an option with a particular exercise price. In the insurance context the exercise price is related to the deductible of the insurance contract. EK consider the problem where the agent has a fixed coinsurance rate, and chooses a deductible, D. They show that the optimal deductible is lower (i.e. more insurance is purchased), when background risk exists than when it does not exist. Again they assume standard risk aversion.

Franke, Stapleton and Subrahmanyam (FSS)(1997a)

FSS consider the effect of background risk on the demand and supply of options. In their setting, individual agents buy contingent claims on a marketed asset (the market portfolio), and also face a non-insurable background risk. Utility is HARA, which implies that utility is also standard in the sense of Kimball. FSS work with the concept of the precautionary premium (Kimball (1990)), and show that, in equilibrium, agents have non-linear sharing rules. For example, FSS (1997a) Theorem 2 shows that, in an economy where agents face background risk, a particular agent who has no background risk will have a concave sharing rule. Such an agent will be a supplier of options.

The existence of background risk provides a role for non-linear payoff contracts (options), as opposed to linear payoff contracts (forwards), in hedging behaviour. It also has implications for the pricing of these non-linear contracts. FSS (1997a) Theorem 3 shows that a general increase in background risk in the economy leads to an increase in the price of put options and a fall in the price of call options. However, in Theorem 4 they also show that all put and call options are over priced in an economy with background risk compared to their prices in a no-background risk economy calibrated to have the same forward price for the marketed asset. Using Weil's style of argument, a calibrator using a model such as the Black-Scholes model and ignoring background risk, would underprice all options in the background risk economy.

2.4 Background Risk and Finance: The State of the Art

Finance specialists should be aware of the significance of background risk to the theory of risk bearing. Having been sceptical about the importance of the issue, my view now is that the consequences of incomplete markets are of fundamental importance to finance. Although finance theory has been well served with the complete markets, expected utlity maximisation paradigm, it requires some modification in the light of background risk. As Weil has shown quite convincingly, if labor income is non-insurable, the implications for the risk premium can be significant. Also, as FSS show, the linear sharing rule, central to finance theory, becomes a non-linear sharing rule in the presence of background risk.

The basic work on extending the theory of risk bearing to multiple risks and incomplete markets is now in place. However, much work is still required on applications to finance problems. For example, Weil raises the possibility that the equity premium is increased by labor income risk, but we have very little idea of the magnitude of the effect, either from a theoretical or an empirical point of view. Also, although FSS analyse the effect of background risk on the sharing rules of agents, we have no complete model of the demand for stocks, bonds and options in a multi-asset, background risk economy. The work on the insurance implications of background risk is also somewhat preliminary in nature. For example, in practice agents have to choose simultaneously both the deductible and the level of coinsurance. Available theoretical models only allow the choice of one variable, holding the other variable constant. While some progress has been made in the case of HARA functions by FSS, work is required on the applications to insurance of their general contingent-claims model. Also, the work of FSS on the effect of background risk on the demand, supply and pricing of options requires both further theoretical analysis and empirical support.

