TWO TOPICS IN FINANCE:

1. WELFARE ASPECTS OF AN ASYMMETRIC INFORMATION RATIONAL EXPECTATIONS MODEL
2. BOND OPTION PRICING, EMPIRICAL EVIDENCE

By

BRUCE JOHN DIETRICH-CAMPBELL

B.Sc., The University of British Columbia, 1977
MBA, The University of British Columbia, 1981

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

in

THE FACULTY OF GRADUATE STUDIES
FACULTY OF COMMERCE AND BUSINESS ADMINISTRATION

We accept this thesis as conforming
to the required standard

THE UNIVERSITY OF BRITISH COLUMBIA
August 1985

In presenting this thesis in partial fulfilment of the requirements for an advanced degree at the University of British Columbia, I agree that the Library shall make it freely available for reference and study. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by the head of my department or by his or her representatives. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

Department of Commerce & Business Admin.

The University of British Columbia
1956 Main Mall
Vancouver, Canada
V6T 1Y3

Date 29 Aug 85

ABSTRACT

Welfare Aspects of an Asymmetric Information Rational Expectations Model

In part 1 of this study I examine several models of competitive markets in which a group of uninformed traders uses the equilibrium price of a traded asset as an indirect source of information known to a group of informed traders. Four different models are compared in two homogeneous information cases plus one asymmetric information case, revealing a) an allocative efficiency benefit resulting from the opportunity to trade current consumption for future consumption, b) a 'dealer' benefit accruing to traders who are able to observe and act on demand fluctuations not apparent to other traders, c) a 'hedging' benefit accruing to all traders, and d) a loss of hedging benefits due to information dissemination before hedge trading can take place. The effect of an increase in precision of information given to informed traders is calculated for the above factors and for net welfare.

Bond Option Pricing, Empirical Evidence

In part 2, a two-factor model using the instantaneous rate of interest and the return on a consol bond to describe the term structure of interest rates – the Brennan-Schwartz model – is used to derive theoretical prices for American call and put options on U.S. government bonds and treasury bills. These model prices are then compared with market prices. The theoretical model used to value the debt options also provides hedge ratios which may be used to construct zero-investment portfolios which, in theory, are perfectly riskless. Several trading strategies based on these 'riskless' portfolios are examined.

TABLE OF CONTENTS

Welfare Aspects of an Asymmetric Information Rational Expectations Model

Bond Option Pricing, Empirical Evidence

LIST OF TABLES

Welfare Aspects of an Asymmetric Information Rational Expectations Model

Bond Option Pricing, Empirical Evidence

LIST OF FIGURES

Welfare Aspects of an Asymmetric Information Rational Expectations Model

OVERVIEW

This thesis presents two unrelated studies in finance. To aid the reader, part A – Welfare Aspects of an Asymmetric Information, Rational Expectations Model – is presented in its entirety before part B – Bond Option Pricing, Empirical Evidence. This means that the references, appendices, tables and figures for part A have been placed *before* the start of part B. Only the references and tables for part B will be found at the end of the thesis.

PART A:

Welfare Aspects of an Asymmetric Information Rational Expectations Model

A.1 INTRODUCTION

This study deals with models of rational expectations markets, that is, with markets in which traders look to the equilibrium price of an asset as a source of information on which to base their own trading decisions. It is true that prices reflect information in Walrasian models, with the difference being that Walrasian traders do not attempt to use the information that is present in prices. To the extent that traders actually do try to use market prices as a indirect source of other traders' information, modelling trading activity using a rational expectations model such as those used here should be a more accurate representation of reality. At the least, we may be able to identify qualitative aspects of rational expectations models not present in their Walrasian counterparts.

Models of Walrasian markets have been analyzed extensively in order to isolate and identify the various dynamics underlying individual welfare changes. The same is not true of the rational expectations models that have appeared in the finance literature. The reason is undoubtedly the complexity of the formulae describing the expected utility of market participants. In this study, I have hopefully illuminated the inner workings of several asymmetric information, rational expectations models and helped somewhat to fill this gap in our knowledge.

The basic features of the models used in this study are the same as used by Grossman and Stiglitz (1980), namely, negative exponential utility functions, normally distributed random variables and the division of traders into two groups: informed traders and uninformed traders. The informed traders all receive a piece of information before trading begins,[1] and the uninformed traders try to use the price of the risky asset as an indirect source of the information that informed traders have received.

In several points, however, my models do differ from the Grossman and Stiglitz model. Unlike their model, the individual traders in my models do not have the choice of being informed or uninformed. The groups are predetermined and traders do not have the option of moving from one group to the other. In addition, the Grossman and Stiglitz model allows consumption to occur only at one point in time. I improve slightly on this by allowing consumption to occur both at the beginning and end of the period which is modelled.

My main concern in this study has been to provide an understanding of the welfare changes which result in these rational expectations models when the quality of information given to the informed trader group is increased.[2] As there are five welfare effects which come to light in this study, all of them requiring reasonably lengthy discussion, I will not attempt to outline the results at this point, and trust that the sections that follow will succeed in that regard.

---

[1] All informed traders receive the *same* piece of information.

[2] By increasing the quality of information, I mean that the correlation between the information and the future payoff on a risky asset is increased.

A.2 MODEL DESCRIPTION

The two elements that are crucial to an analytic solution of the problem addressed here are first of all the negative exponential form of the utility function used and second the assumption that all random variables are joint normally distributed. The reason that these two rather restrictive assumptions are needed has to do with the nature of the problem being solved, namely, its rational expectations character. Simply stated, the market participants condition their actions on an observable indicator – in this case a market price – but this same indicator is endogenously determined by the actions of all the participants in the model. That is, we must solve a 'chicken and the egg' problem or, in more technical terms, a fixed point problem. As a result, the analytic solution found in this study is very sensitive to some of the assumptions, especially the two mentioned above.

A.2.1 Utility Functions

As was mentioned above, it is quite important that the utility function used is negative exponential. This utility function paired with joint normally distributed random variables combine to allow the only known analytically solvable rational expectations model.[3] Even when these two assumptions are made, the solution

[3] This is not true if one allows random variables having discrete probability distributions. See Kraus and Sick (1979) for an example using the power utility

is still quite difficult to obtain. So difficult, in fact, that the models to date have been single period models with consumption allowed only at the end of period. The utility function used in the literature up to this point has been

$$U = -\beta \exp(-aC), \quad a > 0, \quad 1 \geq \beta \geq 0, \quad \text{'standard' model,}$$

where C is consumption at the end of period. Since this has been the standard in the literature to this point, I will call this the 'standard' model.

In addition, I analyze a more complex model allowing consumption at two points in time, namely, both the beginning and end of period. The utility function used for this extension of the standard model is[4]

$$V = -\exp(-aC_0) - \beta \exp(-aC_1), \quad \text{'extended' model,}$$

where C_0 is the beginning of period consumption and C_1 is consumption at the end of period. Because this model is an extension of the standard one, I call it the 'extended' model.

A.2.2 Aggregation and Trading Volume

Unlike the Verrecchia (1982) model where each individual receives an independent piece of information, the Grossman and Stiglitz formulation has all informed

function.

[4] Note that a is the same for both the beginning and end of period terms in the 'extended' model utility function, and the same for all all traders. The variable β is also the same for all traders.

traders receiving the *same* piece of information. Because of this, we make the assumption that we can aggregate over all the informed traders, replacing them with a representative informed trader. Similarly, we use a representative uninformed trader. Apparently this is an innocuous assumption, though there is one aspect of it which makes model interpretation difficult.

The one problem which arises is with regard to trading volume. Since we have replaced groups of individuals with representatives, the model trading volumes represent only the trading that occurs between the two *groups*. They do not include trading which occurs between members of the groups. That is, representative trader trading volume includes only *inter*group trading, not *intra*group trading.[5] Other than this, there should be no other side effect of group aggregation.

A.2.3 Endowments

Previous to this study, there was no question what the endowments of the market participants were. There were only two assets – the riskless technology and the risky asset – and trader i was given an initial endowment of each: m_i of the riskless technology and f_i of the risky asset. Aggregate supply constraints were

[5] Of course, in the particular framework of this study, since all traders have the same utility function there is no intragroup trading, only intergroup trading. Intragroup trading would occur in a model in which traders had different tastes.

imposed to give

$$\frac{1}{n}\sum_{i=1}^{n} m_i = m, \qquad \frac{1}{n}\sum_{i=1}^{n} f_i = f.$$

With this setup, trader i's initial wealth is

$$W_{0i} = m_i + f_i p,$$

where p is the price of the risky asset.[6]

A.2.3.a GROSSMAN AND STIGLITZ MODEL

In the interests of simplicity, Grossman and Stiglitz (1980) limited the initial endowments of traders to the two mentioned above, namely, endowments of the riskless technology and risky asset. However, since participants in the model are given perfect information about the entire structure of the economy – the utility functions of all traders and the distribution functions of all random variables are common knowledge – Grossman and Stiglitz were able to show that without some obscuring 'noise' in the model the result is a fully revealing risky asset price.

That is, when there is no 'noise' the uninformed traders have enough knowledge of the structure of the economy to enable them to figure out *exactly* what the informed traders' demand for the risky asset is as a function of the information

[6] The price of the riskless (constant returns to scale) technology is 1. Prices are in terms of units of a good which may either be consumed or invested, with the only restriction being that consumption cannot occur at the beginning of period in the standard model.

they receive. Since no trading takes place in this rational expectations market except at the equilibrium price, we assume that the equilibrium price is part of the information set of the uninformed. However, because they have exact knowledge of the informed traders' demand function, the uninformed traders can use the equilibrium price to figure out exactly what information the informed traders have received. Unless we add some obscuring 'noise', the equilibrium price is a *sufficient statistic* for the information.

Grossman and Stiglitz chose to introduce the needed 'noise' by making the aggregate supply of the risky asset an unobservable random variable.

$$\frac{1}{n}\sum_{i=1}^{n} f_i = f \sim N\left\{\mu_f, \sigma_f^2\right\}$$

With this change, the equilibrium price reveals to the uninformed traders only a linear combination of aggregate supply noise and the informed traders' information. Using this linear combination, uninformed traders are not able to exactly invert the pricing function to find what information was received by the informed traders. They are only able to calculate a probability distribution for this information. Since the price no longer fully reveals information it is called a partially revealing equilibrium price.

A.2.3.b EFFICIENT MARKET MODEL

There is an alternative way to introduce noise into the supply of the risky asset.[7] We can retain the assumption of a random aggregate supply of the risky asset, and still allow the beginning of period holdings to be non-random. This leads us into a small difficulty of interpretation, since the sum of the initial endowments of risky asset does not equal the aggregate supply. That is,

$$\frac{1}{n}\sum_{i=1}^{n} f_i = f \neq s \sim N\left\{f, \sigma_s^2\right\},$$

where s is the *per trader* supply of risky asset.

Apparently, by making this assumption the model has been opened. That is, an exogenously determined element has been added. One characteristic of this exogenously determined supply component is that its size is independent of the price of the risky asset. It is a perfectly inelastic supply component. This is in stark contrast to the demand/supply functions of our rational expectations traders. These traders condition their demands on the market price of the risky asset. Since this additional supply component is exogenous and perfectly price inelastic, we can infer that another – unpredictable – group of traders has been added to the model, and that they are *not* rational expectations traders.

The interpretation that I like to put on the model is the following. If we are

[7] This model is labelled the 'efficient market model' not because the market modelled here is actually efficient – it is not – but because a group of traders assumes, or acts as if, it were efficient.

attempting to model a real market by using rational expectations traders, we most likely don't really want to model a situation where every trader is a rational expectations trader. In order to draw information from the price at which an asset is trading, a trader has to be intimately familiar with the market he is trading in. If the price acts in a peculiar way, he must first be able to identify peculiar price behaviour when it occurs, and second be able to interpret what any given peculiar price behaviour means. The ordinary trader does not follow the market enough to be able to identify peculiar price behaviour, or to interpret it if he could identify it. The rational expectations traders, therefore, form a central group of traders closely following the price behaviour of an asset.

Normally, when we think of an efficient market we think of the price of the asset reflecting all known information about the asset. The asset price reflects all this information because a small group of traders closely follow the asset's price movements and step in whenever the price departs even slightly from what it 'should be'. They keep the asset price where it 'should be'. The ordinary trader, of course, benefits from this, too. He is assured that the equilibrium price at which an asset trades in an efficient market is the correct price at which to buy the asset. That is, since the asset is presumably always correctly priced, the price the ordinary trader has to pay doesn't matter. It's always guaranteed by the actions of the rational expectations traders to be the right price. If the market is efficient, a 'naive' trader can safely ignore the fact that an asset price conveys information.

This, then, is a justification of the exogenous, perfectly inelastic supply of the risky asset. We can think of it as due to the demand/supply generated by 'naive', efficient market traders.[8] The reason that they are modelled as a *random* supply element is that they are unpredictable as far as the rational expectations traders are concerned. When the rational expectations traders invert the pricing function, the result is a probability distribution for supply/demand aggregated over all the *rational expectations* traders. This group of traders forms a small, stable core of predictable traders, whereas little is known about the much larger group of 'naive', efficient market traders. Because little is known about how they form their demands, they appear to trade randomly.[9]

Given that the beginning of period endowments of risky asset are now free of the additional role of introducing randomness into the model, I decided to set the rational expectations traders' initial endowments equal to the Hakansson, Kunkel and Ohlson (1982) 'no-information' endowments. That is, I set the endowments to what the traders' equilibrium holdings would be *if all rational expectations traders were uninformed* and we allowed them to trade to equilibrium before the beginning of the period.[10] Since we are dealing with the simplified case where all

8 Although I label these traders 'efficient market traders', I do not mean to imply that the market modelled here is actually efficient. What is meant is that these traders assume or act as if it were efficient. As is shown in later sections, this turns out to be an erroneous assumption. The market is not efficient in the model I have labelled the 'efficient market model'.

9 Note that there must be some correlation between the individual demands of these 'naive' traders, assuming that there is a large number of traders in this group. If there were no correlation, or 'fads', then aggregation of their demands – each of which is assumed normally distributed with a mean of zero – would result in an aggregate demand of zero, with *no* variance. That is, in the limit as we aggregate over an infinite number of 'naive' traders, without 'fads' there would be no random supply component.

10 Note that these Hakansson, Kunkel and Ohlson endowments are only applicable in a single-period model. If we wished to view this as a model of one period taken from a multi-period framework, then this additional round of trading can be eliminated. It is not required, as all of the possible information effects that may arise are discussed in this study. Eliminating this round of trading would not

traders have the same absolute risk aversion, this means that all traders receive the same no-information initial endowment of risky asset, namely, the expected beginning of period per trader supply, f.

A.2.3.c HEDGING MODEL

Consumption based trading is an important use of 'real' markets, but we know that it is not the only use. People also trade in the market in order to hedge their positions in other non-tradable assets, or assets which might be traded if one was willing to pay high transaction costs. The reason for hedge based trading is that exposure to risk from holding an asset – which can be traded only by incurring high transaction costs – can be reduced by trading in another asset with transaction costs which are *relatively* low. For simplicity, I will assume here that the assets being hedged are non-tradable assets.

Specifically, I assume that traders have an endowment of an additional, non-tradable asset. This endowment, h_i, has a payoff which is equal to the payoff on the risky asset. This, of course, covers all possible cases, since an asset with a payoff which is *partially* correlated with the payoff on the risky asset can be thought of as a combination of two assets: one having a payoff perfectly *un*correlated with the risky asset payoff and the other perfectly correlated.[11] Trader i's initial wealth,

---

change any of the conclusions of this study.

[11] I do not explicitly model the asset having a payoff perfectly uncorrelated with the

therefore, is

$$W_{0i} = m_i + (f_i + h_i)\, p,$$

where p is, as before, the price of the risky asset.[12]

Randomness is introduced into the model by assuming that the endowments of this non-traded asset are random. That is, at the beginning of the period each trader receives income in the form of non-tradable asset, h_i. The amount of income received by one individual is unknown to the rest of the traders in the market, that is, it appears random to them.[13]

$$\frac{1}{n}\sum_{i=1}^{n} h_i = h \sim N\left\{\mu_h, \sigma_h^2\right\}$$

As proposed by Hellwig (1980), we also assume that no trader can use his own income, h_i, to determine anything about the aggregate quantity, h. To prevent this, Hellwig assumes that there is a very large number of traders in the market, so that each trader's contribution to this aggregate is infinitesimal and, consequently, the correlation between any particular h_i and the aggregate h is also infinitesimal.

There are two pleasing aspects to this model. First, we have avoided the Grossman-Stiglitz assumption of random beginning of period holdings of a *traded* asset, some-

---

risky asset payoff, as there are no effects of information on this asset's contribution to utility. That is, explicitly adding this asset would merely shift the utility function, with the amount of the shift being unaffected by anything else in the model.

12 Since the payoff on the risky asset and h_i are equal, the shadow price of the non-tradable asset h_i will equal the market price of the risky asset.

13 I assume that the trader knows his own non-tradable income. It is sufficient to assume that the rest of the traders are ignorant of it. Note that this model is *conceptually* extendible to a multiperiod framework, whereas the Grossman and Stiglitz model is not (even though the two are mathematically identical).

thing which is difficult to justify if we are thinking in terms of financial assets such as stocks. Second, we have expanded the model to include not only consumption based trading, but also hedge based trading. Notice, however, that since the risky asset and non-tradable asset are perfect substitutes, this model is mathematically identical to the Grossman-Stiglitz model. It provides another way to interpret the Grossman-Stiglitz framework.

A.2.4 The Riskless Technology

Besides the risky asset, there is also a riskless technology in this economy. The riskless technology can be consumed either now or at the end of the period in the 'extended' model, but only at the end of the period in the 'standard' model.[14] The payoff per unit of the riskless, constant returns to scale technology, R, is assumed to be exogenously given.

$$R = 1 + r, \quad r \text{ exogenous}$$

This corresponds to a rate of return, r, which is totally insensitive to supply or demand. That is, the supply of riskless technology is perfectly elastic at this rate of return.

This assumption is needed in order to keep the model relatively simple. As is

[14] The risky asset cannot be consumed, it must be held until the end of the period. At that time it produces a payoff which is consumable.

easily seen, if we were instead to allow the riskless technology price to be sensitive to aggregate demand for the riskless technology, then its price would convey information about that demand. Of course, demand for the riskless technology is a function of the information that informed traders have, so that the uninformed traders in the model could potentially use the price of the riskless technology as an indirect source for this information.[15] To prevent the price of the riskless technology from *fully revealing* this information, we would be forced to introduce randomness into the supply of the riskless technology, thus complicating the model needlessly. Therefore, in order to keep the model relatively simple, we want to restrict the function of information transmission to the price of the risky asset only, which forces us to make the supply of the riskless technology perfectly elastic.

A.2.5 Random Elements

Up to this point, we have already mentioned several random variables in the models. In the section on endowments, section A.2.3, when defining the random initial endowments of f_i in the Grossman and Stiglitz model, the random aggregate supply of the risky asset, s, in the efficient market model, or the random non-tradable endowments, h_i, of the hedging model, we defined the random elements uing normal probability distributions. This is, as mentioned previously, required for the analytic solution of our rational expectations model.

[15] A similar argument holds if the riskless technology price is sensitive to the price of the risky asset.

I have mentioned in several places that there is a risky asset in the models, but have never explicitly defined it. I have also mentioned that the informed group of traders receives information correlated with the future payoff on the risky asset.[16] More precisely stated, let[17]

$$\begin{pmatrix} x \\ \epsilon \end{pmatrix} \sim N\left\{\begin{pmatrix} \mu_x \\ 0 \end{pmatrix}, \sigma_x^2 \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}\right\}, \quad \rho \geq 0,$$

where x is the end of period payoff per unit of the risky asset and ϵ is the information received by the informed group of traders. The correlation between these two variable, ρ, can be thought of as the 'informativeness' of the information. Note also that the end of period payoff on the non-tradable asset endowment or income h_i is equal to the payoff on the risky asset.[18]

As a last note, all the random elements are taken to be independently distributed, except for x and ϵ which, as stated, have a correlation of ρ.

A.2.6 Timing of Events

The exact timing of events may be confusing at first, so I have provided a time line in Figure A.1. The first event to occur is the receipt of the risky and riskless

[16] Recall also that all informed traders receive the same piece of information.

[17] Since the definition of information is somewhat arbitrary, I have chosen to let the variance of ϵ be equal to the variance of x, and its mean be equal to zero. This is justifiable, as the information contained in the random variable ϵ is exactly the same as the information contained in an arbitrary linear combination $a + b\epsilon$.

[18] See section A.2.3.

endowments. We also supply the rational expectations traders with their common knowledge regarding the utility functions of the other participants. We do not, however, provide them with the knowledge that there will be an asymmetric information situation in the future. Next, we allow the rational expectations traders to trade to a Hakansson, Kunkel and Ohlson (1982) 'no-information' equilibrium.[19] [20] Once the Hakansson, Kunkel and Ohlson equilibrium has been reached, we supply traders with their endowments of non-tradable asset (in the hedging model), and with the rest of their common knowledge, namely, who will be members of the informed group plus the distribution functions for all the random variables of the model. Following this, we supply traders with their additional non-tradable endowments (in the hedging model only).