3 Models of the term structure and the pricing of interest rate derivatives.

The investigation of the term structure of interest rates has provided an exciting area for research in recent years. This has been stimulated, to an extent, by the spectacular growth in the use of interest rate derivatives to hedge interest rate risk. I will survey what I regard as the most important theoretical developments and try to present a structure for catergorising the various models. One of the main difficulties facing researchers and practitioners is the multiple purposes for which the models have been built. For example, there seem to me to be at least five different reasons for building term-structure models in Finance. At the highest (or purist) level we may as financial economic theorists be interested in explaining the level and stochastic movement through time of real interest rates. I will refer to these as level 1 models. These models, of which the most celebrated in the recent literature is the Cox, Ingersoll, and Ross (1985) (CIR) model, require a general equilibrium approach. At a second level we may take the production side of the economy as given, as it is in most models in capital market theory, and derive the term structure given assumptions regarding the behaviour of the pricing kernel. This approach has been taken recently, for example by Constantinides (1992) in his derivation of a theory of nominal interest rates. At a third level, there is a long pedigree of models which assume a stochastic process for the spot interest rate and derive bond prices. One of the first and most important of these is the Vasicek (1977) model. Although Vasicek's analysis is quite general, he illustrates it with an example which assumes that the interest rate evolves as a mean reverting normally distributed variable. Models in this category are very much in line with the models of capital market equilibrium (such as the CAPM)which normally take interest rates as exogenous. My fourth level includes most of the recent work in the area and includes models whose purpose is to derive the prices of interest rate derivatives. Of course, models in categories one, two and three can also be used to price derivatives, however the analysis in these cases is neither required nor entirely satisfactory if the pricing of derivatives is the major purpose. The models in category four apply the methodology, originally developed by Black and Scholes (1973) and Cox, Ross, and Rubinstein (1979) for the case of stock options, to the problem of valuing options (both European and American) on bonds. One of the best known models in this category was published by Ho and Lee (1986). These models take the current term structure as given (as the current stock price is taken as given by Black and Scholes) and build an arbitrage-free evolution of interest rates or bond prices (as CRR do for stock prices). Finally, at the fifth (and lowest) level, there is a set of models whose purpose is to engineer a tree of interest rates which reflects observed prices of interest rate derivatives. This tree may be required for purposes of risk management, or to price other exotic or path-dependent derivatives. Such models are built to reflect both the current term structure and the implied volatilities of caps, floors and swaptions. The most well known model in this category is the model of Black, Derman, and Toy (1990). To summarise my categorisation of interest rate models is as follows:-

Level 1. General equilibrium models of interest rates

Level 2. Partial equilibrium models of interest rates

Level 3. Models of bond prices, given interest rates

Level 4. Models of interest rate derivatives, given bond prices

Level 5. Models of interest rates, consistent with derivative prices

In the following survey I will concentrate most on the papers that deal with level 4 models, since most recent work has been in the area of no-arbitrage models and the pricing of interest rate derivatives. Level 1,2, and 3 models are of interest since they lay the foundation for much of this recent work. Level 5 models are also of interest since future developments are likely to follow the lead given by these examples.

Another difficulty in studying term-structure theory arises from the widespread use of continuous time models and notation by many of the originators of the models. Although in some cases the continuous time methodology makes a significant contribution to the analysis, it often serves only to cloud the issues involved. In this survey I will attempt where possible to illustrate the models with discrete time examples and notation. In this regard I am indebted to the prior work of Backus (1994) and Backus and Zin (1993). Also, I have benefited greatly from being able to read the prior survey papers by Strickland (1993) and by Subrahmanyam (1996).

3.1 General equilibrium models of interest rates

The Cox, Ingersoll, and Ross (1985) (CIR) model

The CIR model of the dynamic behaviour of the term structure is derived using an intertemporal, general equilibrium asset pricing model. It is assumed that production technology changes randomly over time and that the utility function of the identical investors is time additive. Further, to obtain a simple closed form solution, CIR assume that the utility function is logarithmic. In their model, the real interest rate, like the technology, follows a single-factor square root process of the (discrete) form

$$r_{t+1} = r_t + \alpha(\theta - r_t) + \sigma r_t^{0.5} \Delta z \quad (4)$$

where r_t is the (continuously compounded) interest rate at time t, α controls the speed at which the rate reverts towards a fixed 'target' rate θ, σ is a constant and Δz is a normally distributed innovation with zero mean and unit variance. Equation (4) is the discrete form of the well known continuous 'square root process', CIR eq (17).

The CIR 'square root' process can be applied to value bonds (see CIR, eq(23)), and options on bonds (see CIR, eq(32)), however, it should be noted that both formulae contain the potentially unmeasurable market price of risk term. While this is a common feature when bond prices are derived from an interest rate process, it is not a characteristic of most option pricing models. Also, the model is a model of real interest rates as opposed to nominal rates. Here lies one of the disadvantages of the CIR model. The CIR square root process is attractive for nominal rates, partly because it precludes negative interest rates. However, the model is one of real rates, and real rates can and do become negative. The CIR model has been extended to nominal rates in an economy where money is neutral, however this is not entirely satisfactory.

A feature of the general equilibrium approach is that very strong assumptions have to be made to obtain simple solutions. For example, time additive logarithmic utility is such a strong assumption that it implies myopia and reduces the investors' problem in a multiperiod world to a series of single period problems. This together with the square root technology assumption guarantees the result that the interest rate will follow the process in equation (4). My own view is that it is normally wise to separate the 'economics' question of what type of general equilibrium can support a given interest rate process from the 'finance' question of how to value bonds and options given an interest rate process.

The Longstaff and Schwartz (1992) (LS) model

One extension of the CIR approach which has produced an interesting and significant result is the two-factor model of LS. Here LS assume that two factors, each following a square root process generate the production in the economy. However, because the rate of interest and the variance of the rate of interest both depend on the two factors, these two abstract factors can be replaced by the short rate and the variance of the short rate. The model has the potential to explain the term premium.

The LS model provides a general equilibrium in which both the short rate and its volatility depend on the two factors in a rather complex manner, as shown in LS equations (14) and (15). They then derive bond prices and the yield on zero-coupon bonds. The term premium shown in LS equation (23) is the key result. It shows that the term premium collapses to zero if the market price of interest rate risk λ is zero in their model. It follows that the LS model cannot provide a two-factor model of interest rates in a risk-neutral world (or under the risk-neutral measure). It provides a possible explanation of that part of the term premium that is related to the volatility of interest rates. It cannot provide an explanation, for example, for inflation related premia, unless these are generated by the uncertainty of inflation, rather than by the expected inflation itself.

3.2 Partial equilibrium models of interest rates

The Constantinides (1992) model

The above discussion of the CIR and LS models reveals that it is an ambitious project to construct a general equilibrium, in which interest rates and then bond prices, are derived. A more modest aim would be to build a model of interest rates starting with a representative agent economy and making a direct assumption about the stochastic evolution of the agent's marginal utility. This partial equilibrium approach could then lead to a realistic process for interest rates. The Constantinides model is based on this idea but uses a more general analysis by making assumptions regarding the pricing kernel. We know from Harrison and Kreps (1979) that a pricing kernel exists if and only if the no-arbitrage condition holds, a weaker assumption than the assumption of a representative agent economy. Constantinides asserts that a pricing kernel exists and assumes that it evolves as a 'squared Gaussian process'. This guarantees that the pricing kernel is a positive variable, and a further parameter restriction guarantees that the nominal interest rate (which is the rate that Constantinides models) is also a non-negative variable. Constantinides then derives bond prices, the yield to maturity curve and the term premium in his model. He shows that even a single-factor model can produce complex shapes of the term structure.

The approach used by Constantinides of modelling the pricing kernel is an interesting one. Backus (1993) uses the same approach to investigate a whole range of models. However I feel that this analysis is little more than pedagogic. There seems little difference to my mind between making an assumption about the pricing kernel that leads to a given interest rate process than assuming the interest rate process in the first place. It seems rather strong therefore when Constantinides titles his paper "A *Theory* of the Nominal Term Structure of Interest Rates". However, Constantinides does provide a new and potentially fruitful line of approach. At worst this paper might convince us that it is better to start our analysis at the next, lower, level.

3.3 Models of bond prices, given interest rates

The Vasicek (1977) model

Vasicek's model is one of the first, and perhaps the most quoted, of a set of models which assume a stochastic process for the evolution of the interest rate and derive the evolution of bond prices and the term structure. In its most general form, the short term interest rate follows a process whose discrete time equivalent is

$$r_{t+1} = r_t + f(t, r_t) + \sigma(t, r_t)\Delta z \tag{5}$$

where r_t is the (continuously compounded) interest rate at time t, $f(t, r_t)$ is the expected change in the rate, σ is a function which determines the volatility and Δz is a normally distributed innovation with zero mean and unit variance.