We can then calculate the 'pre-info' expected utility of wealth for the informed and uninformed traders in the model. That is, we calculate their expected utility *before* receipt of information by the informed traders. Of course, the 'post-info' expected utility of wealth is a function of the actual signal received, so it cannot be used as a basis for conclusions regarding an individual's welfare. Otherwise, our conclusion might also be a function of the actual signal received. It seems more reasonable to base any conclusions on the expected utility of individuals at a point in time where a signal is expected to be but has not yet been received, that is, on the 'pre-info' expected utility, which is calculated in *expectation* of information receipt.

[19] That is, the rational expectations traders trade without knowing that there will be an asymmetric information situation arising in the future.

[20] Note that this round of trading is not present in the Grossman and Stiglitz model, only in the efficient market and hedging models. As mentioned in a previous footnote (section A.2.3.b), this round of trading is not essential in the efficient market model nor the hedging model.

Once a signal has been received by the informed trader group, trading (and consumption in the 'extended' model) may occur. The beginning of period market clearing price of the risky asset is determined by equilibrium conditions and is used by the uninformed traders as an imperfect signal of the information received by the informed traders. It is at this point that we can calculate the 'post-info' expected utility of traders.

Nothing further happens in the model until the end of period, at which time the riskless technology and risky asset generate their payoffs, and final consumption of end of period wealth occurs.

One potential problem with the sequence of events as depicted in Figure A.1 is the lack of an additional round of trading just before the point where a signal is received by the informed trader group. Presumably, the uninformed traders could somehow insure themselves against potential exploitation by the better informed traders if only they had the opportunity of doing so before the information was received. In fact, it is not even necessary to introduce an insurance market into the model. Simply allowing an additional round of trading in the risky asset just before information was received by the informed traders would protect the uninformed traders against any potential exploitation.

The effect of an additional round of trading before information receipt would be to reveal some otherwise unknown random variable to all traders. For example, in the Grossman and Stiglitz model the only reason the market clearing price of the risky asset does not fully reveal to uninformed traders the information received by informed traders is that the price is a function of two random variables – the aggregate supply of risky asset and the information received by informed traders – both of which are unknown to the uninformed traders. If there were an additional round of trading prior to receipt of information, then the clearing price at that point would be a function of only the aggregate supply of risky asset. That is, the price would reveal the value of the aggregate supply to uninformed traders.

When the next round of trading occurred, after information was received by the informed traders, the market clearing price would still be a function of two random variables – aggregate supply and information – but only one would still be unknown to the uninformed traders. The result is that the price would reveal the value of the second random variable to the uninformed traders. That is, it would perfectly reveal the information received by the informed traders. In order to retain a partially revealing market clearing price, we cannot allow another round of trading to occur between receipt of endowments and information. A partially revealing price must be a function of two random variables, both of which are unknown to uninformed traders.[21]

[21] In the efficient market model, we can allow an additional round of trading *between rational expectations traders* without revealing an otherwise unknown random variable to the uninformed. However, we cannot allow a round of trading between rational expectations traders and 'naive' traders if doing so would reveal what the 'naive' trader demand at the next trading point – after information receipt – would be. For example, if this round of trading exhausted 'naive' trader demand, then the next round of trading would take place without 'naive' traders and the price of the risky asset would be fully revealing. In the hedging model, an extra round of trading would reveal to the uninformed traders the size of the average endowment of non-tradable asset, h. Consequently, at the next round of trading, the risky asset price would be fully revealing.

A.3 MODEL DERIVATION

It may be confusing to understand how one can calculate optimal actions using a price function for the risky asset, when that price function is endogenously determined by these actions and has itself not yet been calculated. This, of course, is the essence of the rational expectations problem. The price function is a fixed point solution to the problem. It is that particular price function which, when used to calculate optimal actions, leads to a price function which just happens to be the same as the one we started with.

The models that are derived in the sections below ignore the possibility of informed trader cartels. That is, we deal here only with the case of non-schizophrenic traders *à la* Hellwig (1980).

A.3.1 Initial and Terminal Wealth

Collecting all that we have said about endowments, the initial wealth of one of our traders, trader i, must be

$$W_{0i} = m_i + (f_i + h_i)\, p.$$

If this trader changes his position so that he holds a total of z_i units of the risky asset[22] and consumes C_{0i} at the beginning of the period,[23] then assuming he invests the balance of his tradable wealth in the riskless technology, his holdings at the beginning of period are

z_i units of the risky asset,

h_i units of non-tradable asset h,

$m_i + (f_i - z_i)\,p - C_{0i}$ units of riskless technology.

Since the end of period payoffs on these different assets are, respectively, x, x and R, this individual's end of period wealth would be

$$\begin{aligned} W_{1i} &= (z_i + h_i)\,x + R\,(m_i + (f_i - z_i)\,p - C_{0i}) \\ &= R\,(m_i + f_i p - C_{0i}) + z_i(x - Rp) + h_i x. \end{aligned}$$

The model ends at the end of period, so this is also the end of period consumption, C_{1i}.[24]

A.3.2 'Post-info' Expected Utility

The post-info expected utility of our traders is calculated at the beginning of period, at exactly the same time that the market equilibrium price of the risky asset is determined. The reason the two events occur simultaneously is that we cannot calculate an uninformed trader's expected utility until he has his full

[22] Unlimited short sales are allowed. Thus, the Grossman-Stiglitz and hedging models are mathematically identical even though the hedging model contains an extra asset.

[23] In the 'standard' model, consumption is not allowed at the beginning of period. Just let C_{0i} be zero for this case.

[24] C_i in the 'standard' model.

information set. However, part of the uninformed trader's information set is the market clearing price itself. He calculates his optimal beginning of period consumption and investment using the market clearing price of the risky asset as a signal telling him something about the information received by the informed traders.

Of course, this is the exact nature of the rational expectations model we are solving. The optimal consumption and investment are fixed point functions. We use the optimal consumption and investment decisions of all individuals to determine the market clearing price, and given that market clearing price people are satisfied that the consumption and investment decisions that they made are actually optimal. That is, nobody wants to change his decision.

A.3.2.a PROBABILITY DISTRIBUTION OF FUTURE CONSUMPTION

Using the end of period wealth expression derived above, our description of individual i to this point is

$$U_i = -\beta \exp(-aC_i),$$

$$C_i = R(m_i + f_i p) + z_i(x - Rp) + h_i x,$$

in the 'standard' model and

$$V_i = -\exp(aC_{0i}) - \beta \exp(-aC_{1i}),$$

$$C_{1i} = R(m_i + f_i p - C_{0i}) + z_i(x - Rp) + h_i x,$$

in the 'extended' model.

As all the variables in the model are normally distributed,[25] and end of period consumption is a simple linear combination of these normal variables, end of period consumption must also be a drawing from a normal distribution. In the 'standard' model, the mean and variance of this distribution are

$$\mathbf{E}\,(C_i|I_i) = R(m_i + f_i p) + z_i[\mathbf{E}\,(x|I_i) - Rp] + h_i \mathbf{E}\,(x|I_i)\,,$$
$$\mathbf{var}\,(C_i|I_i) = (z_i + h_i)^2 \mathbf{var}\,(x|I_i)\,,$$

where I_i is the information set available to individual i.[26] In the 'extended' model we have

$$\mathbf{E}\,(C_{1i}|I_i) = R(m_i + f_i p - C_{0i}) + z_i[\mathbf{E}\,(x|I_i) - Rp] + h_i \mathbf{E}\,(x|I_i)\,,$$
$$\mathbf{var}\,(C_{1i}|I_i) = (z_i + h_i)^2 \mathbf{var}\,(x|I_i)\,.$$

In order to calculate the optimal beginning of period consumption (in the 'extended' model) and risky asset investment, we will need to calculate the derivatives of the expressions above with respect to the decision variables C_{0i} and z_i. For convenience, these will be stated here. For the 'standard' model we have

$$\frac{\partial}{\partial z_i}\mathbf{E}\,(C_i|I_i) = \mathbf{E}\,(x|I_i) - Rp,$$
$$\frac{\partial}{\partial z_i}\mathbf{var}\,(C_i|I_i) = 2(z_i + h_i)\,\mathbf{var}\,(x|I_i)\,,$$

[25] I make the assumption here that the price of the risky asset will be a simple linear combination of variables drawn from normal distributions, and will therefore also be a drawing from a normal distribution. We see below that there does exist a fixed point price function satisfying this assumption.

[26] The informed trader information set contains the information received by the informed trader group plus the price of the risky asset, while the information set of uninformed traders only contains the price of the risky asset. In addition, all rational expectations traders have complete knowledge about the utility functions of all traders and distributions of the random elements of the economy.

and in the 'extended' model

$$\frac{\partial}{\partial z_i}\mathbf{E}\left(C_{1i}|I_i\right) = \mathbf{E}\left(x|I_i\right) - Rp,$$
$$\frac{\partial}{\partial z_i}\mathbf{var}\left(C_{1i}|I_i\right) = 2(z_i + h_i)\mathbf{var}\left(x|I_i\right),$$

$$\frac{\partial}{\partial C_{0i}}\mathbf{E}\left(C_{1i}|I_i\right) = -R,$$
$$\frac{\partial}{\partial C_{0i}}\mathbf{var}\left(C_{1i}|I_i\right) = 0.$$

A.3.2.b OPTIMAL INVESTMENT AND CONSUMPTION

In this section it becomes clear exactly why we require negative exponential utility functions and normal random variables. The problem is a standard utility maximization problem with two decision variables, z_i and C_{0i}. In the 'standard' model, individuals solve for J_i^*, their maximum expected 'post-info' utility,

$$J_i^*(I_i) = J_i(z_i^*; I_i) = \max_{z_i} J_i(z_i; I_i),$$
$$J_i(z_i; I_i) = \mathbf{E}\left(U_i(z_i)|I_i\right),$$

and for K_i^* in the 'extended' model,

$$K_i^*(I_i) = K_i(z_i^*, C_{0i}^*; I_i) = \max_{z_i, C_{0i}} K_i(z_i, C_{0i}; I_i),$$
$$K_i(z_i, C_{0i}; I_i) = \mathbf{E}\left(V_i(z_i, C_{0i})|I_i\right).$$

The reason that we can solve this problem analytically is due to the following property of the exponential function:

$$\mathbf{E}\left(\exp(-ax)|I_i\right) = \exp\left(-a\mathbf{E}\left(x|I_i\right) + \frac{1}{2}a^2\mathbf{var}\left(x|I_i\right)\right), \quad \text{for } x \text{ normally distributed.}$$

That is, the expectation of an exponential function is itself exponential, but only if the argument of the exponential function is normally distributed. Now it is clear why these two assumptions are so critical. Actual calculation of individual i's expected utility gives us

$$J_i(z_i; I_i) = -\beta \exp\left(-a\mathbf{E}\left(C_i|I_i\right) + \frac{1}{2}a^2\mathbf{var}\left(C_i|I_i\right)\right),$$

in the 'standard' model, and

$$K_i(z_i, C_{1i}; I_i) = -\exp(-aC_{0i}) - \beta \exp\left(-a\mathbf{E}\left(C_{1i}|I_i\right) + \frac{1}{2}a^2\mathbf{var}\left(C_{1i}|I_i\right)\right),$$

in the 'extended' model.

If we calculate the first order conditions and set them equal to zero we find that the 'standard' model optimal beginning of period investment in the risky asset z_i^* satisfies[27]

$$\begin{aligned}
&\frac{\partial}{\partial z_i}J_i(z_i^*; I_i) = 0\\
&\quad\Rightarrow -J_i^*(I_i)\left[a\frac{\partial}{\partial z_i}\mathbf{E}\left(C_i|I_i\right) - \frac{1}{2}a^2\frac{\partial}{\partial z_i}\mathbf{var}\left(C_i|I_i\right)\right] = 0,\\
&\quad\Rightarrow a[\mathbf{E}\left(x|I_i\right) - Rp] - a^2(z_i^* + h_i)\mathbf{var}\left(x|I_i\right) = 0,\\
&\quad\Rightarrow \qquad z_i^* = \frac{\mathbf{E}\left(x|I_i\right) - Rp}{a\,\mathbf{var}\left(x|I_i\right)} - h_i.
\end{aligned}$$

In the 'extended' model, the optimal beginning of period investment, z_i^*, and

[27] The expressions for the expectations and variances of C_i have been taken from the previous section.

consumption, C^*_{0i}, satisfy

$$\frac{\partial}{\partial z_i} K_i(z^*_i, C^*_{0i}; I_i) = 0$$

$$\Rightarrow -[K^*_i(I_i) + \exp(-aC^*_{0i})]\left[a\frac{\partial}{\partial z_i}\mathbf{E}\left(C_{1i}|I_i\right) - \frac{1}{2}a^2\frac{\partial}{\partial z_i}\mathbf{var}\left(C_{1i}|I_i\right)\right] = 0,$$

$$\Rightarrow \qquad z^*_i = \frac{\mathbf{E}\left(x|I_i\right) - Rp}{a\,\mathbf{var}\left(x|I_i\right)} - h_i,$$

$$\frac{\partial}{\partial C_{0i}} K_i(z^*_i, C^*_{0i}; I_i) = 0$$

$$\Rightarrow a\ \exp(-aC^*_{0i})$$

$$- [K^*_i(I_i) + \exp(-aC^*_{0i})]\left[a\frac{\partial}{\partial C_{0i}}\mathbf{E}\left(C_{1i}|I_i\right) - \frac{1}{2}a^2\frac{\partial}{\partial C_{0i}}\mathbf{var}\left(C_{1i}|I_i\right)\right] = 0,$$

$$\Rightarrow a\ \exp(-aC^*_{0i}) - [K^*_i(I_i) + \exp(-aC^*_{0i})]\left[-aR\right] = 0,$$

$$\Rightarrow \qquad \exp(-aC^*_{0i}) = -\left(\frac{R}{1+R}\right) K^*_i(I_i).$$

Given these expressions for individual i's optimal beginning of period investment and consumption, we can aggregate the demand for the risky asset and impose an aggregate supply constraint. Notice that the price of the risky asset appears in two places. It appears explicitly as 'p' in the equations, and also implicitly as a part of the information set I_i.[28]

A.3.2.c 'POST-INFO' EXPECTED UTILITY FUNCTIONS

The final step in calculating the 'post-info' expected utility functions is simply substitution of the expressions for z^*_i and C^*_{0I} back into the expressions for ex-

[28] As mentioned in a previous footnote, we are assuming that p is a simple linear combination of variables drawn from normal distributions and is therefore also a drawing from a normal distribution.

pected utility which were used to derive them. That is, from the previous sections, we have

$$
\begin{aligned}
J_i(z_i; I_i) &= -\beta \exp\left(-a\mathbf{E}\left(C_i|I_i\right) + \frac{1}{2}a^2\mathbf{var}\left(C_i|I_i\right)\right) \\
&= -\beta \exp\Bigg(- a[R(m_i + f_i p) + z_i(\mathbf{E}\left(x|I_i\right) - Rp) + h_i\mathbf{E}\left(x|I_i\right) \\
&\qquad + \frac{1}{2}a^2\left[(z_i + h_i)^2\mathbf{var}\left(x|I_i\right)\right]\Bigg),
\end{aligned}
$$

$$
\begin{aligned}
K_i(z_i, C_{0i}; I_i) &= -\exp(-aC_{0i}) - \beta \exp\left(-a\mathbf{E}\left(C_{1i}|I_i\right) + \frac{1}{2}a^2\mathbf{var}\left(C_{1i}|I_i\right)\right) \\
&= -\exp(-aC_{0i}) - \beta \exp\Bigg(- a[R(m_i + f_i p - C_{0i}) \\
&\qquad + z_i(\mathbf{E}\left(x|I_i\right) - Rp) + h_i\mathbf{E}\left(x|I_i\right)] \\
&\qquad + \frac{1}{2}a^2\left[(z_i + h_i)^2\mathbf{var}\left(x|I_i\right)\right]\Bigg).
\end{aligned}
$$

Substituting in the optimal values of z_i and C_{0i} gives

$$
\begin{aligned}
J_i^*(I_i) &= -\beta \exp\Bigg(- a[R(m_i + f_i p) + z_i^*(\mathbf{E}\left(x|I_i\right) - Rp) + h_i\mathbf{E}\left(x|I_i\right)] \\
&\qquad + \frac{1}{2}a^2\left[(z_i^* + h_i)^2\mathbf{var}\left(x|I_i\right)\right]\Bigg) \\
&= -\beta \exp\Bigg(-\Bigg[aR[m_i + (f_i + h_i)p] \\
&\qquad + \frac{1}{2}\frac{\left[\mathbf{E}\left(x|I_i\right) - Rp\right]^2}{\mathbf{var}\left(x|I_i\right)}\Bigg]\Bigg),
\end{aligned}
$$

$$
\begin{aligned}
K_i^*(I_i) = &- \exp(-aC_{0i}^*) - \beta \exp\Bigg(- a[R(m_i + f_i p - C_{0i}^*) + z_i^*[\mathbf{E}(x|I_i) - Rp] \\
&\qquad + h_i \mathbf{E}(x|I_i)] \\
&\qquad + \frac{1}{2}a^2[(z_i^* + h_i)^2 \mathbf{var}(x|I_i)]\Bigg) \\
= &- \exp(-aC_{0i}^*) - \beta \exp\Bigg(- \Bigg[aR[m_i + (f_i + h_i)p - C_{0i}^*] \\
&\qquad + \frac{1}{2}\frac{[\mathbf{E}(x|I_i) - Rp]^2}{\mathbf{var}(x|I_i)}\Bigg]\Bigg).
\end{aligned}
$$

The terms in C_{0i}^* can be eliminated by using one of the first order conditions from the previous section, namely,

$$
\exp(-aC_{0i}^*) = -\left(\frac{R}{1+R}\right) K_i^*(I_i),
$$

giving

$$
\begin{aligned}
K_i^*(I_i) = &\left(\frac{R}{1+R}\right) K_i^*(I_i) - \beta \left[-\left(\frac{R}{1+R}\right) K_i^*(I_i)\right]^{-R} \\
&\times \exp\left(-\left[aR[m_i + (f_i + h_i)p] + \frac{1}{2}\frac{[\mathbf{E}(x|I_i) - Rp]^2}{\mathbf{var}(x|I_i)}\right]\right) \\
= &-\left(\frac{1+R}{R}\right)(\beta R)^{1/1+R} \exp\Bigg(-\frac{1}{1+R}\Bigg[aR[m_i + (f_i + h_i)p] \\
&\qquad + \frac{1}{2}\frac{[\mathbf{E}(x|I_i) - Rp]^2}{\mathbf{var}(x|I_i)}\Bigg]\Bigg) \\
= &- \exp\Bigg(-\frac{1}{1+R}\Bigg[aR[m_i + (f_i + h_i)p] \\
&\qquad + \frac{1}{2}\frac{[\mathbf{E}(x|I_i) - Rp]^2}{\mathbf{var}(x|I_i)} \\
&\qquad - \ln(\beta R) - (1+R)\ln\left(\frac{1+R}{R}\right)\Bigg]\Bigg).
\end{aligned}
$$

All that is needed now is an expression for the market clearing price function p.[29]

A.3.3 Market Clearing

Examination of the optimal investment and consumption functions derived in the previous section reveals that the investment decision is independent of the consumption decision. This is the result of using the negative exponential utility function, which has the characteristic that all investment decisions are independent of each other[30] and of total wealth.

This should have become apparent above when it was shown that the optimal investment decision was the same in both the 'standard' and 'extended' models, and contained no references to beginning of period consumption or wealth. Because of this characteristic of the exponential utility function, the equilibrium price of the risky asset may be found without having to worry about simultaneously satisfying a constraint on aggregate consumption.

[29] As p is contained in the information set I_i, $\mathbf{E}(x|I_i)$ and $\mathbf{var}(x|I_i)$ cannot be evaluated until the functional form of p is known.

[30] Assuming, of course, that the assets invested in are independent of each other. That is, that they are not partial substitutes or complements.

If we let

$$z_I^0 = \frac{\bar{x}_I - Rp}{a\sigma_{xI}^2}, \qquad z_U^0 = \frac{\bar{x}_U - Rp}{a\sigma_{xU}^2},$$

where

$$\mathbf{E}(x|I_i) = \begin{cases} \bar{x}_I = \mathbf{E}(x|\epsilon, p), & \text{if } I_i \text{ is the informed trader information set,} \\ \bar{x}_U = \mathbf{E}(x|p), & \text{if } I_i \text{ is the uninformed trader information set,} \end{cases}$$

$$\mathbf{var}(x|I_i) = \begin{cases} \sigma_{xI}^2 = \mathbf{var}(x|\epsilon, p), & \text{if } I_i \text{ is the informed trader information set,} \\ \sigma_{xU}^2 = \mathbf{var}(x|p), & \text{if } I_i \text{ is the uninformed trader information set,} \end{cases}$$

then

$$z_i^* = \begin{cases} z_I^0 - h_i & \text{if trader } i \text{ is informed,} \\ z_U^0 - h_i & \text{if trader } i \text{ is uninformed.} \end{cases}$$

The next step is to sum the individual demands for the risky asset

$$\begin{aligned} \frac{1}{n}\sum_{i=1}^{n} z_i^* &= \frac{1}{n}\left(\sum_{i\in I} z_i^* + \sum_{i\in U} z_i^*\right) \\ &= \frac{1}{n}\left(n_I z_I^0 + n_U z_U^0 - \sum_{i=1}^{n} h_i\right) \\ &= \lambda z_I^0 + (1-\lambda) z_U^0 - h, \end{aligned}$$

where $i \in I$, and $i \in U$ imply sums over, respectively, all informed and uninformed traders, n_I, n_U and n are the number of informed, uninformed and total traders, and λ is the proportion of the total traders who are informed.[31] To find the market equilibrium, we set this average trader demand equal to the per trader supply, s.

$$\lambda z_I^0 + (1-\lambda) z_U^0 = s + h$$

[31] That is, $\lambda = n_I/n$.

A.3.3.a GENERALIZED MODEL

In the sections above, I discussed three different models corresponding to three different sets of assumptions about the initial endowments received by traders. The first was the Grossman and Stiglitz model, which had endowments of the riskless technology, m_i, and of the risky asset, f_i. The average endowment of the risky asset was normally distributed and equal to the per trader supply of the risky asset.

$$\left.\begin{aligned} &\frac{1}{n}\sum_{i=1}^{n} f_i = f = s \sim N\left\{\mu_f, \sigma_f^2\right\} \\ &h_i = 0, \quad h = 0 \end{aligned}\right\} \text{Grossman and Stiglitz}$$

The second model was the efficient market model. It also had initial endowments of only the riskless and risky assets. The difference is that the endowment of risky asset is constant and equal to the Hakannson, Kunkel and Ohlson (1982) 'no-information' endowments. The endowment is *not* equal to the per trader supply.

$$\left.\begin{aligned} &f_i = f \neq s, \quad f \text{ constant} \\ &s \sim N\left\{f, \sigma_s^2\right\} \\ &h_i = 0, \quad h = 0 \end{aligned}\right\} \text{efficient market}$$

The last model was the hedging model. In this model, iitial endowments of a non-tradable asset, h_i, with payoff equal to the payoff on the risky asset are also received. As in the previous model, the endowment of risky asset is constant. In this model, however, the per trader supply is also constant, and equal to the endowment of risky asset.

$$\left.\begin{array}{l} f_i = f = s, \quad f \text{ constant} \\ \dfrac{1}{n}\displaystyle\sum_{i=1}^{n} h_i = h \sim N\left\{\mu_h, \sigma_h^2\right\} \end{array}\right\} \text{hedging model}$$

We can generalize these three models by noting that the important thing in the market equilibrium condition of the previous section was the sum $s + h$. If we define

$$t = s + h,$$

then we have

$$t \sim \left\{\begin{array}{ll} N\{\mu_f, \sigma_f^2\}, & \text{Grossman and Stiglitz model,} \\ N\{f, \sigma_s^2\}, & \text{efficient market model,} \\ N\{f + \mu_h, \sigma_h^2\}, & \text{hedging model,} \end{array}\right.$$

which allows us to express the market clearing condition in a generalized form

$$\lambda z_I^0 + (1 - \lambda) z_U^0 = t, \qquad t \sim N\{\mu_t, \sigma_t^2\},$$

where μ_t and σ_t depend on the particular model used. In the sections that follow, as much analysis as possible is done using this generalized model. Following the general analysis, results are analyzed for the particular cases of the efficient market model and hedging model. As the Grossman and Stiglitz and hedging models are mathematically identical, no further reference will be made to the Grossman and Stiglitz model.[32]

---

[32] I have chosen to use the hedging model interpretation instead of the Grossman and Stiglitz interpretation in order to avoid the assumption of random endowments of a traded asset. This assumption is difficult to justify, especially in a multiperiod situation.

A.3.3.b MARKET CLEARING PRICE

Up to this point, I have avoided making any assumptions about the market clearing price of the risky asset, except for assuming that it is drawn from a normal distribution.[33] In this section we find a market clearing price function which is normally distributed and satisfies the market clearing condition developed in the previous section. Note that even though we can show existence of this fixed point pricing function – by actually calculating it – we will not have shown uniqueness. There is nothing in the theory which rules out the existence of more than one solution to this fixed point problem.

The price function that we find here is a linear combination of random variables.

$$p = p_0 + p_1\epsilon + p_2(t - \mu_t), \quad p_0, p_1, p_2 \text{ non-stochastic}$$

If this pricing function *is* a fixed point solution, then we can use it to calculate z_I^0 and z_U^0 for use in the market clearing condition. Since x is uncorrelated with t, the informed traders' distribution for x, conditional on their information set, is[34]

$$x|\epsilon, p = x|\epsilon \sim N\left\{\mu_x + \rho\epsilon, \sigma_x^2(1-\rho^2)\right\},$$

and

$$\bar{x}_I = \mathbf{E}(x|\epsilon, p) = \mu_x + \rho\epsilon, \qquad \sigma_{xI}^2 = \mathbf{var}(x|\epsilon, p) = \sigma_x^2(1-\rho^2),$$

giving

$$z_I^0 = \frac{\bar{x}_I - Rp}{a\sigma_{xI}^2} = \frac{\mu_x + \rho\epsilon - R\left(p_0 + p_1\epsilon + p_2(t-\mu_t)\right)}{a\sigma_x^2(1-\rho^2)}.$$

[33] This was needed in section A.3.2 for the calculation of 'post-info' expected utility.
[34] Recall that ϵ is the same for all informed traders.

Similarly, since

$$\begin{pmatrix} x \\ p \end{pmatrix} \sim N\left\{\begin{pmatrix} \mu_x \\ p_0 \end{pmatrix}, \begin{pmatrix} \sigma_x^2 & p_1\sigma_x^2\rho \\ p_1\sigma_x^2\rho & p_1^2\sigma_x^2 + p_2^2\sigma_t^2 \end{pmatrix}\right\},$$

for the uninformed traders, we have

$$x|p \sim N\left\{\mu_x + \frac{p_1\sigma_x^2\rho}{p_1^2\sigma_x^2 + p_2^2\sigma_t^2}[p_1\epsilon + p_2(t-\mu_t)], \sigma_x^2\left(1 - \frac{p_1^2\sigma_x^2\rho^2}{p_1^2\sigma_x^2 + p_2^2\sigma_t^2}\right)\right\},$$

$$\bar{x}_U = \mathbf{E}(x|p) = \mu_x + \phi\rho[\epsilon - \theta(t-\mu_t)], \qquad \sigma_{xU}^2 = \sigma_x^2(1-\phi\rho^2),$$

where

$$\theta = -\frac{p_2}{p_1}, \qquad \phi = \frac{p_1^2\sigma_x^2}{p_1^2\sigma_x^2 + p_2^2\sigma_t^2} = \frac{\sigma_x^2}{\sigma_x^2 + \theta^2\sigma_t^2}.$$

This gives

$$z_U^0 = \frac{\bar{x}_U - Rp}{a\sigma_{xU}^2} = \frac{\mu_x + \phi\rho[\epsilon - \theta(t-\mu_t)] + R[p_0 + p_1\epsilon + p_2(t-\mu_t)]}{a\sigma_x^2(1-\phi\rho^2)}.$$

If we substitute these expressions for z_I^0 and z_U^0 into the market clearing equation, we find

$$\begin{aligned}
t &= \lambda z_I^0 + (1-\lambda)z_U^0 \\
&= \lambda\,\frac{\mu_x + \rho\epsilon - R[p_0 + p_1\epsilon + p_2(t-\mu_t)]}{a\sigma_x^2(1-\rho^2)} \\
&\quad + (1-\lambda)\,\frac{\mu_x + \phi\rho[\epsilon - \theta(t-\mu_t)] + R[p_0 + p_1\epsilon + p_2(t-\mu_t)]}{a\sigma_x^2(1-\phi\rho^2)} \\
&= \frac{\lambda}{a\sigma_x^2(1-\rho^2)}\Big\{[(\mu_x - Rp_0) + \nu\,(\mu_x - Rp_0)] \\
&\qquad + [(\rho - Rp_1) + \nu\,(\phi\rho - Rp_1)]\,\epsilon \\
&\qquad + (-Rp_2) + \nu\,(-\phi\rho\theta - Rp_2)]\,(t-\mu_t)\Big\},
\end{aligned}$$

where

$$\nu = \left(\frac{1-\lambda}{\lambda}\right)\left(\frac{1-\rho^2}{1-\phi\rho^2}\right).$$

Equating the constant terms on each side of the equation,[35] the terms in ϵ and those in $(t-\mu_t)$ gives us expressions for the price function parameters.

$$p_0 = \frac{1}{R}\left(\mu_x - \frac{a\sigma_x^2(1-\rho^2)}{\lambda(1+\nu)}\mu_t\right)$$

$$p_1 = \frac{\rho}{R}\left(\frac{1+\phi\nu}{1+\nu}\right)$$

$$p_2 = -\frac{1}{R(1+\nu)}\left(\frac{a\sigma_x^2(1-\rho^2)}{\lambda} + \nu\phi\rho\theta\right)$$

These parameters can be simplified slightly by recalling the definition of θ, namely, $\theta = -p_2/p_1$. Using the expressions for p_1 and p_2 we find that

$$\theta = \frac{a\sigma_x^2(1-\rho^2)}{\lambda\rho},$$

which we can use to express the price parameters as

$$p_0 = \frac{1}{R}\left(\mu_x - \frac{\theta\rho}{1+\nu}\mu_t\right),$$

$$p_1 = \frac{\rho}{R}\left(\frac{1+\phi\nu}{1+\nu}\right),$$

$$p_2 = -\theta p_1.$$

Notice that we can give ϕ a natural interpretation as the 'informativeness of the price system' [36] as it is equal to the square of the correlation between the price, p, and the signal, ϵ, received by the informed traders.

$$\mathbf{corr}^2(p,\epsilon) = \frac{\mathbf{cov}^2(p,\epsilon)}{\mathbf{var}(p)\mathbf{var}(\epsilon)} = \frac{(p_1\sigma_x^2)^2}{(p_1\sigma_x^2 + p_2^2\sigma_t^2)\,\sigma_x^2} = \phi$$

35 Use $\mu_t + (t-\mu_t)$ instead of t as the left-hand side of the equation.
36 See Grossman and Stiglitz (1980), p. 399.

A.3.4 'Pre-info' Expected Utility

In the sections above, we first derived expressions for traders' 'post-info' expected utility functions assuming only that the price was drawn from a normal distribution. We then found a fixed point market clearing price function which was normally distributed. The final step is to substitute the price function into the expressions for 'post-info' expected utility and take expectations over all the random variables which are not known at the 'pre-info' point in time. That is, we find each trader's expected utility in expectation of the receipt of information.

In order to accomplish this, we need to make use of the following properties of the exponential function:

$$\begin{aligned}\mathbf{E}\left[\exp\left(-[(\gamma b)^2 + \psi b + \omega]\right)\right] &= \frac{1}{\sqrt{1+2\gamma^2}} \exp\left(\frac{1}{2}\frac{\psi^2}{1+2\gamma^2} - \omega\right) \\ &= \exp\left(-\left[\omega - \frac{1}{2}\frac{\psi^2}{1+2\gamma^2} + \frac{1}{2}\ln(1+2\gamma^2)\right]\right), \\ &\quad \text{for } b \sim N\{0,1\},\end{aligned}$$

and

$$\begin{aligned}\mathbf{E}\left[\exp\left(-[(\Gamma' b)^2 + \Psi' b + \Omega]\right)\right] &= \frac{1}{\sqrt{|I+2\Gamma\Gamma'|}} \exp\left(\frac{1}{2}\Psi'(I+2\Gamma\Gamma')^{-1}\Psi - \Omega\right) \\ &= \exp\left(-\left[\Omega - \frac{1}{2}\Psi'(I+2\Gamma\Gamma')^{-1}\Psi + \frac{1}{2}\ln|I+2\Gamma\Gamma'|\right]\right), \\ &\quad \text{for } b \sim N\{0,I\}.\end{aligned}$$

We must reexpress the 'post-info' expected utility functions in terms of standard normal variables and then use the above expressions to find the 'pre-info' expected utility.

A.3.4.a UNINFORMED TRADER

From section A.3.2.c, the 'post-info' expected utility of an uninformed trader in the 'standard' model is

$$J^*_{iU} = -\beta \exp\left(-\left[aR[m_i + (f_i + h_i)p] + \frac{1}{2}\frac{(\bar{x}_U - Rp)^2}{\sigma^2_{xU}}\right]\right),$$

and in the 'extended' model

$$K^*_{iU} = -\exp\left(-\frac{1}{1+R}\left[aR[m_i + (f_i + h_i)p] + \frac{1}{2}\frac{(\bar{x}_U - Rp)^2}{\sigma^2_{xU}} - \ln(\beta R) - (1+R)\ln\left(\frac{1+R}{R}\right)\right]\right).$$

Since the only difference between the 'post-info' and 'pre-info' information sets of an uninformed trader is that the 'post-info' set contains the market clearing price of the risky asset, p, while the 'pre-info' set does not, we can find this trader's 'pre-info' expected utility by calculating

$$\overline{J}^*_{iU} = \mathbf{E}(J^*_{iU}) \qquad \text{or} \qquad \overline{K}^*_{iU} = \mathbf{E}(K^*_{iU}),$$

where the expectation is taken over p.

First, since p is not a standard normal variable,

$$\begin{aligned} p &\sim N\{p_0, p_1^2(\sigma_x^2 + \theta^2\sigma_t^2)\} \\ &\sim N\{p_0, p_1^2\sigma_x^2/\phi\}, \end{aligned}$$

we define a transformation of p which is.

$$b = \left(\frac{p - p_0}{p_1}\right)\frac{\sqrt{\phi}}{\sigma_x} \sim N\{0, 1\}$$

That is, we can substitute

$$p = p_0 + p_1 \frac{\sigma_x}{\sqrt{\phi}} b$$

into the expressions above, in order to have the 'post-info' expected utility functions expressed in terms of the standard normal variable b.

This is not all we need, however. We also must have an expression for $\bar{x}_U$ in terms of b. Since we showed in section A.3.3.b that

$$\bar{x}_U = \mu_x + \phi\rho[\epsilon - \theta(t - \mu_t)], \qquad \sigma_{xU}^2 = \sigma_x^2(1 - \phi\rho^2),$$

it is easily shown that

$$\bar{x}_U = \mu_x + \sigma_x \rho \sqrt{\phi}\, b.$$

Making the necessary substitutions, we find that

$$\begin{aligned} J_{iU}^* = -\beta \exp\Bigg(- \Bigg[aR \left(m_i + (f_i + h_i) \left(p_0 + p_1 \frac{\sigma_x}{\sqrt{\phi}} b \right) \right) \\ + \frac{1}{2} \frac{1}{\sigma_x^2(1 - \phi\rho^2)} \left((\mu_x - Rp_0) + \frac{\sigma_x}{\sqrt{\phi}} (\phi\rho - Rp_1)\, b \right)^2 \Bigg] \Bigg), \end{aligned}$$

which is equivalent to

$$J_{iU}^* = -\beta \exp(-[(\gamma b)^2 + \psi b + \omega]),$$

where

$$\begin{aligned} \gamma^2 &= \frac{1}{2} \frac{(\phi\rho - Rp_1)^2}{\phi(1 - \phi\rho^2)}, \\ \psi &= \frac{\sigma_x}{\sqrt{\phi}} \left(\frac{(\mu_x - Rp_0)(\phi\rho - Rp_1)}{\sigma_x^2(1 - \phi\rho^2)} + aR(f_i + h_i)p_1 \right), \\ \omega &= \frac{1}{2} \frac{(\mu_x - Rp_0)^2}{\sigma_x^2(1 - \phi\rho^2)} \\ &\quad + aR[m_i + (f_i + h_i)p_0]. \end{aligned}$$

Using the formula given in the previous section, we now know that the 'pre-info' expected utility of an uninformed trader in the 'standard' model is

$$\overline{J}^*_{iU} = \mathbf{E}(J^*_{iU}) = -\exp\left(-\left[\omega - \ln\beta - \frac{1}{2}\frac{\psi^2}{1+2\gamma^2} + \frac{1}{2}\ln(1+2\gamma^2)\right]\right),$$

where γ, ψ and ω are as given above.

The same procedure can be followed for the 'extended' model, resulting in

$$\overline{K}^*_{iU} = \mathbf{E}(K^*_{iU}) = -\exp\left(-\left[\frac{1}{1+R}\left(\omega - \ln(\beta R) - (1+R)\ln\left(\frac{1+R}{R}\right)\right)\right.\right.$$
$$\left.\left. - \frac{1}{2}\frac{\psi^2/(1+R)^2}{1+2\gamma^2/(1+R)} + \frac{1}{2}\ln\left(1+\frac{2\gamma^2}{1+R}\right)\right]\right),$$

where γ, ψ and ω are the same as given above for the 'standard' model.

A.3.4.b INFORMED TRADER

From section A.3.2.c, the 'post-info' expected utility of an informed trader in the 'standard' model is

$$J^*_{iI} = -\exp\left(-\left[aR[m_i + (f_i + h_i)p] - \ln\beta + \frac{1}{2}\frac{(\bar{x}_I - Rp)^2}{\sigma^2_{xI}}\right]\right),$$

and in the 'extended' model

$$K^*_{iI} = -\exp\left(-\frac{1}{1+R}\left[aR[m_i + (f_i + h_i)p] + \frac{1}{2}\frac{(\bar{x}_I - Rp)^2}{\sigma^2_{xI}}\right.\right.$$
$$\left.\left. - \ln(\beta R) - (1+R)\ln\left(\frac{1+R}{R}\right)\right]\right).$$

The difference between the 'post-info' and 'pre-info' information sets of an informed trader is that the 'post-info' set contains the informed trader signal, ϵ, and the market clearing price of the risky asset, p, while the 'pre-info' set does not. If we note that the price of the risky asset is a simple linear combination of ϵ and t, knowing both ϵ and p is equivalent to knowing both ϵ and t. Therefore, the expectation of J^*_{iI} taken over all ϵ and p is identical to the expectation taken over all ϵ and t. The latter pair of variables will be used, as it is easier to calculate the expectations

$$\overline{J}^*_{iI} = \mathbf{E}(J^*_{iI}) \qquad \text{and} \qquad \overline{K}^*_{iI} = \mathbf{E}(K^*_{iI}),$$

where the expectation is taken over all ϵ and t.

First, since the vector of ϵ and t is not standard normal,

$$\begin{pmatrix} \epsilon \\ t \end{pmatrix} \sim N\left\{ \begin{pmatrix} 0 \\ \mu_t \end{pmatrix}, \begin{pmatrix} \sigma_x^2 & 0 \\ 0 & \sigma_t^2 \end{pmatrix} \right\},$$

we define a transformation which is.

$$b = \begin{pmatrix} \sigma_x & 0 \\ 0 & \sigma_t \end{pmatrix}^{-1} \begin{pmatrix} \epsilon \\ t - \mu_t \end{pmatrix} \sim N\{0, I\}$$

That is, we substitute

$$\begin{aligned} p &= p_0 + p_1(1, -\theta) \begin{pmatrix} \epsilon \\ t - \mu_t \end{pmatrix} \\ &= p_0 + p_1(1, -\theta) \begin{pmatrix} \sigma_x & 0 \\ 0 & \sigma_t \end{pmatrix} b \\ &= p_0 + p_1(\sigma_x, -\sigma_t\theta)\, b, \end{aligned}$$

into the expressions above, in order to have the 'post-info' expected utility functions expressed in terms of the standard normal variable b.

As before, we also must have an expression for $\bar{x}_I$ in terms of b. Using the expressions derived in section A.3.3.b,

$$\bar{x}_I = \mu_x + \rho\epsilon, \qquad \sigma_{xI}^2 = \sigma_x^2(1-\rho^2),$$

we can show that

$$\bar{x}_I = \mu_x + (\sigma_x \rho, 0)\, b.$$

If we make the necessary substitutions, we find

$$J_{iI}^* = -\exp\Bigg(-\Bigg[aR[m_i + (f_i + h_i)p_0] - \ln\beta + aR(f_i + h_i)p_1(\sigma_x, -\sigma_t\theta)\, b$$
$$+ \frac{1}{2}\frac{1}{\sigma_x^2(1-\rho^2)}\left[(\mu_x - Rp_0) + (\sigma_x(\rho - Rp_1), \sigma_t R p_1 \theta)\, b\right]^2\Bigg]\Bigg).$$

This is equivalent to

$$J_{iI}^* = -\beta\, \exp(-[(\Gamma' b)^2 + \Psi' b + \Omega]),$$

where

$$\Gamma = \frac{1}{\sqrt{2\sigma_x^2(1-\rho^2)}} \begin{pmatrix} \sigma_x(\rho - Rp_1) \\ \sigma_t R p_1 \theta \end{pmatrix},$$

$$\Psi = \frac{(\mu_x - Rp_0)}{\sigma_x^2(1-\rho^2)} \begin{pmatrix} \sigma_x(\rho - Rp_1) \\ \sigma_t R p_1 \theta \end{pmatrix} + aR(f_i + h_i)p_1 \begin{pmatrix} \sigma_x \\ -\sigma_t \theta \end{pmatrix},$$

$$\Omega = \frac{1}{2}\frac{(\mu_x - Rp_0)^2}{\sigma_x^2(1-\rho^2)} + aR[m_i + (f_i + h_i)p_0].$$

Using the formula given in section A.3.4, we now know that the 'pre-info' expected utility of an informed trader in the 'standard' model is

$$\overline{J}_{iI}^* = \mathbf{E}(J_{iI}^*) = -\exp\left(-\left[\Omega - \ln\beta - \frac{1}{2}\Psi'(I + 2\Gamma\Gamma')^{-1}\Psi + \frac{1}{2}\ln|I + 2\Gamma\Gamma'|\right]\right),$$

where Γ, Ψ and Ω are as given above.