Vasicek assumes that the price of a bond with a maturity $T > t$ is determined only by the spot rate over the period $(t - T)$. He then uses an arbitrage argument (similar to that used in the Arbitrage Pricing Model, APT) to establish that there is a unique market price of risk across bonds of all maturities. It follows that the bond price for a particular maturity must follow the differential equation, Vasicek, equation (15). A specific example, which is used by Vasicek to illustrate his general technique, is often referred to as 'the Vasicek model'. In this case the short rate follows a mean reverting process, somewhat similar to that of CIR. The discrete equivalent of the process is

$$r_{t+1} = r_t + \alpha(\theta - r_t) + \sigma\Delta z \tag{6}$$

where all the variables are defined in equation (4).

When the short rate process is given by (6), the short rate is normally distributed and bond prices are lognormally distributed. It is not surprising therefore that under these assumptions bond options can be priced using the Black-Scholes model with a volatility input which reflects the rate of mean reversion α. Using (6), Vasicek is able to obtain a closed form solution for the price of all bonds and the term structure.

A number of criticisms have been levelled at the Vasicek model over the years, and some extensions have been suggested. Many of these are really comments on Vasicek's mean-reverting, Gaussian example. For example, a typical comment is that the model does not exclude negative interest rates, and hence could fail an arbitrage-free test if interpreted as a model of nominal rates. However, for reasonable parameter inputs the probability of negative rates is insignificant. Another comment is that the mean reversion parameter is a constant, and cannot reflect the current term structure. Hull and White (1990) proposed an 'extended Vasicek model'. This provides an interesting extension of Vasicek's example, allowing time varying mean reversion. The proposed process is

$$r_{t+1} = r_t + \alpha(t)(\theta(t) - r_t) + \sigma \Delta z. \tag{7}$$

This can hardly be called an extension of Vasicek's paper, since equation (7) is a special case of equation (5). In fact, the contribution of this extension of Vasicek's example is not that it extends our knowledge of bond pricing. The contribution is that it shows how to build a term structure model which reflects existing bond prices and which can be used to value interest rate dependent contingent claims.

Another 'extension' of the Vasicek model is due to Dothan (1978). Dothan assumes that the interest rate is lognormally distributed. Again this is really just another example of the general Vasicek solution. However Dothan's model has provided the basis for various lognormal interest rate contingent claim valuation models.

3.4 Models of interest rate derivatives, given bond prices

The Ho and Lee (1986) (HL) model

The HL model was the first no-arbitrage model of the term structure. Its purpose is to value interest rate dependent contingent claims given the existing term structure. The prior models had all been asset pricing models whose principal aim was to determine the term structure. The key idea in the HL paper is that the discount function (the price of unit zero-coupon bonds of all maturities) is given today, and should move through time in an arbitrage free manner. The HL model of interest rates describes the stochastic evolution of the discount function under the EMM, not under the actual measure. As in the case of the Black-Scholes model for the valuation of stock options, only the EMM is required if the purpose is to value interest rate derivatives. Under the EMM, the no-arbitrage condition is very simple. At each node, the forward price of any bond must equal the expected value of its price at the next time period, using the pseudo probabilities of the EMM.

In order to understand the idea of modelling interest rates under the EMM, I found it helpful to refer to the closely related work of Heath, Jarrow and Morton (HJM) (1990a), (1990b), and (1992). The formalities are dealt with in HJM (1992), where the conditions for the existence and uniqueness of an EMM are derived and where it is shown that derivatives can be valued using the EMM. However, the exact condition that must hold for the term structure to be arbitrage free is best understood by referring to HJM (1990b), p.58-64. The HL model is a single factor model, where the discount function follows a binomial process. At each node on the tree, all the bonds are therefore perfectly correlated (in terms of logarithms). HL constrain the tree to be recombining (referred to, confusingly in my view, as the path independent condition) and this implies that the volatility of the short rate is constant in their model. The crucial no-arbitrage condition is clearly stated in HL, equation (10), as a restriction on pseudo probability π.