The same procedure can be followed for the 'extended' model, resulting in

$$\overline{K}^*_{iI} = \mathbf{E}(K^*_{iI}) = -\exp\left(-\left[\frac{1}{1+R}\left(\Omega - \ln(\beta R) - (1+R)\ln\left(\frac{1+R}{R}\right)\right)\right.\right.$$
$$-\frac{1}{2}\frac{\Psi'}{1+R}\left(I + 2\frac{\Gamma\Gamma'}{1+R}\right)^{-1}\frac{\Psi}{1+R}$$
$$\left.\left.+\frac{1}{2}\ln\left|I + 2\frac{\Gamma\Gamma'}{1+R}\right|\right]\right),$$

where Γ, Ψ and Ω are the same as given above for the 'standard' model.

A.4 MODEL INTERPRETATION

Each step in the unfolding of a model may be difficult in its own way. The derivation we have just gone through was difficult because the algebraic manipulations were complicated and unenlightening. This next step is also difficult, but in a different way. We have to use the formulas developed in the previous step to make model predictions or descriptions, and then these predictions or descriptions must be interpreted. We must present an intuitive argument which makes the same predictions as the purely mathematical one and which correctly captures the interactions between elements in the mathematical model.

The first requirement, namely, providing an intuitive argument giving the same predictions as the mathematical model is not the difficult part. What is difficult is providing a correct intuitive argument, where by correct I mean not only paralleling some final prediction, but also paralleling the actual dynamics whereby that prediction is produced. In the following sections, I aim to develop an understanding of the dynamics of the models, that is, of the interactions of each of the assumptions making up the models. This understanding will automatically build up to understanding of the more complex model predictions.

With this aim in mind, I have divided the interpretation into three cases. In case 1, I present characteristics of the models when all traders are informed (that is, we do *not* have an asymmetric information model), and there is *no* random supply (that is, endowments of the non-tradable asset, h_i, and risky asset supply are *not* random). This case introduces the first of the factors which underlie the dynamics of these models: the allocative efficiency benefit. As information quality increases we may find that utility increases due to a more efficient allocation between current and future consumption. As shown in Tables A.1 and A.2, this benefit does not arise in the 'standard' model – there is no current consumption in that model. In the 'extended' model, however, this factor can be identified.

The second case adds one element of complexity by introducing a random supply. In this case, however, we still do not have an asymmetric information model. This case introduces the rest of the factors needed to understand the model dynamics. An analysis of the efficient market model brings to light the reason why the market modelled is not actually efficient. We see that the relation between the 'naive' efficient market traders and the rational expectations traders is similar to that of an ordinary trader to his dealer. As the dealer has better information than his clients – in particular, knowledge of unexpected demand/supply variantions – he is able to use his own inventory of risky asset to supply unexpectedly high demand and absorb unexpectedly high supply. I name this benefit the rational expectations traders receive in return for this service the dealer benefit (see Table A.1).

Two additional factors are found in the analysis of the hedging model in this second case. As one of the motivations for trading is the desire to hedge one's position in other non-tradable assets, it is not surprising to find a factor which we can identify as a benefit from the opportunity to hedge. This hedging benefit is shown to be analogous to the dealer benefit which arose in the efficient market model. In addition, a factor is found which reflects the risk of market revaluation of one's endowment of non-tradable asset. After information release, when one does finally have the opportunity to hedge, the resulting benefits are diminished relative to what they would have been had no information been released. This decrease in the benefits from hedging I name revaluation risk.

In the third case, the uninformed rational expectations trader is introduced, thus giving us an asymmetric information case. This is the most interesting case, but also the most difficult to analyze. Tables A.1 and A.2 summarize the results of the analysis of this case for the efficient market model and hedging model, respectively.

A.4.1 Case 1: Homogeneous Information, Non-Random Supply

In terms of model variables, this case has no uninformed traders,[37]

$$\lambda = 1,$$

and has endowments of only the riskless and risky assets. The endowment of the risky asset is the Hakannson, Kunkel and Ohlson 'no information' endowment.

$$f_i = f, \qquad h_i = 0$$

The generalized randomness variable t is also constant and equal to the endowment of risky asset.

$$t = f, \qquad \mu_t = f, \qquad \sigma_t = 0$$

[37] By non-random supply I mean that the endowments of the non-tradable asset, h_i, and the risky asset supply are constant.

Several immediate consequences of these assumptions are[38]

$$\theta = \frac{a\sigma_x^2(1-\rho^2)}{\lambda\rho} = \frac{a\sigma_x^2(1-\rho^2)}{\rho},$$

$$\phi = \frac{\sigma_x^2}{\sigma_x^2 + \theta^2\sigma_t^2} = 1,$$

$$\nu = \left(\frac{1-\lambda}{\lambda}\right)\left(\frac{1-\rho^2}{1-\phi\rho^2}\right) = 0,$$

$$p_0 = \frac{1}{R}\left(\mu_x - \frac{\theta\rho}{1+\nu}\mu_t\right) = \frac{1}{R}[\mu_x - a\sigma_x^2(1-\rho^2)\, f],$$

$$p_1 = \frac{\rho}{R}\left(\frac{1+\phi\nu}{1+\nu}\right) = \frac{\rho}{R},$$

$$p_2 = -\theta p_1.$$

It is a simple matter to substitute the above into the general expressions for the 'pre-info' expected utility of an informed trader to give

$$\Gamma = \frac{1}{\sqrt{2\sigma_x^2(1-\rho^2)}}\begin{pmatrix}\sigma_x(\rho - Rp_1)\\ \sigma_t Rp_1\theta\end{pmatrix} = \begin{pmatrix}0\\0\end{pmatrix},$$

$$\Psi = \frac{(\mu_x - Rp_0)}{\sigma_x^2(1-\rho^2)}\begin{pmatrix}\sigma_x(\rho - Rp_1)\\ \sigma_t Rp_1\theta\end{pmatrix} + aR(f_i + h_i)p_1\begin{pmatrix}\sigma_x\\ -\sigma_t\theta\end{pmatrix} = \begin{pmatrix}af\rho\sigma_x\\ 0\end{pmatrix},$$

$$\begin{aligned}\Omega &= aR[m_i + (f_i + h_i)p_0] + \frac{1}{2}\frac{(\mu_x - Rp_0)^2}{\sigma_x^2(1-\rho^2)}\\ &= aRm_i + af\mu_x - \frac{1}{2}a^2\sigma_x^2(1-\rho^2)\, f^2,\end{aligned}$$

which, when substituted into the 'pre-info' expected utility function for the 'standard' model, give

$$\begin{aligned}\overline{J}_{iI}^* &= -\exp\left(-\left[\Omega - \ln\beta - \frac{1}{2}\Psi'(I + 2\Gamma\Gamma')^{-1}\Psi + \frac{1}{2}\ln|I + 2\Gamma\Gamma'|\right]\right)\\ &= -\exp\left(-\left[aRm_i + af\mu_x - \frac{1}{2}a^2\sigma_x^2 f^2 - \ln\beta\right]\right).\end{aligned}$$

[38] Recall that $\phi = \mathbf{corr}^2(p, \epsilon)$, so that $\phi = 1$ means that the price fully reveals the information ϵ.

Similarly, when substituted into the 'pre-info' expected utility function for the 'extended' model, we find

$$
\begin{aligned}
\overline{K}^*_{iI} = &-\exp\Bigg(-\Bigg[\frac{1}{1+R}\left(\Omega - \ln(\beta R) - (1+R)\ln\left(\frac{1+R}{R}\right)\right) \\
&\qquad - \frac{1}{2}\frac{\Psi'}{1+R}\left(I + 2\frac{\Gamma\Gamma'}{1+R}\right)^{-1}\frac{\Psi}{1+R} \\
&\qquad + \frac{1}{2}\ln\left|I + 2\frac{\Gamma\Gamma'}{1+R}\right|\Bigg]\Bigg) \\
= &-\exp\Bigg(-\Bigg[\frac{1}{1+R}\left(\Omega - \ln(\beta R) - (1+R)\ln\left(\frac{1+R}{R}\right)\right) \\
&\qquad - \frac{1}{2}\frac{a^2\sigma_x^2\rho^2 f^2}{(1+R)^2}\Bigg]\Bigg) \\
= &-\exp\Bigg(-\frac{1}{1+R}\Bigg[aRm_i + af\mu_x - \ln(\beta R) - (1+R)\ln\left(\frac{1+R}{R}\right) \\
&\qquad - \frac{1}{2}a^2\sigma_x^2 f^2\left(1 - \frac{R}{(1+R)}\rho^2\right)\Bigg]\Bigg).
\end{aligned}
$$

Now that we have explicit expressions for the 'pre-info' expected utility, we can see what the effects are when we increase the quality or 'informativeness' [39] of the signal given to these informed traders. It is easy to see that

$$
\frac{\partial}{\partial\rho^2}\overline{J}^*_{iI} = 0, \qquad \frac{\partial}{\partial\rho^2}\overline{K}^*_{iI} \geq 0.
$$

That is, using the terminology of Hakannson, Kunkel and Ohlson (1982) or Hirshleifer (1971), the social value of information is always zero in the 'standard' model and positive in the 'extended' model.[40]

[39] The correlation between the payoff on the risky asset and the signal received by the informed traders, ρ, will be referred to as the quality or 'informativeness' of the signal. This is not the same as the 'informativeness' of the price system, ϕ.

[40] This result for the 'extended' model was shown by Epstein and Turnbull (1980).

This result, of course, already throws into doubt the intuition that better information makes traders better off. As Hirshleifer (1971) pointed out, however, this result should not be too surprising. Information really has no intrinsic worth – these traders can't eat it – and only has a *derived* value when it can have an effect on the allocation of goods in the economy. The flat expected utility curve in the 'standard' model merely points out that our informed traders can't use their information to create a better allocation of their wealth. If they receive information that the future payoff on the risky asset will be poor, their immediate reaction is to sell at the currently high price and buy more riskless technology with the proceeds. But, since *all* traders are informed, they all want to sell. This depresses the price of the risky asset far enough that everyone decides to retain their holdings.

That explains the flat expected utility curve of the 'standard' model, but the same argument does not appear to apply in the 'extended' model. The argument used to explain the 'standard' model offers only one possible reason for an increase in utility as the quality of information increases. The better information must be allowing a more efficient allocation of wealth. Of course, there is no possibility for a trader to change his holdings of risky asset, since this is a homogeneous information economy, so the effect must be due to the only other investment in the economy: the riskless technology.

What is happening is that the allocation of wealth between beginning of period consumption and investment in the riskless technology is more efficient given better information.[41] This is an effect of not fixing the aggregate supply of riskless

[41] Returning to Table A.1, the results of this case are shown in the first column.

technology. If the supply were fixed, then we would have the same situation as we have with the risky asset, namely, attempts to change holdings of the riskless technology leading only to a price adjustment. The effect, therefore, is not an unrealistic one. We would expect the number of shares of a particular company to be insensitive to demand, which is consistent with the modelling of the risky asset, but the total supply of alternative investments might be sensitive to demand.[42]

We can make an analogy between this and the flexibility of investment plans. We would not expect information to have any value in an economy with totally inflexible investment plans or opportunities. There has to be the possibility of increasing investment in some assets and cutting back investment in others before we would expect information effects in a homogeneous information economy. This possibility is provided by the perfectly elastic supply of the riskless technology, though it is not the fact that the supply is perfectly elastic that is important. What is important is that the supply of riskless technology is *not* perfectly *in*elastic.[43]

In conclusion, we can expect the 'standard' and 'expected' models to provide different conclusions. In the 'standard' model, because traders consume only at one point in time there is no possibility of trading off current consumption against future consumption. That is, there are no possibilities for increasing allocative efficiency. In the 'extended' model, since consumption occurs at two points in time,

---

This case does not involve dealer benefit.

[42] The supply curve for the risky asset, or for a stock, is vertical. The supply curve for total available investments is unlikely to be vertical.

[43] That is, what is important is that the supply curve for the riskless asset is not perfectly vertical, not that it is perfectly flat.

there is the possibility of foregoing current consumption in return for increased future consumption. At least, the opportunity is present as long as there is at least one investment vehicle which is not in fixed supply. In the 'extended' model this function is provided by the riskless technology.

A.4.2 Case 2: Homogeneous Information, Random Supply

In analyzing this case,[44] I first present the mathematical analysis in terms of the generalized model. In separate sections following this, the results are interpreted for the efficient market model and the hedging model.

As in the previous case, we still have no uninformed traders, that is,

$$\lambda = 1,$$

but the generalized random variable t is no longer constant,

$$t \sim N\{\mu_t, \sigma_t^2\}.$$

The assumptions regarding endowments depend upon the specific model chosen.

[44] By random supply, I mean that either the supply of the risky asset is random, as in the efficient market model, or the endowments of the non-tradable asset, h_i, are random, as in the hedging model.

The immediate consequences of these assumptions are[45]

$$
\begin{aligned}
\theta &= \frac{a\sigma_x^2(1-\rho^2)}{\lambda\rho} = \frac{a\sigma_x^2(1-\rho^2)}{\rho}, \\
\nu &= \left(\frac{1-\lambda}{\lambda}\right)\left(\frac{1-\rho^2}{1-\phi\rho^2}\right) = 0, \\
p_0 &= \frac{1}{R}\left(\mu_x - \frac{\theta\rho}{1+\nu}\,\mu_t\right) = \frac{1}{R}[\mu_x - a\sigma_x^2(1-\rho^2)\,\mu_t], \\
p_1 &= \frac{\rho}{R}\left(\frac{1+\phi\nu}{1+\nu}\right) = \frac{\rho}{R}, \\
p_2 &= -\theta p_1.
\end{aligned}
$$

Notice that the price function of the risky asset is identical to what it was in case 1, where we had $\sigma_t = 0$. In the previous case, however, the value of p_2 really didn't matter, since $(t - \mu_t)$ was constrained to be zero. What we see in this case, therefore, is the addition of a non-zero term to the pricing function.

Once again, we must substitute the above into the general expressions for the

45 Unlike the previous case, the price is not fully revealing. That is, $0 \le \phi < 1$.

'pre-info' expected utility of an informed trader.

$$\Gamma = \frac{1}{\sqrt{2\sigma_x^2(1-\rho^2)}}\begin{pmatrix}\sigma_x(\rho - Rp_1)\\ \sigma_t Rp_1\theta\end{pmatrix} = \begin{pmatrix}0\\ a\sigma_x^2\sigma_t(1-\rho^2)\end{pmatrix}$$

$$\begin{aligned}\Psi &= \frac{(\mu_x - Rp_0)}{\sigma_x^2(1-\rho^2)}\begin{pmatrix}\sigma_x(\rho - Rp_1)\\ \sigma_t Rp_1\theta\end{pmatrix} + aRp_1(f_i+h_i)\begin{pmatrix}\sigma_x\\ -\sigma_t\theta\end{pmatrix}\\ &= \begin{pmatrix}a(f_i+h_i)\rho\sigma_x\\ a^2\sigma_x^2\sigma_t(1-\rho^2)[\mu_t-(f_i+h_i)]\end{pmatrix}\end{aligned}$$

$$\begin{aligned}\Omega &= aR[m_i + (f_i+h_i)p_0] + \frac{1}{2}\frac{(\mu_x - Rp_0)^2}{\sigma_x^2(1-\rho^2)}\\ &= aRm_i + a\mu_x(f_i+h_i)\\ &\quad + \frac{1}{2}a^2\sigma_x^2(1-\rho^2)[\mu_t^2 - 2\mu_t(fi+h_i)]\end{aligned}$$

We can find the expected utility for the 'standard' model by noting that

$$\Psi'(I+2\Gamma\Gamma')^{-1}\Psi = a^2\sigma_x^2\rho^2(f_i+h_i)^2 + \frac{a^4\sigma_x^4\sigma_t^2(1-\rho^2)^2}{1+a^2\sigma_x^2\sigma_t^2(1-\rho^2)}[\mu_t-(f_i+h_i)]^2,$$

$$\begin{aligned}&\Omega - \frac{1}{2}\Psi'(I+2\Gamma\Gamma')^{-1}\Psi\\ &\quad = aRm_i + a\mu_x(f_i+h_i) - \frac{1}{2}a^2\sigma_x^2(f_i+h_i)^2\\ &\quad\quad + \frac{1}{2}a^2\sigma_x^2(1-\rho^2)[\mu_t-(f_i+h_i)]^2\left(1 - \frac{a^2\sigma_x^2\sigma_t^2(1-\rho^2)}{1+a^2\sigma_x^2\sigma_t^2(1-\rho^2)}\right)\\ &\quad = aRm_i + a\mu_x(f_i+h_i) - \frac{1}{2}a^2\sigma_x^2(f_i+h_i)^2\\ &\quad\quad + \frac{1}{2}\left(\frac{a^2\sigma_x^2(1-\rho^2)}{1+a^2\sigma_x^2\sigma_t^2(1-\rho^2)}\right)[\mu_t-(f_i+h_i)]^2.\end{aligned}$$

From this, it is easily seen that both[46]

$$\frac{\partial}{\partial\rho^2}\left(\Omega - \frac{1}{2}\Psi'(I+2\Gamma\Gamma')^{-1}\Psi\right)\quad \begin{cases}=0, & \text{efficient market model,}\\ <0, & \text{hedging model,}\end{cases}$$

and

$$\frac{\partial}{\partial\rho^2}\ln|I+2\Gamma\Gamma'| = \frac{\partial}{\partial\rho^2}\ln[1+a^2\sigma_x^2\sigma_t^2(1-\rho^2)] < 0.$$

[46] The efficient market model has $\mu_t - (f_i+h_i) = 0$.

Since

$$\overline{J}^*_{iI} = -\exp\left(-\left[\Omega - \ln\beta - \frac{1}{2}\Psi'(I+2\Gamma\Gamma')^{-1}\Psi + \frac{1}{2}\ln|I+2\Gamma\Gamma'|\right]\right),$$

for the 'standard' model, we have unambiguously[47]

$$\frac{\partial}{\partial\rho^2}\overline{J}^*_{iI} < 0.$$

Similarly, for the 'extended' model, we have

$$\begin{aligned}
&\Psi'\left(I+2\frac{\Gamma\Gamma'}{1+R}\right)^{-1}\Psi \\
&\qquad = a^2\sigma_x^2\rho^2(f_i+h_i)^2 + \frac{(1+R)a^4\sigma_x^4\sigma_t^2(1-\rho^2)^2}{(1+R)+a^2\sigma_x^2\sigma_t^2(1-\rho^2)}\left[\mu_t-(f_i+h_i)\right]^2,
\end{aligned}$$

and

$$\begin{aligned}
&\frac{\Omega}{1+R} - \frac{1}{2}\frac{\Psi'}{1+R}\left(I+2\frac{\Gamma\Gamma'}{1+R}\right)^{-1}\frac{\Psi}{1+R} \\
&\quad = \frac{1}{1+R}\left(aRm_i + a\mu_x(f_i+h_i) - \frac{1}{2}\left(\frac{1}{1+R}\right)a^2\sigma_x^2(f_i+h_i)^2\right) \\
&\quad + \frac{1}{2}\frac{a^2\sigma_x^2(1-\rho^2)}{(1+R)+a^2\sigma_x^2\sigma_t^2(1-\rho^2)}\left[\mu_t-(f_i+h_i)\right]^2 \\
&\quad - \frac{1}{2}\frac{R}{(1+R)^2}a^2\sigma_x^2(1-\rho^2)(f_i+h_i)^2.
\end{aligned}$$

From this we can see that

$$\begin{aligned}
&\frac{\partial}{\partial\rho^2}\left(\frac{\Omega}{1+R} - \frac{1}{2}\frac{\Psi'}{1+R}\left(I+2\frac{\Gamma\Gamma'}{1+R}\right)^{-1}\frac{\Psi}{1+R}\right) \\
&\qquad = -\frac{1}{2}\frac{(1+R)a^2\sigma_x^2}{((1+R)+a^2\sigma_x^2\sigma_t^2(1-\rho^2))^2}\left[\mu_t-(f_i+h_i)\right]^2 \\
&\qquad + \frac{1}{2}\frac{R}{(1+R)^2}a^2\sigma_x^2(f_i+h_i)^2 \\
&\qquad \begin{cases} >0, & \text{efficient market model,} \\ <>0, & \text{hedging model,} \end{cases}
\end{aligned}$$

---

[47] This result and those following are interpreted in the immediately following sections.

which may be positive or negative, and

$$\frac{\partial}{\partial\rho^2}\ln\left|I+2\frac{\Gamma\Gamma'}{1+R}\right| = \frac{\partial}{\partial\rho^2}\ln\left(1+\left(\frac{1}{1+R}\right)a^2\sigma_x^2\sigma_t^2(1-\rho^2)\right)$$
$$= -\frac{a^2\sigma_x^2\sigma_t^2}{(1+R)+a^2\sigma_x^2\sigma_t^2(1-\rho^2)}$$
$$< 0.$$

Since the first partial derivative above may be positive or negative, and

$$\overline{K}_{iI}^* = -\exp\left(-\left[\frac{1}{1+R}\left(\Omega-\ln(\beta R)-(1+R)\ln\left(\frac{1+R}{R}\right)\right)\right.\right.$$
$$-\frac{1}{2}\frac{\Psi'}{1+R}\left(I+2\frac{\Gamma\Gamma'}{1+R}\right)^{-1}\frac{\Psi}{1+R}$$
$$\left.\left.+\frac{1}{2}\ln\left|I+2\frac{\Gamma\Gamma'}{1+R}\right|\right]\right),$$

the partial derivative,

$$\frac{\partial}{\partial\rho^2}\overline{K}_{iI}^* <> 0,$$

may likewise be positive or negative.[48]

A.4.2.a EFFICIENT MARKET MODEL

In this model, the endowment of risky asset is constant and equal for all traders:[49]

$$f_i = f = \text{constant}$$

Even though the endowments are not random, the needed randomness is provided by a group of non-rational expectations traders outside the model, contributing a

[48] This is true for both the efficient market model and the hedging model.
[49] The endowment h_i is zero.

random supply component.

$$t = s \sim N\{f, \sigma_s^2\}$$

The previous section seems quite clear about its prediction for the 'standard' model. There is no possibility of an increase in expected utility for these traders if we increase the quality of their information. From the discussion of the previous case, however, this is not surprising. There is, after all, no possibility of trading off current consumption - none is allowed - against future consumption. If we look at the terms in the expression for expected utility,[50]

$$\begin{aligned}
&\frac{\partial}{\partial \rho^2}\left(\Omega - \frac{1}{2}\Psi'(I + 2\Gamma\Gamma')^{-1}\Psi\right) \\
&\quad = \frac{\partial}{\partial \rho^2}\left(aRm_i + a\mu_x f - \frac{1}{2}a^2\sigma_x^2 f^2\right) \\
&\quad = 0,
\end{aligned}$$

$$\frac{\partial}{\partial \rho^2}\ln|I + 2\Gamma\Gamma'| = \frac{\partial}{\partial \rho^2}\ln(1 + a^2\sigma_x^2\sigma_s^2(1-\rho^2)) < 0,$$

we see that there is one term causing a decrease in expected utility as we increase quality of information.