The HL model is log-binomial in the price of the zero-coupon bonds. Interest rates in the model have a continuous time limit such that they are normally distributed with constant variance. The rates are calibrated to the current structure of forward rates in the market, which is a feature shared by most subsequent models in the literature. The resulting process for the short rate is of the form:

$$r_{t+1} = r_t + f(t) + \sigma \Delta z \tag{8}$$

Here, $f(t)$ reflects the current term structure of forward rates, but does not allow mean reversion (since it is not a function of r_t), and σ is a constant implying normality of the rates.

The purpose of the HL model is to value contingent claims. The Lemma in HL, equation (33) indicates that claims can be valued recursively, starting at the expiry date of the claim and working back state-by-state to the current time. As HL state in discussing the practical use of their model 'interest rate claims are priced (in their model) *relative* to the term structure. It should be noted that the true value of the HL model is in the valuation of American-style claims (i.e. claims with path-dependent payoffs), since European-style claims can be valued under these, and indeed under more general, assumptions using the Black-Scholes and related models.

The HL model is limited by rather strong assumptions. However, most of these have been relaxed by HJM. The most important limitations are the assumption of normality of the short rate, the constant volatility, and the lack of mean reversion, assumptions which make the allowable interest rate processes far less general than those in Vacicek, for example. However, the significance of the paper is reflected by the fact that since the publication of the HL model, few models are proposed that are not built using the no-arbitrage methodology.

The Heath, Jarrow and Morton (HJM) (1990a), (1990b), and (1992) model

Since the publication of the three papers, the HJM model has become the premier model in the literature for the valuation of interest rate derivatives. The general equilibrium model of CIR and the partial equilibrium models of the Vasicek (1977) and Brennan and Schwartz (1982)type have as their principle purpose the valuation of bonds and the determination of the term structure. These models are neither necessary nor sufficient for the valuation of interest rate contingent claims. They require knowledge of the market price of risk, which, as HL and HJM show is not required for the valuation of options. Hence the models' intermediate step of valuing bonds is not necessary. Also, the models are not sufficient for the accurate valuation of contingent claims, since they do not (except in the case of the extended Hull and White model) take into account the current term structure observable in the market. The HJM model's popularity stems from the fact that it is like a generalised Black-Scholes model which applies to bond options of the American as well as the European variety.

In order to appreciate the HJM model I recommend starting with HJM (1990b), concentrating mainly on HJM (1990a), and if you wish to research in this area, move on to HJM (1992) which formalises the argument and establishes the continuous time limiting case. HJM (1990b) states clearly that the purpose of the model is to generalise the HL model in a number of directions. My view is that HJM generalise HL in the following significant ways:-

- Continuous time limit of HL is derived
- Multiple risk factors are included
- Non-normally distributed interest rates are allowed for

The significance of the continuous time limit of HL is that in the limit it is clear that only one parameter, the volatility σ has to be estimated. This is because, in the limit, when the interest rate follows an additive brownian motion, the EMM has the same volatility as the actual distribution. However, the more important generalisations, which make the model of potential practical significance, are the extension to multiple risk factors and non-normal rates. Most participants in the market for interest rate derivatives use a model for pricing European-style options which implies that interest *rates* are lognormally distributed. Hence they wish to build term-structure models with similar properties. Also the pricing of interest rate caps and floors and swaptions implies that more than one factor is required to describe the process for interest rates.

HJM achieve their goal of generalising HL by the clever technique of modelling forward rates rather than spot rates or bond prices. In their notation $f(t, T)$ is the forward rate at t for delivery at T of a short-term loan. In the single-factor case, for example, they assume that

$$f(t+1,T) = f(t,T) + \delta(t,T) + \sigma(t,T,f(t,T))\Delta z \tag{9}$$

In the general two-factor case, the process is described clearly in HJM (1990a), equation(5). For the special case where the two factors evolve as a three-state process, refer to HJM (1990b), equation (15). The HJM argument for modelling forward rates is that forward rates are likely to be stationary whereas bond prices are not. However, the advantages of modelling forward rates are not as clear as HJM feel, as pointed out by Hull in his published comment on HJM (1990b). Also, it is interesting to note that it is only because they chose to model forward rates, defined as continuously compounded rates, that the particular adjustment term δ appears in equation (9).

The HJM model is undoubtedly a large step forward, even considering the fact that the HL model contains many of its features. However, there are some questions regarding the HJM approach, which leave interesting issues for future research. At a theoretical level, it is not at all clear what a given forward rate process implies. Since HJM require the process for a forward with a fixed date T, the process is likely to have non-stationary variance. It is more transparent to start with an assumption regarding the spot rate as in the Hull and White extended Vasicek model. Also, it is not clear that it is optimal to model the continuously compounded interest rate. It is far simpler to model the behaviour of the 'banker's discount rate' as in Stapleton and Subrahmanyam (1997) or to model the annual yield rate. Also there is the question of implementation of the model. Since from an academic point of view, the only point of the HL or HJM models is to price claims whose payoff is path dependent, then it has to be shown that the HJM model can be implemented efficiently. There are problems, first in the single factor case when volatility changes over time, and secondly there are severe computational problems in evaluating path-dependent claims in multi-factor models.

3.5 Models of interest rates, consistent with derivative prices

The Black, Derman, and Toy (1990) (BDT) model

The BDT model takes as inputs the current term structure (as in HL and HJM) and the current term structure of yield volatilites. The output of the model is a single-factor stochastic term structure with short rates which limit to the lognormal distribution. Yield volatilities in the BDT model are volatilities of the yield on zero-coupon bonds of different maturities. The general idea behind the model is to produce a term structure model, for the pricing of path-dependent contingent claims for example, which is consistent with current market data from the pricing of European-style options. At one level, the difference between the BDT approach and that of HL and HJM is between what the models take as inputs and outputs. In principle, it should be possible to build a HJM model, measure the yield volatilities as outputs, and iterate until these are the same as the market volatilities taken as given in the BDT model.

BDT describe a procedure for building a tree of short rates which is consistent with no arbitrage and which limits in distribution to the lognormal, i.e. it is a multiplicative binomial tree. This assumption has the advantage of excluding the possibility of negative interest rates, and more importantly, of being consistent with the pricing of European-style interest rate options. As in HL, the tree of rates has the re-combining property, that precludes mean reversion of the rates. The main problem with the model is the somewhat bizarre outputs that can result from the input of seemingly reasonable data on yield volatilities. As in the case of the HJM model, it is not clear that BDT have the inputs and outputs the right way around. The Hull and White approach of modelling rates by adapting the Vasicek model is perhaps more promising.

The Black and Karasinsky (1991) (BK) model

After recognising some of the problems with the BDT model, mentioned above, BK propose an alternative procedure for modelling the interest rate process. Again the short-term interest rate has a distribution which limits to the lognormal. The process is approximately given by:

$$lnr_{t+1} = lnr_t + \alpha(t)(ln\theta(t) - lnr_t) + \sigma(t)\Delta z. \tag{10}$$

Here the logarithm of the rate mean reverts at a rate $\alpha(t)$ to a varying mean $\theta(t)$. The parameter $\sigma(t)$ is the volatility of the logarithm of the rate, and it is possibly time varying.

The idea is again to match both the term structure at time 0 and the term structure of volatilities. In the case of the BK model, the volatility input however is a set of 'local', i.e. conditional volatilities. These cannot be observed, but are implied by the cap and floor volatilites, and the degree of mean reversion. The inputs of the BK model are, therefore, the conditional volatilities and the degree of mean reversion. The output of the model is a tree of rates from which it is possible to compute yields on bonds and yield volatilities. The method used by BK, which allows them to construct a recombining tree with probabilities equal to 0.5 is innovative. They capture mean reversion by varying the time spacing.