In order to interpret this effect, notice that expected utility *increases* as we increase the supply variance, σ_s^2. What we are seeing here is simply due to a difference in outlook. In this model, the 'naive' or efficient market trader has the belief that as long as he is buying at the equilibrium price then the purchase price does not matter. He believes the asset is 'correctly' priced at equilibrium.

50 Substitution of $(\mu_t - (f_i + h_i)) = 0$ has been made. These expressions are from the previous section.

The rational expectations traders have a different view of the situation. In particular, they know that the equilibrium pricing function for the risky asset is

$$p = p_0 + p_1\epsilon + p_2(t - \mu_t),$$

where $p_1 = 0$ when $\rho = 0$, increasing to $1/R$ as ρ approaches 1. Also, $p_2 < 0$ when $\rho = 0$, increasing to $p_2 = 0$ as ρ increases to 1. It is clear, then, that the equilibrium price tells the rational expectations trader something that the 'naive' trader doesn't even consider, namely, the aggregate supply of the risky asset.

The risky asset, like any other asset, has a price that falls when supply rises and rises when supply falls, something that the rational expectations traders realize. Basically, they sell the risky asset to the 'naive' traders when there is a small supply of it or buy it when supply is high and make a profit on doing so.[51] They realize that the price fluctuates due to both information and supply effects while the 'naive' investor believes that the price only changes because of new information entering the market.

When ρ equals 1, the rational expectation traders have perfect information about the future payoff on the risky asset, thus making it *riskless* in their eyes. Because we implicitly assumed competition between traders, they will ensure that this now riskless technology is priced at the riskless technology price. This happens to make the price perfectly insensitive to supply fluctuations, just like the riskless technology price. What is happening, therefore, is that as the risky asset becomes

[51] Essentially, the rational expectations traders are acting as dealers. From their position, they can note fluctuations due to 'fads', or other effects, and use their stock of risky asset to satisfy demand when it is abnormally high and absorb abnormally high supply. In return for this service they realize a profit.

less risky – in the eyes of the informed rational expectations traders – its price becomes less sensitive to supply variability, which reduces the potential gains that the rational expectations traders stand to make by buying low or selling high to the 'naive' traders. That is, 'pre-info' expected utility decreases as ρ is increased because the dealer benefit decreases.[52]

What we find when we turn to the 'extended' model is not surprising. The equations show that the term tending to increase expected utility due to a more efficient allocation of current and future consumption is once more present,

$$\begin{aligned}\frac{\partial}{\partial \rho^2}\left(\frac{\Omega}{1+R} - \frac{1}{2}\frac{\Psi'}{1+R}\left(I + 2\frac{\Gamma\Gamma'}{1+R}\right)^{-1}\frac{\Psi}{1+R}\right) \\ = \frac{1}{2}\frac{R}{(1+R)^2}a^2\sigma_x^2 f^2 \\ > 0,\end{aligned}$$

and that we still have the term from the 'standard' model which depresses expected utility as we increase quality of information.

$$\begin{aligned}\frac{\partial}{\partial \rho^2}\ln\left|I + 2\frac{\Gamma\Gamma'}{1+R}\right| &= \frac{\partial}{\partial \rho^2}\ln\left(1 + \left(\frac{1}{1+R}\right)a^2\sigma_x^2\sigma_s^2(1-\rho^2)\right) \\ &= -\frac{a^2\sigma_x^2\sigma_s^2}{(1+R) + a^2\sigma_x^2\sigma_s^2(1-\rho^2)} < 0\end{aligned}$$

The result, therefore, depends on the size of the supply variance, σ_s^2. For small σ_s^2, the benefits due to trading with the 'naive' traders are small to begin with, so

[52] Returning to Table A.1, we have now moved from the first column – allocative efficiency – to the second – dealer benefit. To this point, we have discussed only the top half of the column, that is, the 'standard' model. Note that the positive signs shown in this column apply only to case 3, which involves asymmetric information. In the current, homogeneous information case, this column becomes strictly negative.

the loss of dealer benefit due to better quality information is small. For σ_s^2 small enough, increasing the quality of information will cause an increase in expected utility due to a better allocation of current and future consumption. Otherwise, the net effect will be a decrease in utility due to decreased supply sensitivity of the risky asset price, and a consequent drop in the benefits of trading with the group of 'naive' traders.

A.4.2.b HEDGING MODEL

The 'standard', hedging model also allows no possibility of an increase in 'pre-info' expected utility if we increase the quality of information. As was discussed in the efficient market model section above, this is due to not having the opportunity to trade off current consumption against future consumption. The reason in this model for a decrease in expected utility given better quality information is that we have replaced the situation we had in case 1, namely, constant endowments of the risky asset, with a situation where the endowments are random.

Since endowments are random, trading between rational expectations traders will be necessary to bring about market equilibrium. This is unlike the previous efficient market model case above, where no trading between rational expectations traders took place, only trading between rational expectations traders and 'naive' traders. Along with the need to trade with other rational expectations traders in the hedging model comes vulnerability to information which changes those traders' perceptions of the asset one wishes to trade.

No matter what endowment one has, it is certain that one will have to do some trading with other rational expectations traders. Before one has the opportunity to trade, however, information will be disseminated which tells everyone either that the asset you want to trade is desirable or that it is undesirable. One is exposed to the risk of revaluation of one's endowment. As the quality of information given out is increased, the extremes of revaluation become more probable, thus increasing the revaluation risk. This, naturally, is the reason for the decrease in expected utility in the 'standard' model.[53]

In the 'extended' model, we have the same tendency to a decrease in expected utility as we increase the revaluation risk by increasing the 'informativeness' of information. In this model, however, we see a counteracting increase in expected utility. As we identified in the previous case, this potential increase in expected utility is due to the ability to trade off current and future consumption by altering one's investment in the riskless technology.

[53] In theory, we would expect an insurance market to arise to allow risk-sharing of this revaluation risk. As noted in a previous footnote (section A.3), however, we cannot allow this insurance trading to take place after traders have received their non-tradable asset endowments. Such a round of insurance trading would reveal the average non-tradable endowment, h, to the uninformed traders. Consequently, at the next round of trading – after receipt of information by the informed traders – the risky asset price would be fully revealing. The only way to allow insurance without destroying the partially revealing price of the risky asset is to allow traders to purchase an insurance contract *before* endowments of the risky asset have been received. Furthermore, settling up on these insurance contracts could not take place until *after* the beginning of period trading in the risky asset had taken place.

By examining the actual equations,[54]

$$\frac{\partial}{\partial\rho^2}\left(\frac{\Omega}{1+R}-\frac{1}{2}\frac{\Psi'}{1+R}(I+2\frac{\Gamma\Gamma'}{1+R})^{-1}\frac{\Psi}{1+R}\right)$$
$$=-\frac{1}{2}\frac{(1+R)a^2\sigma_x^2\sigma_h^2}{((1+R)+a^2\sigma_x^2\sigma_h^2(1-\rho^2))^2}\frac{(\mu_h-h_i)^2}{\sigma_h^2}$$
$$+\frac{1}{2}\frac{R}{(1+R)^2}a^2\sigma_x^2(f+h_i)^2$$
$$<> 0,$$

we see that the positive term identified in the previous section as the benefit due to increased allocative efficiency is still present. In addition, we have a second, negative, term dependent on the deviation of one's endowment from the expected endowment. Notice that for σ_h^2 small, the negative term of the partial derivative above will likewise be small.[55] The positive term also changes, but in the limit as σ_h^2 approaches zero, the positive term dominates. This holds true even if we add in the negative effect from the following term.

$$\frac{\partial}{\partial\rho^2}\ln\left|I+2\frac{\Gamma\Gamma'}{1+R}\right|=\frac{\partial}{\partial\rho^2}\ln\left(1+\left(\frac{1}{1+R}\right)a^2\sigma_x^2\sigma_h^2(1-\rho^2)\right)$$
$$=-\frac{a^2\sigma_x^2\sigma_h^2}{(1+R)+a^2\sigma_x^2\sigma_h^2(1-\rho^2)}<0$$

This expression also vanishes as σ_h^2 approaches zero.

In conclusion, we can say that information will have value in the 'extended' model only if the variance of the average non-tradable asset endowment, σ_h^2, is not too large. This value arises from the ability to improve one's allocation of current and future consumption. As the endowment variance increases, however, the exposure to revaluation risk increases, decreasing expected utility. This may lead to a

[54] For this model, $t \sim N\{f+\mu_h,\sigma_h^2\}$ and $f_i=f$, so that $\mu_t-(f_i+h_i)=\mu_h-h_i$ and $\sigma_t=\sigma_h$.

[55] While varying σ_h^2, we must keep $(\mu_h-h_i)/\sigma_h$ constant.

decrease in expected utility as quality of information is increased.[56]

In the analysis above, two negative contributions from the partial derivatives were lumped into what I named revaluation risk. But one of those factors, namely, the partial derivative of the logarithmic expression, was also present in the previous efficient market model analysis. In that analysis, the explanation of the negative effect on utility was that the rational expectations traders derived a benefit from trading with the group of 'naive' traders,[57] and that this benefit decreased as information quality increased. There is no outside group of 'naive' traders in this model, but the logarithmic term and its negative partial derivative are still present.

Even though there is no outside group of 'naive' traders, the explanation of this term is similar. Simply stated, because every trader's *pre-trading* shadow price of the risky asset is different,[58] [59] every trader benefits from the opportunity to trade. Traders with shadow prices lower than the market price benefit by selling risky asset. Traders with shadow prices higher than the market price benefit

---

[56] Returning again to Table A.2, we have moved to the third column describing revaluation risk.

[57] This was named the dealer benefit.

[58] *After* trading, of course, the price of the risky asset and each trader's shadow price are equal.

[59] The efficient market model may also be interpreted in terms of shadow prices. When the 'naive' trader supply component is positive, we can imagine that the 'naive' trader shadow price of the risky asset is zero. That is, they wish to sell at any price. The price of the risky asset does not fall to zero because the 'naive' trader supply component is finite. That is, the 'naive' trader is never the marginal trader. When the 'naive' trader supply component is negative (ie. they want to buy risky asset), then we can imagine that their shadow price of the risky asset is infinitely positive. That is, they wish to buy at any price.

by buying risky asset. The differences in shadow prices, of course, arise from the different trader endowments of the non-tradable asset. That is, this benefit is result of having the opportunity to hedge one's position in the non-tradable asset, h_i. As this hedging benefit is the result of differences in pre-trade shadow prices, anything which diminishes these differences reduces the hedging benefit. As we saw in the efficient market model, however, as the quality of information increases, the risky asset becomes more and more riskless. At the same time, the differences between shadow prices diminish, until in the limit where $\rho = 1$ we have all the shadow prices exactly equal to the price of the riskless asset. Therefore, this factor identifies the future benefit to be derived from trading with a market of individuals having varying shadow prices for the risky asset. As information quality increases, the differences between shadow prices diminish, causing the hedging benefit to decrease.

The final term to consider is the negative term from the first partial derivative above. This term is the only element of expected utility which depends on the actual endowment, h_i, and represents the exposure to revaluation risk discussed above. One receives an endowment, but information will be revealed causing everyone to revalue the asset you wish to trade.

In summary, we have[60]

$$\frac{\partial}{\partial \rho^2}\overrightarrow{K^*_{iI}} = \frac{\partial}{\partial \rho^2}(\text{revaluation risk term}) + \frac{\partial}{\partial \rho^2}(\text{allocation efficiency term}) + \frac{\partial}{\partial \rho^2}(\text{hedging benefit term}),$$

where we have shown that exposure to revaluation risk increases as information quality increases, thus depressing expected utility.

$$\frac{\partial}{\partial \rho^2}(\text{revaluation risk term}) < 0$$

Also, allocative efficiency increases as information quality increases, thus increasing expected utility,

$$\frac{\partial}{\partial \rho^2}(\text{allocative efficiency term}) > 0,$$

and the hedging benefit decreases as information quality increases.[61]

$$\frac{\partial}{\partial \rho^2}(\text{hedging benefit term}) < 0$$

The net effect depends on the size of the endowment variance, σ_h^2, which in turn determines how much hedging will take place in the model. If the market is used mainly as a vehicle for hedging other assets (σ_h^2 large), then better quality information may not be desired by traders. The reason is that it exposes them to revaluation risk. If hedging is not a very important use of the market (σ_h^2 small),

[60] The analysis for the standard model is the same except that the allocative efficiency term is lacking. 'Pre-info' expected utility therefore unambiguously falls in the 'standard' model when information quality increases.

[61] This term is shown in Table A.2 with a positive or negative derivative for the informed trader. The positive sign is possible only in an asymmetric information model. It is not present in a homogeneous information case such as this.

then better quality information is desirable, as it leads to increased allocative efficiency.

Certainly this could have effects on the amount of information produced. If information produced immediately becomes public (this *is*, after all, a homogeneous information model), then our traders might not produce information to the point where the marginal cost of information equals the marginal benefit due to increased allocative efficiency. Obviously, if hedging is an important use of the market, then we have to also note the effects of information on hedging possibilities.

A.4.3 Case 3: Asymmetric Information, Randomness Present

In this last case we finally introduce the uninformed rational expectations trader. The result is that the expressions for 'pre-info' expected utility become very complicated. Luckily, some degree of simplification is possible and allows a few conclusions to be extracted. Unfortunately, a complete analysis is not possible.

As was done in case 2, the analysis will first be presented for the general model. The interpretation of the results will be given in the sections immediately following the general analysis. In addition, the 'standard' model will be fully treated before continuing on to the 'extended' model.

A.4.3.a STANDARD MODEL

As was shown in section A.3.4.a above, the 'pre-info' expected utility of an uninformed trader in the 'standard' model is

$$\overline{J}^*_{iU} = \mathrm{E}(J^*_{iU}) = -\exp\left(-\left[\omega - \ln\beta - \frac{1}{2}\frac{\psi^2}{1+2\gamma^2} + \frac{1}{2}\ln(1+2\gamma^2)\right]\right),$$

where

$$\gamma^2 = \frac{1}{2}\frac{(\phi\rho - Rp_1)^2}{\phi(1-\phi\rho^2)},$$

$$\psi = \frac{\sigma_x}{\sqrt{\phi}}\left(\frac{(\mu_x - Rp_0)(\phi\rho - Rp_1)}{\sigma_x^2(1-\phi\rho^2)} + aRp_1(f_i + h_i)\right),$$

$$\omega = \frac{1}{2}\frac{(\mu_x - Rp_0)^2}{\sigma_x^2(1-\phi\rho^2)} + aR[m_i + (f_i + h_i)p_0].$$

It is shown in appendix 1 that[62]

$$\frac{\partial}{\partial\rho^2}\left(\omega - \frac{1}{2}\frac{\psi^2}{1+2\gamma^2}\right) \begin{cases} = 0, & \text{efficient market model, } \lambda > 0, \\ < 0, & \text{hedging model, } \lambda > 0, \end{cases}$$

and

$$\frac{\partial}{\partial\rho^2}\ln(1+2\gamma^2) \le 0, \quad \lambda > 0.$$

[62] The second derivative below is negative in both models except for the case $\rho = 1$, at which point it is zero.

The conclusion we reach is that the 'pre-info' expected utility of the uninformed trader in the 'standard' model unambiguously decreases as we increase the quality of information.[63] [64]

$$\frac{\partial}{\partial \rho^2}\overline{J}^*_{iU} \leq 0$$

What about the 'pre-info' expected utility of an informed trader in the 'standard' model? To handle this question we can borrow a result from Grossman and Stiglitz (1980, pp. 406-407), namely,[65]

$$\frac{\overline{J}^*_{iI}}{\overline{J}^*_{iU}} = \sqrt{\frac{1-\rho^2}{1-\phi\rho^2}},$$

which gives us

$$\overline{J}^*_{iI} = -\exp\left(-\left[\omega - \ln\beta - \frac{1}{2}\frac{\psi^2}{1+2\gamma^2} + \frac{1}{2}\ln\left(\left(\frac{1-\phi\rho^2}{1-\rho^2}\right)(1+2\gamma^2)\right)\right]\right).$$

If the added logarithmic term also has a negative partial derivative, then the 'pre-info' expected utility of both informed and uninformed traders decreases as quality of information increases.

This is, however, not the case. It can be shown that the partial derivative of this term may be positive or negative (see appendix 1).

$$\frac{\partial}{\partial \rho^2}\left(\frac{1-\phi\rho^2}{1-\rho^2}\right) <> 0$$

As a result, the partial derivative of the informed trader's expected utility is, likewise, either positive or negative.

$$\frac{\partial}{\partial \rho^2}\overline{J}^*_{iI} <> 0$$

---

[63] The interpretation of this result is presented in the following sections.

[64] Note that the derivative below is zero *only* in the efficient market model at the point where $\rho = 1$. Otherwise it is negative.

[65] This esult can be shown to also hold in the general model.

A.4.3.a.i Efficient Market Model

The previous section is quite definite about the effect of better quality information on the 'pre-info' expected utility of uninformed traders in this asymmetric information, 'standard' model. Their expected utility is unambiguously decreased when the informed traders receive better quality information. This is not surprising, as we saw in case 2 that even when every trader is informed the result is still a decrease in expected utility as the quality of information increases.

In fact, if we subtract the utility function terms derived for the homogeneous information case from the terms from the current asymmetric information case (see appendix 2), we can sign the differences as follows:[66]

$$\left(\omega - \frac{1}{2}\frac{\psi^2}{1+2\gamma^2}\right) - \left(\Omega - \frac{1}{2}\Psi'(I+2\Gamma\Gamma')^{-1}\Psi\right) = 0,$$

$$\ln(1+2\gamma^2) - \ln|I+2\Gamma\Gamma'| \leq 0.$$

As a result, the uninformed trader is unambiguously worse off here than he would be in the homogeneous information case.[67] The reason for the decrease in utility is not that the uninformed have poorer quality information than they would in the homogeneous information case. After all, in the 'standard' model better quality information *decreases* expected utility.

The reason for the decrease in utility lies in a decrease in the dealer benefit that our rational expectations traders receive from trading with the outside group of 'naive'

[66] The second difference below is negative for both the efficient market model and the hedging model, except when $\rho = 0$ or $\rho = 1$, at which points it is zero.

[67] Except, of course, when $\rho = 0$ or $\rho = 1$, at which points he is just as well off.

investors.[68] Note that since this benefit is the only reason that the uninformed traders are trading at all, it is not possible for all of the benefit to disappear. If it did, then the uninformed traders would presumably not trade, there now being no reason for them to do so. Since this situation does not arise, we are assured that the benefit to trading never completely disappears.

The dealer benefit decreases because the uninformed traders are not able to distinguish perfectly between above average demand due to 'naive' traders versus above average demand on the part of the informed traders. When they sell to the 'naive' investors they end up better off than if they hadn't; when they sell to the informed traders they end up worse off than if they hadn't. On the whole, however, they end up better off than if they stopped trading altogether. The benefits that they give away to the informed are less than the benefits that they receive from trading with the 'naive' traders.