The papers by BDT and BK leave many questions unanswered. The methods are somewhat ad hoc and give the impression of being cobbled together. Many details of how the methods are implemented are not given, and the reader is left with the task of figuring out the methods based on odd numerical examples. However, a number of specific questions can be raised. First, it is not clear why BK insist on using probabilities of 0.5 and hence use their spacing change method to capture mean reversion. Also, it is not clear how the BK methodology can be generalised to two or more factors. If we wish to capture both cap/floor and swaption volatilities then it is unlikely that a single-factor model will suffice. A model, similar to that of BK, but with two lognormal factor rates, has been proposed by Stapleton and Subrahmanyam (1997). In this model the mean reversion and the changing volatility can be captured by adjusting the probabilities on the binomial tree.

3.6 Interest Rate Models: The State of the Art

We have discussed a wide range of models of interest rates, ranging from the equilibrium model of CIR to the practitioner models of BDT and BK. What overall conclusions can be drawn about the state of the art and about future research? First, we have seen a change of emphasis in recent work towards models which are primarily concerned with the pricing and risk management of interest rate derivatives. These are no-arbitrage models which take the current term structure as given and model interest rates under the EMM. The dominant model is that of HJM. However, the HJM framework, modelling forward rates has disadvantages as well as advantages. We will have to see whether this remains the dominant paradigm after more models built on the more intuitive lines of Vasicek have been developed. I suspect that in the multi-factor case the HJM framework is a rather tortuous view of the world and will eventually be replaced by simpler models.

At the financial engineering level, I expect much progress in the next few years in building better single-factor as well as tractable multi-factor models. The papers by BDT and BK are preliminary, and will surely be improved on. Also, I expect empirical work to throw light on the question of how many factors should be modelled when a term-structure model is required for a particular purpose.

With regard to equilibrium models, I feel it is a mistake to regard the CIR and similar models as superior to the no-arbitrage models of the HL type. I do not expect to see much effort made to construct full equilibrium models consistent with general models of the HJM variety. I think equilibrium analysis, referring back to the preferences of agents, is instructive in reminding us that there may well be preference restrictions embedded in the assumptions of no-arbitrage models. However, even if the assumptions behind the Black-Scholes model may imply some restrictions on the type of underlying economy, we do not have to build an equilibrium model every time we wish to value an option on a stock. The same is true in the case of interest rates and interest rate options. We do not have have to build an equilibrium model every time we wish to value an American-style option using the HJM model.

References: Background Risk and its Implications for Finance

1. Eeckhoudt, L., Gollier, C. and Schlesinger, H. (1996), "Changes in Background Risk and Risk Taking Behaviour", *Econometrica*, **64**, 683-689.

2. Eeckhoudt, L. and M.S. Kimball (1992), "Background Risk, Prudence and the Demand for Insurance", in *Contributions to Insurance Economics*, G. Dionne Ed., Kluwer Academic Press.

3. G Franke, R.C. Stapleton and M.G. Subrahmanyam (1997), "Standard Risk Aversion and the Demand for Risky Assets in the Presence of background Risk", working paper

4. G Franke, R.C. Stapleton and M.G. Subrahmanyam (1997a), "Who buys and who sells options: The role and pricing of options in an economy with background risk" working paper

5. Gollier, C. and Pratt, J. W. (1996), "Risk Vulnerability and the Tempering Effect of Background Risk", *Econometrica*, **64**, 1109-1123.

6. Kihlstrom R.E., Romer D. and Williams S. (1981), "Risk Aversion with Random Initial Wealth", *Econometrica*, **49**, 911-920.

7. Kimball, M.S. (1990), "Precautionary Saving in the Small and in the Large", *Econometrica*, **58**, 53-73.

8. Kimball, M.S. (1993), "Standard Risk Aversion", *Econometrica*, **61**, 589-611.

9. Nachman, D.C. (1982), "Preservation of 'More Risk Averse' under Expectations", *Journal of Economic Theory*, **28**, 361-368.

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

11. Pratt, J.W. and Zeckhauser, R.J. (1987), "Proper Risk Aversion", *Econometrica*, **55**, 143-154.

12. Ross, S. (1981), "Some stronger measures of risk aversion in the small and large with applications", *Econometrica*, **49**, 621-638.

13. Rothschild, M. and Stiglitz, J.E. (1970), "Increasing Risk I: A Definition", *Journal of Economic Theory*, **2**, 225-243.

14. Weil, P. (1992), "Equilibrium Asset Prices with Undiversifiable Labor Income Risk," *Journal of Economic Dynamics and Control*, **16**, 769-790.

References: No-arbitrage models of the term structure and the pricing of interest rate derivatives.

1. Backus, D.K. (1994), “Bond Pricing in Discrete Time”, New York University, mimeo.
2. Backus, D.K., and S.E. Zin (1993), “Reverse Engineering the Yield Curve”, New York University, mimeo.
3. Black, F., E. Derman, and W. Toy (1990), “A One-Factor Model of Interest Rates and its Application to Treasury Bond Options”, *Financial Analysts Journal*, **46**, 33–339.
4. Black, F. and P. Karasinski (1991), “Bond and Option Pricing when Short Rates are Lognormal”, *Financial Analysts Journal*, **47**, 52–59
5. Constantinides, G.M., (1992), “A Theory of the Nominal Term Structure of Interest Rates,” *Review of Financial Studies*, **5**, 4, 531–552.
6. Cox, J.C., J.E. Ingersoll, and S.A. Ross (1985), “A Theory of the Term Structure of Interest Rates”, *Econometrica*, **53**, 385-407.
7. Dothan, L.U., (1978) “On the Term Structure of Interest Rates”, *Journal of Financial Economics*, **6**, 59–69.
8. Harrison, J.M. and D.M. Kreps (1979), “Martingales and Arbitrage in Multiperiod Securities Markets”, *Journal of Economic Theory*, **20**, 381-408.
9. Heath, D., R.A. Jarrow, and A. Morton (1990a), “Bond Pricing and the Term Structure of Interest Rates: A Discrete Time Approximation,” *Journal of Financial and Quantitative Analysis*, **25**, 4, December, 419–440.
10. Heath, D., R.A. Jarrow, and A. Morton (1990b), “Contingent Claim Valuation with a Random Evolution of Interest Rates,” *Review of Futures Markets*, **9**, 1, 55–75.
11. Heath, D., R.A. Jarrow, and A. Morton (1992), “Bond Pricing and the Term Structure of Interest Rates: A New Methodology for Contingent Claims Valuation,” *Econometrica*, **60**, 1, January, 77–105.
12. Ho, T.S.Y. and S.B. Lee (1986), “Term Structure Movements and Pricing of Interest Rate Claims,” *Journal of Finance*, **41**, December, 1011–1029.
13. Hull, J. and A. White (1990), “Pricing Interest-Rate-Derivative Securities”, *Review of Financial Studies*, **3**, 573–592.

14. Longstaff, F.A. and E.S. Schwartz (1992), "Interest Rate Volatility and the Term Structure: a Two-Factor General Equilibrium Model", *Journal of Finance*, **47**, 1259–1282.

15. Stapleton, R.C. and M.G. Subrahmanyam (1997) "Arbitrage Restrictions and Multi-Factor Models of the Term Structure of Interest Rates," working paper

16. Strickland, C. (1993), "A Comparison of Models for Pricing Interest Rate Derivative Securities", University of Warwick, mimeo.

17. Subrahmanyam, M.G. (1996), "The Term Structure of Interest Rates: Alternative Approaches and Their Implications for the Valuation of Contingent Claims", *The GENEVA PAPERS on Risk and Insurance Theory*, Spring, 1996.

18. Vasicek, O., (1977), "An Equilibrium Characterization of the Term Structure", *Journal of Financial Economics*, **5**, 177–188.