What about the 'pre-info' expected utility of the informed traders in this asymmetric information model? We know at least one thing, namely, that at the two end points, $\rho = 0$ and $\rho = 1$, the expected utility of the informed and uninformed converge at the values found in the homogeneous information case. Since the uninformed trader expected utility lies below what it did in the homogeneous information case, we would expect the utility of the informed traders to lie above what we found in the homogeneous information case.

---

[68] Recall from previous sections that the logarithmic term in the expected utility function was identified with the dealer benefit.

This turns out to be the case, as can be verified by examining the terms of the informed traders' 'pre-info' expected utility in this case. If we once again subtract the utility terms derived for the homogeneous information case from the terms of this asymmetric information case (see appendix 2), we can again sign the differences.[69]

$$\left(\omega - \frac{1}{2}\frac{\psi^2}{1+2\gamma^2}\right) - \left(\Omega - \frac{1}{2}\Psi'(I+2\Gamma\Gamma')^{-1}\Psi\right) = 0,$$

$$\ln\left(\left(\frac{1-\phi\rho^2}{1-\rho^2}\right)(1+2\gamma^2)\right) - \ln|I+2\Gamma\Gamma'| \geq 0.$$

As a result, the informed trader is unambiguously better off here than he would be in the homogeneous information case.[70]

Further, we can show that the partial derivative of the logarithmic term is guaranteed to be negative if the quality of information is raised high enough.[71]

$$\frac{\partial}{\partial\rho^2}\left(\frac{1-\phi\rho^2}{1-\rho^2}\right) < 0, \qquad \text{for} \quad \frac{a\sigma_x\sigma_s}{a\sigma_x\sigma_s+\lambda} < \rho^2 < 1$$

This is consistent with a decrease of expected utility to the level of the homogeneous information model as ρ approaches one.

The informed trader 'pre-info' expected utility is higher in this efficient market model than it is in the homogeneous information case because the informed traders are able to capture more of the dealer benefit here than they could when everyone was informed. This is due to the difficulty the uninformed have in distinguishing

69 The first difference is identical to the one for the uninformed trader above. The second difference is positive for both the efficient market model and the hedging model, except when $\rho = 0$ or $\rho = 1$, at which points it is zero.

70 Except, of course, when $\rho = 0$ or $\rho = 1$, at which points he is just as well off.

71 Substitution has been made for $\sigma_t = \sigma_s$.

between out of the normal demand due to 'naive' investors versus informed traders. There are, therefore, two effects occurring simultaneously. As ρ increases from zero, the portion of the dealer benefit captured by the informed traders increases.[72] At the same time, the total dealer benefit decreases.

Certainly the informed traders are better off than they would be if everyone were informed, but are they better off than they would be if everyone were *uninformed*? In the homogeneous information case we saw that better quality information caused a drop in utility, so that if traders had the option of forming an enforceable cartel, they would do so. The cartel would in effect be an agreement not to use the signal, ϵ, thus moving everyone to the higher expected utility point where $\rho = 0$. Would the informed traders also want to form such a cartel in this asymmetric information case?

This question can be answered by looking at the partial derivative of the informed trader's expected utility function at the point $\rho = 0$. If it is possible for it to be positive, then it is possible for the informed traders to reach a higher expected utility than they would have if everyone was uninformed. (In the 'standard' model with homogeneous information this is the highest utility level possible.)

Using the expressions derived in appendix 1, it can be shown that the partial derivative of the 'pre-info' expected utility of an informed trader in this asymmet-

---

[72] This occurs as ρ increases from zero. When ρ approaches one the opposite occurs: the portion of the benefit captured by informed and uninformed traders once again equalizes.

ric information model may be positive or negative at $\rho = 0$. The derivatives of the terms contained in the informed trader's utility function are[73]

$$\frac{\partial}{\partial \rho^2}\left(\omega - \frac{1}{2}\frac{\psi^2}{1+2\gamma^2}\right) = 0,$$

$$\frac{\partial}{\partial \rho^2}\left(\frac{1}{2}\ln(1+2\gamma^2)\right) = -\frac{1}{2}\frac{\lambda(\lambda + 2a\sigma_x^2\sigma_s^2)}{1+a\sigma_x^2\sigma_s^2}, \qquad \text{at } \rho = 0,$$

$$\frac{\partial}{\partial \rho^2}\left(\frac{1}{2}\ln\left(\frac{1-\phi\rho^2}{1-\rho^2}\right)\right) = \frac{1}{2}, \qquad \text{at } \rho = 0.$$

As λ approaches zero, the second term also approaches zero. As σ_s approaches zero, it approaches $-\lambda^2/2$, which is greater than $-1/2$.

The last term, however, is $1/2$ when ρ is zero. Both of these terms are due to the presence of the dealer benefit. The negative term is due to the shrinkage of the total benefit as information quality increases, while the positive term reflects the fact that the informed trader group initially captures a greater portion of the benefit.[74] We see, therefore, that if the benefits to trading with the outside group of 'naive' traders are too great, or there are too many informed traders, then the loss of these benefits due to increased information quality will outweigh the gain in benefits due to being one of the informed. If, however, λ and σ_s are sufficiently small, then the gain of benefits from being one of the informed traders dominates.

In case 2, we saw that when all traders were informed, then increasing the quality

[73] Since this is the efficient market model, substitution has been made for $\mu_t - (f_i + h_i) = 0$ and $\sigma_t = \sigma_s$.

[74] Returning to Table A.1, we have here the explanation for the '+/–' entries in the dealer benefit column for the informed trader. Naturally, in a homogeneous information model the portion of the benefit captured by the informed traders cannot increase – all traders are informed – so the positive term is lacking. That is, in the homogeneous information case the derivative becomes unambiguously negative.

of information caused a decrease in expected utility. Here we see that when *not* all traders are informed, those that are informed may receive an increase in expected utility when information quality increases. Naturally, as we increase the proportion of informed traders, λ, we expect to eventually see this possibility disappear.[75]

In fact, we can state that when up to half of the traders are informed, then the informed traders will experience a net increase in expected utility when information quality is increased. This can be seen by adding together the two partial derivatives above.

$$\frac{\partial}{\partial \rho^2}\left(\frac{1}{2}\ln\left(\frac{1-\phi\rho^2}{1-\rho^2}\right)(1+2\gamma^2)\right) = \frac{1}{2}\,\frac{1+a^2\sigma_x^2\sigma_s^2 - 2(a^2\sigma_x^2\sigma_s^2)\lambda - \lambda^2}{1+a^2\sigma_x^2\sigma_s^2}$$
$$> 0, \quad \text{for } \lambda < \tfrac{1}{2}, \text{ at } \rho = 0$$

Therefore, we see that the informed trader in the asymmetric information model may not only be better off than he would be if he were uninformed, he may be better off than he would be if everyone, including himself, were uninformed. This, of course, is not a situation conducive to the stability of a cartel. As we saw, the negative contribution to the partial derivative of the informed trader expected utility depended on the proportion of traders informed. When that proportion decreases to zero as a result of a cartel being set up, this term disappears.

As a result, only the positive contribution to the partial derivative is present when a cartel has been created. There is always a great incentive to be the *only* informed

---

75 As λ approaches 1, the model approaches the homogeneous information case. In the homogeneous information case we know that better quality information reduces informed traders' expected utility.

trader and reap huge benefits. Because all the potentially informed traders have this incentive, we expect the classic solution to such a prisoners' dilemna, namely, that all cartel participants cheat, causing the cartel to fail.

A.4.3.a.ii Hedging Model

The analysis for the hedging model parallels that of the efficient market model in the previous section. The difference is that in this model there is no outside group of 'naive' traders. Instead, each trader receives an endowment, h_i, which he hedges by trading in the risky asset market. As discussed in section A.4.2.b, this leads to a hedging benefit from trading which is analogous to the benefit from trading with a group of 'naive' traders.

In addition, however, we introduce the risk of endowment revaluation into the model.[76] The expression describing the change in revaluation risk as information quality is changed is the following partial differential (see appendix 1).

$$\frac{\partial}{\partial \rho^2}\left(\omega - \frac{1}{2}\frac{\psi^2}{1+2\gamma^2}\right) < 0, \quad \text{for } \lambda > 0$$

The important points to note about this term are that it is negative and that it approaches zero as λ or σ_h approach zero. This means that all of the conclusions

[76] See section A.4.2.b.

regarding the uninformed trader which were made in the previous section on the efficient market model, also hold in this model. His expected utility decreases as information quality increases, and always lies below the level we would find if every trader were informed.

The conclusions regarding the informed trader are also not changed. If hedging is an unimportant use of the market (σ_h small), or there are few informed traders (λ small), then the benefit of receiving a greater *portion* of the hedging benefits due to better quality information outweighs the losses of expected utility due to increased revaluation risk and decreased *total* hedging benefits. That is, we once again find a situation where the benefit of being the only informed trader is large, thus ruling out the possibility of a cartel.

If hedging is an important use of the market (σ_h large), or there are a large number of traders informed (λ large), then receiving a larger portion of the hedging benefit does not compensate for the losses resulting from an increase in information quality. A cartel would still not be successful, however, since imposing a cartel immediately sets λ equal to zero. As we have seen when λ is zero, the negative terms in the partial derivative of expected utility (such as the term above) disappear. This leaves just the positive attraction of receiving a larger portion of the hedging benefit, thus creating a large incentive for all cartel members to cheat and become informed.

A.4.3.b EXTENDED MODEL

In the previous sections we have been able to do a reasonably thorough analysis of the 'pre-info' expected utility functions for informed and uninformed traders. Once we attempt to pass to the 'extended' model, however, the equations become relatively intractable. As a result, most of the analysis to follow will concentrate on the efficient market model. Because $\mu_t - (f_i + h_i) = 0$ in this model, some simplification of the equations is possible.

A.4.3.b.i Efficient Market Model

The buildup of models looked at up to this point gives us confidence in predicting what to expect when extending the 'standard' model. To explain the dynamics of the efficient market, 'standard' model, we only needed to use two concepts, namely, the loss of *total* dealer benefit as information quality increased, and the unequal division of dealer benefit, with informed traders receiving more of this benefit than uninformed traders.[77]

We also saw that as the proportion of informed traders decreased, the discrepancy

[77] As ρ increases from 0, the division of the dealer benefit becomes more unequal. As ρ increases to 1, the division once more equalizes.

between the amount of dealer benefit received by informed versus uninformed traders increased. This lead to instability of any cartel that might be proposed for informed traders.

What difference do the previous models predict when we extend the 'standard' model? From what we have seen, the only effect on the efficient market model is that a new benefit appears. This new benefit arises from the possibility of trading off current consumption against future consumption by trading in the riskless technology.

Unlike the 'standard' model, therefore, we should not expect to find an unambiguously negative partial derivative for the uninformed trader expected utility as we vary ρ^2. The partial derivative should have the possibility of being positive, to allow for the fact that counteracting the decrease in dealer benefit lost as information quality is increased, we have more efficient allocation of consumption.

We can verify this prediction, even though the proof is long (see appendix 3). The prediction, in fact, makes sense intuitively, since we would expect the positive benefit of improved consumption allocation to show up most strongly when the 'informativeness' of the risky asset price, ie. ϕ, is highest. Since the price becomes fully revealing when $\rho = 1$, it makes sense to find these benefits showing up as we approach a fully revealing price.

Naturally, since we previously showed that the homogeneous information case expected utility was higher in the 'extended' model than in the 'standard' model, we know that at $\rho = 1$ the uninformed traders must have higher expected utility in this asymmetric information, 'extended' model than they have in the asymmetric information, 'standard' model.[78] This by itself, however, will not guarantee that we can find a situation where uninformed trader expected utility increases as information quality increases.

In order to have the partial derivative above *guaranteed* positive at $\rho = 1$, we need to have the losses of dealer benefit caused by increased information quality tailing off to *zero* as information quality increases to 1. This is, in fact, the situation we have, as is shown below (see appendix 3).[79]

$$\frac{\partial}{\partial \rho^2} \ln(1 + R + 2\gamma^2) = \frac{2}{1 + R + 2\gamma^2} \frac{\partial \gamma^2}{\partial \rho^2} = 0, \quad \text{at } \rho = 1$$

Therefore, at the same time as the loss of dealer benefit decreases to zero, the benefit due to better consumption allocation is in its region of greatest increase, resulting in a guaranteed upswing in uninformed trader expected utility as information quality increases to 1.

Unfortunately, the expressions for the 'pre-info' expected utility of the informed trader in this asymmetric information, 'extended' model are not analytically tractable, so no further analysis can be presented here. We would expect, however, that an informed trader in the 'extended' model would have higher expected

[78] This is because the risky asset price is fully revealing at $\rho = 1$, thus creating a homogeneous information situation.

[79] Recall that the logarithmic term has been associated with the dealer benefit in previous sections.

utility than the corresponding informed trader in the 'standard' model, due to the added benefit from better allocation of consumption.

The reason for expecting this is the fact that the consumption and investment decisions were shown in section A.3.3 to be independent of one another given negative exponential utility. Changing the consumption allocation opportunities should, therefore, not affect the investment decision. Since we are improving the consumption opportunities only, we should expect an increase in utility. This should hold for both the informed and uninformed traders in this model.

A.4.3.b.ii Hedging Model

Unfortunately, this model which is potentially the most interesting is also the most difficult. The expressions for informed and uninformed trader 'pre-info' expected utility are unmanageable, and leave us no choice but to speculate using the concepts built up in previous sections.

As we have seen, the dynamics of this model depend on four factors: the decrease of hedging benefits as information quality increases, increase of revaluation risk as information quality increases, increased efficiency of consumption allocation as information quality increases, and the fact that the hedging benefit realized by an informed trader is higher than the hedging benefit realized by an uninformed trader.

When we pass from the efficient market model to the hedging model, the only qualitative change that is expected is due to the addition of revaluation risk.[80] We would expect both informed and uninformed trader expected utility to be lower in the hedging model than in the efficient market model, with the divergence between the two models increasing as information quality increases.[81] Because of the multiplicity of effects, further analysis is not possible.

---

[80] The effects due to the hedging benefit are analogous to the effects due to the dealer benefit in the efficient market model.

[81] Revaluation risk increases as information quality increases.

A.5 SUMMARY AND CONCLUSIONS

This study examines in detail several asymmetric information, rational expectations models similar to the Grossman and Stiglitz (1980) model. It provides a detailed discussion of the welfare effects following an increase in the quality of information given to the informed trader group in these models.

In case 1 – homogeneous information, non-random supply – an allocative efficiency effect was identified in the 'extended' model but not in the 'standard' model. In the 'extended' model, when traders received better quality information, they were able to allocate their wealth more efficiently between current and future consumption, thereby increasing their welfare. The vehicle allowing this trade-off was shown to be the riskless technology. It was also pointed out that replacing the riskless technology with a fixed supply of riskless asset would prevent such an allocative efficiency effect and make the 'standard' and 'extended' models qualitatively equivalent.

In case 2 – homogeneous information, random supply – a dealer benefit was identified in the efficient market model. The price of the risky asset was shown to fall when supply rose and rise when supply fell, something that the rational expectations 'dealers' were aware of, but which was unknown to or unobservable by the 'naive' efficient market traders. Because they were able to observe unusually high demand when it occured, and satisfy that demand from their own stock of risky asset (and, conversely, absorb abnormally high supply of the risky asset), the rational expectations traders made a profit by trading with the 'naive' traders.

In the 'extended' model, as the quality of information given to traders increases, the size of this dealer benefit decreases, resulting in an ambiguous net effect on welfare. The net effect was shown to depend on the size of the supply variance due to 'naive' traders, σ_s^2. If the dealer benefit is small (σ_s^2 small), then the loss of benefit is small as information quality increases, and may be outweighed by the increase in allocative efficiency benefits, leading to a net increase in welfare. For a large dealer benefit (σ_s^2 large), however, the loss of benefit is large and may outweigh the increased allocative efficiency effects, leading to a net decrease in welfare.

In the hedging model of case 2, a benefit analogous to the dealer benefit of the efficient market model was identified. This benefit is the result of differences between traders' pre-trade shadow prices for the risky asset (due to different endowments of the non-tradable asset). Because trading with a person who has a different shadow price for the risky asset provides both you and the other person with an increase in welfare, this hedging benefit accrues to all parties. Increasing the quality of information given to traders, however, diminishes the differences between pre-trade shadow prices and results in a decrease in this hedging benefit.

In addition to the hedging benefit, another factor was identified in the hedging model. Because the endowments held by rational expectations traders differ, trading between rational expectations traders will take place. This is unlike the situation in the efficient market model, where rational expectations traders trade only with the 'naive' trader group, not between themselves. Each trader is thus exposed to the risk that his endowment will be revalued when information is disseminated to the other rational expectations traders. As the quality of information given out increases, the extremes of revaluation become more probable, thereby decreasing expected utility (by Jensen's inequality).

The net effect on expected utility in the hedging model following an increase in information quality depends on the size of the endowment variance, σ_h^2, which in turn determines the extent of hedge based trading which occurs in the model. If the market is used extensively for hedging (σ_h^2 large), then even though better quality information results in an allocative efficiency benefit, it will also result in a large decrease in hedge-based usage of the market (ie. a decrease in hedging benefit and increase in revaluation risk). This effect may outweigh the allocative efficiency increase, leading to a net drop in welfare.[82]

In case 3, the uninformed rational expectations trader is introduced, resulting in an asymmetric information model. Tables A.1 and A.2 summarize the effects of an increase of information quality on the different factors outlined above, and the

[82] For example, if market prices from one sector of an economy were used as indicators by another sector of the economy, the government might feel that it would be to the common good to collect and disseminate information. Traders in the different markets would have access to this information and would trade on the basis of it, thus producing prices which reflected a greater amount of information than would otherwise be the case. We can see from the hedging market analysis that traders in some markets could be against such a scheme. Certainly, the result would be prices which convey better information, thereby increasing investment efficiency in the economy, but at the same time hedging opportunities would be decreased.

net effect on expected utility. In the 'standard', efficient market model, it was shown that the uninformed trader is unambiguously worse off than he would be in a homogeneous information scenario. The informed investor was shown to be unambiguously better off. The reason for this is that the dealer benefit – which is shared equally in a homogeneous information situation – is unequally distributed in this asymmetric information case, due to the difficulty which uninformed traders have in distinguishing between abnormal demand based on 'naive' trader activity and demand due to receipt of good information by informed traders.

It was also shown in the 'standard' model that it is possible to find informed traders enjoying a higher level of utility than they would have if everyone was uninformed (which is the highest utility level possible in a homogeneous information case). This was shown to be possible even when up to half of the traders in the model belonged to the informed group. As the size of the informed trader group decreases, the benefits of being one of the remaining informed traders increases until, at the limit where no trader is informed, we find that becoming the *only* informed trader is guaranteed to result in an increase in utility. This finding rules out the possibility of the informed traders voluntarily deciding not to receive information.[83]

These arguments can be applied to the 'extended', efficient market model by using the fact that the investment and consumption decisions in these models

[83] This argument is only valid if the information to be received is not of perfect quality. Given perfect information, the price of the risky asset becomes perfectly revealing regardless of the fact that only one trader is informed. Given less than perfect information, since there are an infinite number of traders in the model this problem does not arise.

are independent of one another. Since the 'extended' model adds a consumption opportunity, while leaving the investment opportunities of the 'standard' model unchanged, we would expect only that the utility curves for both informed and uninformed traders would lie above their counterparts in the 'standard' model, with the separation between these pairs of curves increasing as information quality (and thus allocative efficiency) increases.

Turning finally to the hedging model, although we cannot analytically verify that the conclusions reached in the efficient market model also hold here, extension of these conclusions to the hedging model seems intuitively correct. The only aspect which must be taken into account when passing from the efficient market model to the hedging model is the factor of revaluation risk.[84] When addressing the question of whether or not becoming the only informed trader would result in an increase in utility, we must look to the effects of revaluation risk on the conclusions reached above.

Given that there are an infinite number of traders in the model, we can assume that the informativeness of the price system, ϕ, is infinitesimal when only one trader is informed. This, of course, is the reason that being the only informed trader in the efficient market model guaranteed an increase in utility. Since revaluation risk only occurs in situations where other traders are able to obtain information about the risky asset, we see that revaluation risk will also be infinitesimal when only one trader is informed. The reason is that the only source of information

[84] The hedging benefit of the hedging model is analogous to the dealer benefit of the efficient market model.

that is available to the uninformed traders is the risky asset price, which is only infinitesimally revealing.

A.6 REFERENCES

Epstein, L.G. and S.M. Turnbull, 1980, Capital asset prices and the temporal resolution of uncertainty, Journal of Finance 35, No. 3, 627-643.

Grossman, S.J. and J.E. Stiglitz, 1980, On the impossibility of informationally efficient markets, American Economic Review 70, No. 3, 393-408.

Hakansson, N.H., J.G. Kunkel and J.A. Ohlson, 1982, Sufficient and necessary conditions for information to have social value in pure exchange, Journal of Finance 37, No. 5, 1169-1181.

Hellwig, M.F., 1980, On the aggregation of information in competitive markets, Journal of Economic Theory 22, 477-498.

Hirshleifer, J., 1971, The private and social value of information and the reward to inventive activity, American Economic Review 61, 561-574.

Kraus, A. and G.A. Sick, 1979, Communication of aggregate preferences through market prices, Journal of Financial and Quantitative Analysis 14 (Proceedings), No. 4, 695-703.

Verrecchia, R.E., 1982, Information acquisition in a noisy rational expectations economy, Econometrica 50, No. 6, 1415-1430.

A.7 APPENDICES

A.7.1 Appendix 1. Case 3, Standard Model: Derivatives

In this appendix several partial derivatives are calculated for use in sections A.4.3.a, A.4.3.a.i and A.4.3.a.ii. First, we consider terms from the 'pre-info' expected utility function of the uninformed trader in the 'standard' model. Following this, the informed trader utility function is considered.

The expressions in section A.4.3.a do not simplify very easily. We can, however, perform some simplification by introducing a new variable, ξ,

$$\xi = \phi\theta^2 \frac{\sigma_x^2}{\sigma_t^2} = 1 - \phi.$$

Using this new variable, we find that

$$\phi\rho - Rp_1 = -\frac{\lambda\rho\xi(1-\phi\rho^2)}{1-\rho^2+\lambda\rho^2\xi},$$

$$Rp_0 = \mu_x - \frac{a\sigma_x^2(1-\rho^2)(1-\phi\rho^2)}{1-\rho^2+\lambda\rho^2\xi}\,\mu_t,$$

$$aRp_1(f_i+h_i) = \frac{a\rho(\phi(1-\rho^2)+\lambda\xi)}{1-\rho^2+\lambda\rho^2\xi}\,(f_i+h_i).$$

Substituting these expressions into those for γ^2, ψ and ω, we find

$$\gamma^2 = \frac{1}{2}\frac{\lambda^2\rho^2\xi^2(1-\phi\rho^2)}{\phi(1-\rho^2+\lambda\rho^2\xi)^2},$$

$$\psi = \frac{\sigma_x}{\sqrt{\phi}}\Bigg(-\frac{a\lambda\rho\xi(1-\rho^2)(1-\phi\rho^2)}{(1-\rho^2+\lambda\rho^2\xi)^2}\mu_t + \frac{a\rho(\phi(1-\rho^2)+\lambda\xi)}{1-\rho^2+\lambda\rho^2\xi}(f_i+h_i)\Bigg),$$

$$\omega = \widehat{\omega} + aRm_i + a\mu_t(f_i+h_i),$$

$$\widehat{\omega} = -\frac{a^2\sigma_x^2(1-\rho^2)(1-\phi\rho^2)}{1-\rho^2+\lambda\rho^2\xi}\mu_t(f_i+h_i) + \frac{1}{2}\frac{a^2\sigma_x^2(1-\rho^2)^2(1-\phi\rho^2)}{(1-\rho^2+\lambda\rho^2\xi)^2}\mu_t^2.$$

Using these expressions, it can be shown that

$$1+2\gamma^2 = \frac{\phi(1-\rho^2+\lambda\rho^2\xi)^2+\lambda^2\rho^2\xi^2(1-\phi\rho^2)}{\phi(1-\rho^2+\lambda\rho^2\xi)^2} = \frac{\phi(1-\rho^2)^2+2\phi(1-\rho^2)\lambda\rho^2\xi+\lambda^2\rho^2\xi^2}{\phi(1-\rho^2+\lambda\rho^2\xi)^2},$$

and

$$\widehat{\omega}(1+2\gamma^2) - \frac{1}{2}\psi^2 = \frac{1}{2}\frac{a^2\sigma_x^2(1-\rho^2)^2(1-\phi\rho^2)}{(1-\rho^2+\lambda\rho^2\xi)^2}[\mu_t-(f_i+h_i)]^2 - \frac{1}{2}\frac{a^2\sigma_x^2}{\phi(1-\rho^2+\lambda\rho^2\xi)^2}[\phi(1-\rho^2)^2+2\phi(1-\rho^2)\lambda\rho^2\xi+\lambda^2\rho^2\xi^2](f_i+h_i)^2,$$

so that

$$\widehat{\omega} - \frac{1}{2}\frac{\psi^2}{1+2\gamma^2} = \frac{1}{2}a^2\sigma_x^2\left(\frac{\phi(1-\rho^2)^2(1-\phi\rho^2)}{\phi(1-\rho^2)^2+2\phi(1-\rho^2)\lambda\rho^2\xi+\lambda^2\rho^2\xi^2}[\mu_t-(f_i+h_i)]^2 - (f_i+h_i)^2\right),$$

and

$$\omega - \frac{1}{2}\frac{\psi^2}{1+2\gamma^2} = aRm_i + a\mu_t(f_i+h_i) - \frac{1}{2}a^2\sigma_x^2(f_i+h_i)^2$$
$$+\frac{1}{2}\frac{a^2\sigma_x^2\sigma_t^2\phi(1-\rho^2)^2(1-\phi\rho^2)}{\phi(1-\rho^2)^2+2\phi(1-\rho^2)\lambda\rho^2\xi+\lambda^2\rho^2\xi^2}\frac{[\mu_t-(f_i+h_i)]^2}{\sigma_t^2}.$$

At this point we would like to take the partial derivative of the above expression with respect to ρ^2. First, however, we define yet another variable, α,

$$\alpha = a^2\sigma_x^2\sigma_t^2(1-\rho^2),$$

which gives us

$$\phi = \frac{\lambda^2\rho^2}{\alpha(1-\rho^2)+\lambda^2\rho^2}, \qquad \xi = \frac{\alpha(1-\rho^2)}{\alpha(1-\rho^2)+\lambda^2\rho^2},$$

and

$$\phi(1-\rho^2)^2(1-\phi\rho^2) = \frac{\lambda^2\rho^2(1-\rho^2)^3(\alpha+\lambda^2\rho^2)}{[\alpha(1-\rho^2)+\lambda^2\rho^2]^2},$$

$$\phi(1-\rho^2)^2+2\phi(1-\rho^2)\lambda\rho^2\xi+\lambda^2\rho^2\xi^2$$
$$= \frac{\lambda^2\rho^2(1-\rho^2)^2[(1-\rho^2)(\alpha+\lambda^2\rho^2)+(\alpha+\lambda\rho^2)^2]}{[\alpha(1-\rho^2)+\lambda^2\rho^2]^2},$$

giving

$$\frac{\phi(1-\rho^2)^2(1-\phi\rho^2)}{\phi(1-\rho^2)^2+2\phi(1-\rho^2)\lambda\rho^2\xi+\lambda^2\rho^2\xi^2}$$
$$= \frac{(1-\rho^2)(\alpha+\lambda^2\rho^2)}{(1-\rho^2)(\alpha+\lambda^2\rho^2)+(\alpha+\lambda\rho^2)^2}.$$

It can be shown tat the partial derivative of the expression above is

$$\frac{\partial}{\partial\rho^2}\frac{\phi(1-\rho^2)^2(1-\phi\rho^2)}{\phi(1-\rho^2)^2+2\phi(1-\rho^2)\lambda\rho^2\xi+\lambda^2\rho^2\xi^2}$$
$$= -\frac{\lambda(\alpha+\lambda\rho^2)[(\alpha+\lambda^2\rho^2)+(1-\lambda)\alpha]}{[(1-\rho^2)(\alpha+\lambda^2\rho^2)+(\alpha+\lambda\rho^2)]^2}$$
$$< 0, \quad \text{for } \lambda > 0,$$

which tells us that

$$\frac{\partial}{\partial \rho^2}\left(\omega - \frac{1}{2}\frac{\psi^2}{1+2\gamma^2}\right) \quad \begin{cases} = 0, & \text{efficient market model, } \lambda > 0, \\ < 0, & \text{hedging model, } \lambda > 0. \end{cases}$$

Similarly, defining yet another variable, δ,

$$\delta = \alpha(1 - \rho^2 + \lambda\rho^2) + \lambda^2\rho^2,$$

we can express γ^2 as

$$\gamma^2 = \frac{1}{2}\frac{(1-\rho^2)\alpha^2(\alpha + \lambda^2\rho^2)}{\delta^2}.$$

The partial derivative of this with respect to ρ^2 can be shown to be[85]

$$\begin{aligned} \frac{\partial}{\partial \rho^2}\gamma^2 &= -\frac{1}{2}\frac{\alpha^2\lambda}{\delta^3}[2\alpha^2 + \alpha\lambda(1 + 3\rho^2 + \lambda\rho^2) + \lambda^3\rho^2(1+2\rho^2)] \\ &\leq 0, \quad \text{for } \lambda > 0. \end{aligned}$$

This tells us that

$$\frac{\partial}{\partial \rho^2}\ln(1 + 2\gamma^2) \leq 0.$$

Turning now to the utility of the informed trader, we see from section A.4.3.a that the informed trader 'pre-info' utility for the 'standard' model differs from that of the uninformed trader by the addition of the term

$$\ln\left(\frac{1-\phi\rho^2}{1-\rho^2}\right).$$

If we make the necessary substitutions, we find that

$$\frac{1-\phi\rho^2}{1-\rho^2} = \frac{\alpha + \lambda^2\rho^2}{\alpha(1-\rho^2) + \lambda^2\rho^2},$$

[85] The derivative below is negative except for the case $\rho = 1$, at which point it is zero.

which has a partial derivative which may be either positive or negative.

$$\frac{\partial}{\partial\rho^2}\left(\frac{1-\phi\rho^2}{1-\rho^2}\right)=\frac{\alpha[\alpha(1-\rho^2)-\lambda^2\rho^4]}{(1-\rho^2)[\alpha(1-\rho^2)+\lambda^2\rho^2]^2}<>0$$

A.7.2 Appendix 2. Case 3, Standard Model: Differences

In this appendix, terms from the 'pre-info' expected utility function in the 'standard' model, homogeneous information case (case 2) are subtracted from the corresponding terms in the asymmetric information utility functions. The results are used in section A.4.3.a.i.

From appendix 1, the relevant parts of the uninformed trader expected utility in this asymmetric information case are:

$$\begin{aligned}\omega-\frac{1}{2}\frac{\psi^2}{1+2\gamma^2}&=aRm_i+a\mu_t(f_i+h_i)-\frac{1}{2}a^2\sigma_x^2(_i+h_i)^2\\&\quad+\frac{1}{2}\frac{\alpha(\alpha+\lambda^2\rho^2)}{(\alpha+\lambda^2\rho^2)(1-\rho^2)+(\alpha+\lambda\rho^2)^2}\frac{[\mu_t-(f_i+h_i)]^2}{\sigma_t^2},\\ \ln(1+2\gamma^2)&=\ln\left(1+\frac{\alpha^2(1-\rho^2)(\alpha+\lambda^2\rho^2)}{\delta^2}\right).\end{aligned}$$

The corresponding terms from the homogeneous information case of section A.4.2

are[86]

$$
\begin{aligned}
\Omega - \frac{1}{2}\Psi'(I+2\Gamma\Gamma')^{-1}\Psi & \\
&= aRm_i + a\mu_x(f_i+h_i) - \frac{1}{2}a^2\sigma_x^2(f_i+h_i)^2 \\
&\quad + \frac{1}{2}\left(\frac{\alpha}{1+\alpha}\right)\frac{[\mu_t-(f_i+h_i)]^2}{\sigma_t^2},
\end{aligned}
$$

$$
\ln|I+2\Gamma\Gamma'| = \ln(1+\alpha).
$$

If we subtract the terms derived for the homogeneous information case from the terms of the current asymmetric case, we find the differences below[87]

$$
\begin{aligned}
&\frac{1}{2}\left(\frac{\alpha(\alpha+\lambda^2\rho^2)}{(\alpha+\lambda^2\rho^2)(1-\rho^2)+(\alpha+\lambda\rho^2)^2} - \frac{\alpha}{1+\alpha}\right)\frac{[\mu_t-(f_i+h_i)]^2}{\sigma_t^2} \\
&\quad = \frac{1}{2}\frac{\alpha^2\rho^2(1-\lambda)^2}{(1+\alpha)[(\alpha+\lambda^2\rho^2)(1-\rho^2)+(\alpha+\lambda\rho^2)^2]}\frac{[\mu_t-(f_i+h_i)]^2}{\sigma_t^2} \\
&\quad \begin{cases} = 0, & \text{efficient market model,} \\ \geq 0, & \text{hedging model,} \end{cases}
\end{aligned}
$$

$$
\ln\left(\frac{1+\alpha^2(1-\rho^2)(\alpha+\lambda^2\rho^2)/\delta^2}{1+\alpha}\right) \leq 0.
$$

Turning now to the informed trader 'pre-info' expected utility,

$$
\begin{aligned}
\omega - \frac{1}{2}\frac{\psi^2}{1+2\gamma^2} &= aRm_i + a\mu_t(f_i+h_i) - \frac{1}{2}a^2\sigma_x^2(f_i+h_i)^2 \\
&\quad + \frac{1}{2}\frac{\alpha(\alpha+\lambda^2\rho^2)}{(\alpha+\lambda^2\rho^2)(1-\rho^2)+(\alpha+\lambda\rho^2)^2}\frac{[\mu_t-(f_i+h_i)]^2}{\sigma_t^2}
\end{aligned}
$$

$$
\ln\left(\left(\frac{1-\phi\rho^2}{1-\rho^2}\right)(1+2\gamma^2)\right) = \ln\left(\left(\frac{\alpha+\lambda^2\rho^2}{\alpha(1-\rho^2)+\lambda^2\rho^2}\right)\left(1+\frac{\alpha^2(1-\rho^2)(\alpha+\lambda^2\rho^2)}{\delta^2}\right)\right)
$$

[86] These expressions may be found using an approach similar to that used in appendix 1.

[87] The first difference below is positive in the hedging model, except when $\rho = 0$ or $\rho = 1$, at which points it is zero. The second difference is negative for both models except when $\rho = 0$ or $\rho = 1$, at which points it is zero.

In the efficient market model, we can ignore the first term, since $\mu_t - (f_i + h_i) = 0$. If we subtract the corresponding logarithmic term for the homogeneous case from the one above, it can be shown that the difference is unambiguously positive.[88]

$$\ln\left(\left(\frac{\alpha+\lambda^2\rho^2}{\alpha(1-\rho^2)+\lambda^2\rho^2}\right)\left(1+\frac{\alpha^2(1-\rho^2)(\alpha+\lambda^2\rho^2)}{\delta^2}\right)\right) - \ln(1+\alpha) \geq 0$$

A.7.3 Appendix 3. Case 3, Extended Model: Derivatives

In this appendix several partial derivatives are calculated for use in section A.4.3.b.i. The utility function of interest is the 'pre-info' expected utility of the uninformed trader in the efficient market, 'extended' model.

From appendix 1, we have[89]

$$\gamma^2 = \frac{1}{2}\frac{\lambda^2\rho^2\xi^2(1-\phi\rho^2)}{\phi(1-\rho^2+\lambda\rho^2\xi)^2},$$

$$\psi = a\sigma_x\rho\sqrt{\phi}(1+2\gamma^2)f,$$

$$\omega = \widehat{\omega} + aRm_i + af^2,$$

$$\widehat{\omega} = -\frac{1}{2}a^2\sigma_x^2[1-\phi\rho^2(1+2\gamma^2)]f^2.$$

[88] Except at the points $\rho = 0$ and $\rho = 1$, where the difference is zero.

[89] As this appendix is concerned only with the efficient market model, substitution has been made for $\mu_t = f$, $f_i = f$, $h_i = 0$ and $\sigma_t = \sigma_s$. Simplification has been performed.

Using the above, we can easily find that

$$\frac{1}{1+R}\left(\omega - \frac{1}{2}\frac{\psi^2}{1+R+2\gamma^2}\right)$$
$$= aRm_i + af^2 - \frac{1}{2}a^2\sigma_x^2 f^2$$
$$+ \frac{1}{2}a^2\sigma_x^2 R\phi\rho^2\left(\frac{1+2\gamma^2}{1+R+2\gamma^2}\right)f^2.$$

The first step to finding the effect of an increase in information quality on the uninformed trader expected utility is to find the partial derivative of the above. Using the results shown in appendix 1,

$$\frac{\partial\gamma^2}{\partial\rho^2} < 0,$$

we can show that

$$\frac{\partial(\phi\rho^2)}{\partial^2} > 0.$$

Given the above, we can see that

$$\frac{\partial}{\partial\rho^2}\left[\phi\rho^2\left(\frac{1+2\gamma^2}{1+R+\gamma^2}\right)\right]$$
$$= \frac{2R\phi\rho^2}{(1+r+2\gamma^2)^2}\frac{\partial\gamma^2}{\partial\rho^2} + \frac{1+2\gamma^2}{1+R+\gamma^2}\frac{\partial(\phi\rho^2)}{\partial\rho^2}$$
$$<> 0.$$

For example, at $\rho = 0$, $\partial(\phi\rho^2)/\partial\rho^2 = 0$, so that

$$\frac{\partial}{\partial\rho^2}\left[\phi\rho^2\left(\frac{1+2\gamma^2}{1+R+\gamma^2}\right)\right] = 0, \quad \text{at } \rho = 0.$$

Combining this with

$$\frac{\partial}{\partial\rho^2}\left(\frac{1}{2}\ln(1+R+2\gamma^2)\right) < 0, \quad \text{at } \rho = 0,$$

we have

$$\frac{\partial}{\partial \rho^2}\overline{K}^*_{iU} < 0, \quad \text{at } \rho = 0.$$

However, at $\rho = 1$, we have

$$\frac{\partial \gamma^2}{\partial \rho^2} = 0, \quad \frac{\partial(\phi\rho^2)}{\partial \rho^2} = 1, \quad \gamma^2 = 0, \quad \phi\rho^2 = 1, \qquad \text{at } \rho = 1,$$

so that

$$\frac{\partial}{\partial \rho^2}\left[\phi\rho^2\left(\frac{1+2\gamma^2}{1+R+\gamma^2}\right)\right] = \frac{1}{1+R}, \quad \text{at } \rho = 1.$$

Combined with

$$\frac{\partial}{\partial \rho^2}\left(\frac{1}{2}\ln(1+R+2\gamma^2)\right) = 0, \quad \text{at } \rho = 1,$$

we have

$$\frac{\partial}{\partial \rho^2}\overline{K}^*_{iU} > 0, \quad \text{at } \rho = 1.$$

Table A.1.

Decomposition of utility functions into component factors for the asymmetric information case of the efficient market model. Entries in the table are the signs of the partial derivative with respect to ρ^2 of the component factors. The last column shows the sign of the partial derivative of the utility function itself with respect to ρ^2.

		alloc[a] effic	dealer[b] benefit	net[c]
'standard' model	uninformed[d]		−	−
	informed		+/−[e]	+/−[e]
'extended' model	uninformed	+	−	+/−
	informed	+	+/−[e]	+/−

[a] The allocative efficiency benefit is only present in the 'extended' model.
[b] The dealer benefit is only present in the efficient market model. Its analog in the hedging model is the hedging benefit (see Table A.2).
[c] This column shows the net effect on the utility function of an increase in ρ^2.
[d] Uninformed refers to the representative uninformed trader utility function, informed to the representative informed utility.
[e] Note that this ambiguity of sign is present only in the asymmetric information case. In the homogeneous information case, all entries referring to this note are strictly negative.

Table A.2.

Decomposition of utility functions into component factors for the asymmetric information case of the hedging model. Entries in the table are the signs of the partial derivative with respect to ρ^2 of the component factors. The last column shows the sign of the partial derivative of the utility function itself with respect to ρ^2.

		alloc[a] effic	hedging[b] benefit	reval[c] risk	net[d]
'standard'	uninformed[e]		–	–	–
model	informed		+/–[f]	–	+/–[f]
'extended'	uninformed	+	–	–	+/–
model	informed	+	+/–[f]	–	+/–

[a] The allocative efficiency benefit is only present in the 'extended' model.

[b] The hedging benefit is only present in the hedging model. Its analog in the efficient market model is the dealer benefit (see Table A.1).

[c] Revaluation risk is only present in the hedging model.

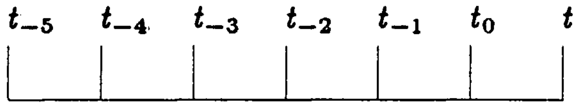
[d] This column shows the net effect on the utility function of an increase in ρ^2.

[e] Uninformed refers to the representative uninformed trader utility function, informed to the representative informed utility.

[f] Note that this ambiguity of sign is present only in the asymmetric information case. In the homogeneous information case, all entries referring to this note are strictly negative.

Figure A.1.

The sequence of events taking place in the models.



t_{-5} Endowments of the risky asset and riskless technology are received at this point. Common knowledge of all traders' utility functions is disseminated.

t_{-4} Trading to a Hakansson, Kunkel and Ohlson (1980) 'no-information' equilibrium position is allowed (in the efficient market and hedging models only, not in the Grossman and Stiglitz model).

t_{-3} Endowments of the non-tradable asset are received (in the hedging model). Common knowledge about who will be in the informed trader group plus the distribution functions of all random variables is disseminated.

t_{-2} Calculation of 'pre-info' expected utility.

t_{-1} Receipt of information by the informed trader group.

t_0 This is the beginning of period. Trading in the risky asset takes place. Consumption takes place (in the 'extended' model). 'Post-info' expected utility is calculated.

t_1 This is the end of period. The risky asset and riskless technology payoffs are received and consumed.

(This page intentionally left blank.)

PART B:

Bond Option Pricing, Empirical Evidence

B.1 INTRODUCTION

On October 22, 1982, US government bond, note and treasury bill option contracts began trading on the American Stock Exchange and the Chicago Board Options Exchange. Because these contracts have default-free government instruments as their underlying securities, the Brennan and Schwartz (1983a) two-factor model for pricing default-free, interest rate dependent options was used to compute theoretical prices for comparison with the actual market quotations now available.

It is only recently that the Brennan and Schwartz two-factor model was extended (Brennan and Schwartz (1983a)) to the valuation of interest rate dependent options.[1] In addition to this model, two other models – Courtadon (1982) and Ball and Torous (1983) – have been recently proposed for the valuation of interest rate dependent options. These models, however, are not considered in this study for the reasons outlined below. The Ball and Torous model uses the prices of two pure discount bonds as state variables to provide an analytic solution for the value of a European option on a pure discount bond. Although this model provides an analytic solution method, it does not appear extendible to the valuation of American options written on coupon bonds. As the Brennan and Schwartz two-factor model is based on numerical solution procedures for deriving option values, it is not limited in this fashion.

---

[1] The extension of contingent claims theory to interest rate dependent options was preceded by applications in the area of valuation of interest rate dependent claims such as government bonds. See Cox, Ingersoll and Ross (1978), Vasicek (1977), Richard (1976) and Brennan and Schwartz (1977, 1979, 1980, 1982, 1983b).

The Courtadon model, like the Brennan and Schwartz model, is also based on numerical solution procedures. In Brennan and Schwartz (1983a) it is shown that the Courtadon single-factor model can be viewed as a special case of the Brennan and Schwartz model. In their comparison, however, Brennan and Schwartz assumed that the stochastic process parameters and market preference parameters were given, and therefore the same for both models. In practice, these parameters are not given, and should be estimated separately for both models. It was felt that the substantial amount of effort required for reestimation of parameters for the Courtadon model would be beyond the scope of this study, and so a direct comparison of the Brennan-Schwartz and Courtadon models has been left as a topic for future research.

The tests that were performed in this study examined whether profits could be made by writing options when the Brennan and Schwartz model indicated that they were overvalued and buying them when undervalued. The trading strategy consisted of forming a theoretically riskless, zero-investment arbitrage portfolio, where the proper proportions, or hedge ratios, of assets held in the portfolio were calculated using results from the Brennan and Schwartz theoretical framework. It was found that the trading strategies did generate arbitrage profits, but that these profits were not sufficient to cover reasonable transactions costs that would be incurred if the strategies were actually implemented.[2] The Brennan and Schwartz model prices appear to be sufficiently accurate to justify practical use of the model for valuing interest rate dependent options.

---

[2] As noted in the conclusions to this study, care mst be exercised when interpreting the presence of these apparent before-transactions-costs arbitrage profits.

B.2 PRICING THEORY

The model presented in this section is similar to the multi-factor model for pricing contingent claims developed by Cox, Ingersoll and Ross (1978). Unlike Cox, Ingersoll and Ross, however, who develop a full general equilibrium model of an economy, the theory presented here relies on arbitrage arguments. The basic assumption of the model is that the underlying uncertainty in the economy can be modelled by a multivariate Wiener process $\mathbf{w}(t)$ evolving stochastically through time,[3] [4] and that there is an n-vector of state variables $\mathbf{x}(t)$ which are related to the Wiener processes by means of the Ito stochastic differential equation

$$d\mathbf{x} = \beta(\mathbf{x}, \mathbf{y}, t)\, dt + \eta(\mathbf{x}, \mathbf{y}, t)\, d\mathbf{w}(t),$$

where the Wiener process is characterized by

$$\mathrm{E}(d\mathbf{w}) = \mathbf{0}, \quad d\mathbf{w}\, d\mathbf{w}' = \mathbf{I}\, dt,$$

and the m-vector $\mathbf{y}(\mathrm{t})$ of non-stochastic state variables is described by

$$d\mathbf{y} = \gamma(\mathbf{y}, t)\, dt.$$

In order to simplify matters I also assume that $\eta\eta'$ is of full rank, and without loss of generality also let $\mathbf{x}$ and $\mathbf{w}$ both be n-vectors. The only other critical

[3] A very good introductory work on stochastic processes, also providing a review of the literature, is Maliaris and Brock (1982).

[4] I also make the standard assumptions that there are no taxes or transactions costs, no restrictions on short sales, and that trading is allowed to place at any point in time, ie. continuously.

assumption is that the assets I am pricing do indeed have prices that are functions only of the state variables $\mathbf{x}$, $\mathbf{y}$ and t.[5]

$$\mathbf{z} = \mathbf{z}(\mathbf{x}, \mathbf{y}, t)$$

For example, Brennan and Schwartz have typically used the instantaneous riskless rate of interest, r, and the yield on a consol bond, l, as their two state variables when pricing default-free government bonds. Hopefully, most of the uncertainty that affects the prices of the wide variety of government bonds is also reflected in the movement of these two state variables.[6] Indeed, the good fit of the Brennan-Schwartz two-factor model to actual market bond prices shows that this is not such a bad assumption to make, but if they had chosen instead to use the price of gold and the Dow-Jones market index as state variables perhaps their results would have been different.

B.2.1 Asset Pricing Theory

In the non-stochastic situation where we have $\mathbf{z}$ as a function of $\mathbf{y}$ and t, we can

[5] In theory, this is a perfectly valid assumption to make. The problem in practice is to *identify* exactly what $\mathbf{x}$ and $\mathbf{y}$ are.

[6] Note that using two other factors related to r and l by an invertible function would be equivalent to using r and l. That is, it is not essential that r and l be the correct underlying factors, but just that they are related to the two underlying factors. For example, use of $x = r - l$ and $y = \ln(l)$ as two factors would be theoretically equivalent to the use of r and l.

find the differential of $\mathbf{z}$ by straightforward partial differentiation as follows:

$$d\mathbf{z}(\mathbf{y},t) = \nabla_y \mathbf{z}\, d\mathbf{y} + \mathbf{z}_t\, dt, \quad (\mathbf{z} \text{ non-stochastic}),$$

where $\nabla'_y = (\partial/\partial y_1, \ldots, \partial/\partial y_m)$ is a vector operator, and $\nabla_y \mathbf{z}$ is shorthand for the matrix $\nabla'_y \otimes \mathbf{z}$.

When $\mathbf{z}$ is a stochastic variable, however, the situation becomes more complicated and we must use Ito's Lemma for the differential.

$$\begin{aligned} d\mathbf{z}(\mathbf{x},\mathbf{y},t) &= \nabla_x \mathbf{z}\, d\mathbf{x} + \nabla_y \mathbf{z}\, d\mathbf{y} + \mathbf{z}_t\, dt + \frac{1}{2}\mathrm{tr}(\eta\eta'\nabla_x\nabla'_x)\mathbf{z}\, dt \\ &= \left(\nabla_x \mathbf{z}\beta + \nabla_y \mathbf{z}\gamma + \mathbf{z}_t + \frac{1}{2}\mathrm{tr}(\eta\eta'\nabla_x\nabla'_x)\mathbf{z}\right) dt + \nabla_x \mathbf{z}\eta\, d\mathbf{w} \\ &= \mathbf{Z}(\mu\, dt + \mathbf{s}\, d\mathbf{w}) \end{aligned}$$

where $\mathrm{tr}(\eta\eta'\nabla_x\nabla'_x) = \sum_{i,j}(\eta\eta')_{ij}\,\partial^2/\partial x_i \partial x_j$ (ie. the trace of $\eta\eta'\nabla_x\nabla'_x$) is a scalar operator acting on all the elements of $\mathbf{z}$, and

$$\begin{aligned} \mathbf{Z} &= \mathrm{diag}(z_i), \\ \mu &= \mathbf{Z}^{-1}\left(\frac{1}{2}\mathrm{tr}(\eta\eta'\nabla_x\nabla'_x)\mathbf{z} + \nabla_x \mathbf{z}\,\beta + \nabla_y \mathbf{z}\,\gamma + \mathbf{z}_t\right), \\ \mathbf{s} &= \mathbf{Z}^{-1}\nabla_x \mathbf{z}\,\eta. \end{aligned}$$

The next few steps are the heart of the pricing theory and involve imposing a 'no arbitrage' rule on the assets. This is nothing new to finance and can be found as far back as Debreu (1959). More recently, we see the arbitrage condition in Black and Scholes' (1973) seminal option pricing paper, and find it forming the central core of Ross' (1976) arbitrage pricing theory.

In its usage here we form an arbitrage portfolio with total investment $p = \delta'\mathbf{1}$, where δ is a vector of the dollar amounts invested in the different assets.[7] The

[7] That is, δ_i is the dollar amount invested in asset i which has the unit price z_i.

return on this portfolio is, therefore,

$$\frac{dp}{p} = \delta' \mathbf{Z}^{-1}(d\mathbf{z} + \mathbf{c}\, dt) = \delta'(\mu + \mathbf{Z}^{-1}\mathbf{c})dt + \delta' \mathbf{s}\, d\mathbf{w},$$

where $\mathbf{c}$ is a vector of payouts *per unit* of asset (for example the coupon payment on a bond).

Now, if we let S be the subspace spanned by the columns of $\mathbf{s}$, any δ chosen from the orthogonal complement of S will give

$$\delta' \mathbf{s} = \mathbf{0}, \quad \forall \delta \in S^{\perp},$$

making the portfolio return totally non-stochastic, that is, riskless. By the no arbitrage rule, the return on this portfolio must be exactly the return that one would receive from a riskless investment, that is, $\delta' \mathbf{1}\, r\, dt$. Therefore, we must have

$$\delta'(\mu + \mathbf{Z}^{-1}\mathbf{c} - r\mathbf{1}) = 0, \quad \forall \delta \in S^{\perp}.$$

This can only be true for *all* $\delta \in S^{\perp}$, however, if we have

$$\mu + \mathbf{Z}^{-1}\mathbf{c} - r\mathbf{1} \in S,$$

in which case we can state that there is a vector function $\lambda(\mathbf{x}, \mathbf{y}, t)$ which satisfies

$$\mu + \mathbf{Z}^{-1}\mathbf{c} - r\mathbf{1} = \mathbf{s}\, \lambda.$$

Cox, Ingersoll and Ross (1978) showed that the function λ can be interpreted as the vector of prices that the market assigns to the uncertainty in the economy represented by the underlying Wiener processes, or, for short, the 'market price of risk function'.

At this point the theory is basically finished, since the arbitrage result above gives us a partial differential equation for any asset price z. Before replacing μ and $\mathbf{s}$ in the arbitrage result above, however, since I am mainly interested in assets which have a maturity date, I will replace time t with time left to maturity τ. Note that this results in $\mathbf{z}_\tau = -\mathbf{z}_t$ since $d\tau = -dt$. Combining this with the definitions of μ and $\mathbf{s}$, the partial differential equation for an individual asset price becomes

$$\frac{1}{2}\mathrm{tr}(\eta\eta'\nabla_x\nabla_x')z + \nabla_x z(\beta - \eta\lambda) + \nabla_y z\,\gamma + c - rz = z_\tau.$$

The only difficulty now is that we don't have expressions for the functions $\beta(\mathbf{x}, \mathbf{y}, t)$ and $\eta(\mathbf{x}, \mathbf{y}, t)$ arising from the stochastic differential equation for $\mathbf{x}$, $\gamma(\mathbf{y}, t)$ from the equation in $\mathbf{y}$, nor for the market price of risk function $\lambda(\mathbf{x}, \mathbf{y}, t)$. If we knew these functions, then in principle the partial differential equation would be solved.[8] There is nothing else to be done about β, η or γ here, but Brennan and Schwartz (1979) made an important observation about the market price of risk function. As they pointed out, if the price of one of the assets in the economy is known as a function of the state variables and time, then we can very simply identify a linear combination of the risk prices λ.

That is, given a known price function z we can calculate μ and $\mathbf{s}$ to produce

$$\mu + c/z - r = \mathbf{s}\,\lambda,$$

which is a linear combination of the market prices of risk. If we had a full basis, $\mathbf{z}_b$, of these known asset price functions then we could replace λ altogether by

$$\lambda = \mathbf{s}_b^{-1}(\mu_b + \mathbf{Z}_b^{-1}\mathbf{c}_b - r\mathbf{1}),$$

[8] In practice, of course, we would still have the problem of estimating the solution.

and reduce the asset pricing partial differential equation to

$$\mu + c/z - r = \mathbf{s}\,\mathbf{s}_b^{-1}(\mu_b + \mathbf{Z}_b^{-1}\mathbf{c}_b - r\mathbf{1}).$$

The only exception to this rule concerns assets which have prices dependent on the instantaneous riskless rate, r. Any asset with a price dependent only on the instantaneous riskless rate must have $\mu + c/z - r = 0$, thus making these assets useless for determining λ even if the riskless rate r is one of the state variables.

B.2.1.a THE BRENNAN-SCHWARTZ MODEL

In this section we look at a special case of the theory which has been proposed for pricing options on default-free bonds, namely, the Brennan-Schwartz model, and see how simplifications were achieved by using several quite reasonable assumptions. First, as mentioned above, Brennan and Schwartz chose as their state variables the instantaneous riskless rate, r, and the yield on a consol bond, l.[9]

$$\mathbf{x} = \begin{pmatrix} r \\ l \end{pmatrix}$$

Already, a very important choice has been made. As pointed out in Brennan and Schwartz (1979), since we know that the price function of a consol bond, V, is

$$V(\mathbf{x}, t) = 1/l,$$

9 There are no non-stochastic state variables, $\mathbf{y}$.

choosing l as a state variable allows us to simplify the partial differential equation by solving for a linear combination of the price of risk function λ. In order to simplify matters even more, Brennan and Schwartz made the further assumption that β and η are time-independent and that the correlation between the stochastic processes for r and l is independent of the levels of r and l, that is,

$$\beta(\mathbf{x},t) = \beta(\mathbf{x}), \quad \eta(\mathbf{x},t) = \eta(\mathbf{x}),$$

$$\eta(\mathbf{x}) = \sigma(\mathbf{x})\,\mathbf{G}, \quad \sigma(\mathbf{x}) = \begin{pmatrix} \sigma_1(\mathbf{x}) & 0 \\ 0 & \sigma_2(\mathbf{x}) \end{pmatrix}, \quad \mathbf{G}\mathbf{G}' = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix},$$

where ρ is a constant correlation coefficient. This allows us to define a transformation of λ,

$$\lambda_x = \begin{pmatrix} \lambda_r \\ \lambda_l \end{pmatrix} = \mathbf{G}\,\lambda,$$

giving

$$\eta\lambda = \sigma\,\lambda_x = \begin{pmatrix} \sigma_1\lambda_r \\ \sigma_2\lambda_l \end{pmatrix}.$$

Now, when we evaluate the partial differential equation for the known consol bond price function we find that

$$\nabla_x V = (0, -\frac{1}{l^2}), \quad V_r = 0, \quad \mathrm{tr}(\eta\eta'\nabla_x\nabla_x')V = 2\frac{\sigma_2^2}{l^3} \qquad \text{and} \qquad c = 1 = l\,V,$$

so that

$$\begin{aligned} &\frac{1}{2}\mathrm{tr}(\eta\eta'\nabla_x\nabla_x')V + \nabla_x V(\beta - \sigma\lambda_x) + c - rV \\ &\qquad = \frac{\sigma_2^2}{l^3} - \frac{1}{l^2}(\beta_2 - \sigma_2\lambda_l) + (l-r)\frac{1}{l} \\ &\qquad = V_\tau = 0. \end{aligned}$$

This gives us an expression for β_2,

$$\beta_2 = l\,(\frac{\sigma_2^2}{l^2} + l - r) + \sigma_2\lambda_l,$$

and allows us to eliminate λ_l from the partial differential equation. The simplified equation is

$$\frac{1}{2}\sigma_1^2 z_{rr} + \rho\sigma_1\sigma_2 z_{rl} + \frac{1}{2}\sigma_2^2 z_{ll} + (\beta_1 - \sigma_1\lambda_r)z_r + l(\sigma_2^2/l^2 + l - r)z_l + c - rz = z_\tau.$$

It might seem that there is little advantage in all of this maneuvering to replace the unknown λ_l since l is itself unobservable. There is, after all, no consol bond outstanding in the United States or Canada. In effect the problem boils down to a choice between either (a) using l as a state variable in the theory and then finding some observable proxy to l in order to estimate the covariance function σ of the stochastic process of r and l or (b) using the yield of an outstanding bond as a state variable and then estimating the additional price of risk function. The advantage of simplicity seems to lie with the first choice. For example, if we used the yield on an issued bond as a state variable we would have to worry about the changing time left to maturity of the bond as time passed and what effect that change would have on the parameters we are trying to estimate.

As is clear from the development of the theory, the resulting partial differential equation applies to *all* assets with prices which are dependent only on the state variables and time. We use the same differential equation to find the prices of different coupon bonds and also options on these bonds. The difference, therefore, lies in the boundary conditions that we impose on the solution.

For a discount bond with price $\delta(r, l, \tau)$, time to maturity τ, and a principal value of \$100,[10] the boundary condition is simply the payout that the holder receives at maturity of the bond,

$$\delta(r, l, 0) = 100.$$

When pricing a european call, $C_E(r, l, \tau)$, or put, $P_E(r, l, \tau)$ once again the boundary condition is simply the payout that the holder receives at maturity of the option

$$C_E(r, l, 0; K) = \max(0, B(r, l, \tau_B; c) - K),$$

$$P_E(r, l, 0; K) = \max(0, K - B(r, l, \tau_B; c)),$$

where K is the exercise price of the option and $B(r, l, \tau_B; c)$ is the price of underlying bond having a coupon rate c and τ_B time left to maturity as of the date the option matures.

The only options on government issues that are currently traded, however, are American options, that is, options that can be exercised at any point in time. Since this is the case, we need an extra boundary condition for an American option preventing its price from falling below what the holder would receive if he exercised the option. That is, at all times up to and including maturity we must have

$$C(r, l, \tau; K) \geq \max(0, B(r, l, \tau + \tau_B; c) - K),$$

$$P(r, l, \tau; K) \geq \max(0, K - B(r, l, \tau + \tau_B; c)),$$

with equality holding at maturity, $\tau = 0$, of the option.

The above boundary conditions hold for the usual type of option, namely, an

[10] In this study I have simplified matters by standardizing all bonds to a face value of \$100.

option on a specific underlying bond. There are also options being traded where the underlying instrument changes over time. In particular, we need to value options where the security deliverable upon exercise has a fixed time to maturity, $\bar{\tau}$. For these 'fixed maturity' options, we have the following boundary conditions,

$$C(r,l,\tau;K) \geq \max(0, B(r,l,\bar{\tau};c) - K),$$
$$P(r,l,\tau;K) \geq \max(0, K - B(r,l,\bar{\tau};c)),$$

with equality once again holding at maturity.[11]

The only question remaining is how to value a coupon bond. I will follow Brennan and Schwartz in this matter, and assume that we can neglect any tax effects. With this being the case, a coupon bond becomes simply a portfolio of discount bonds with the value

$$B(r,l,\tau;c) = \sum_{\tau_c \leq \tau} c\,\delta(r,l,\tau-\tau_c) + \delta(r,l,\tau)$$

where τ_c are the times to maturity of the different bond coupon payouts, and c is the coupon rate of the bond.

[11] Strictly speaking, the asset pricing theory as developed above does not hold if the underlying asset changes continuously, as it does here. This is because the theoretical arguments are based on forming an arbitrage portfolio and holding it for an instant of time. One of the assets included in this portfolio is supposed to be the asset underlying the option, so an implicit assumption is that the underlying asset can be held for an instant of time. This, of course, is not true if the underlying asset changes continually. If, however, we modify the boundary condition so that it is imposed only at a countable number of instants – that is, exercise is only allowed at a countable number of instants – then the 'underlying asset' at any point in time is the asset which is deliverable at the next permissible exercise time, and the theory is once more valid.

Another special case of the general asset pricing model is the Black and Scholes (1973) single-factor pricing model. When Black and Scholes brought this model forward, their main concern was the pricing of derivative assets, so their model contained only a single stochastic state variable x, which was taken to be the price of the underlying asset. This single state variable simplicity has a price, of course, as the random nature of the riskless rate, r, is left undescribed. Since r is present in the asset pricing partial differential equation, we may find that our asset prices are in fact sensitive to unexpected changes in r.

The Black-Scholes asset pricing partial differential equation, therefore, is

$$\frac{1}{2}\eta^2 z_{xx} + (\beta - \eta\lambda) z_x + \gamma\, z_r + c - rz = z_\tau,$$

where x is the price of the asset underlying the derivative asset z, c is the per unit payout on z and

$$dx = \beta(x)\, dt + \eta(x)\, dw, \quad dr = \gamma(r, t)\, dt.$$

If we value the underlying asset x itself using this partial differential equation, since $x_x = 1$, $x_{xx} = 0$, $x_r = 0$ and $x_\tau = 0$, we have

$$\beta - \eta\lambda + c_x - rx = 0,$$

where c_x is the per unit payout on the asset x. This simplifies the equation to

$$\frac{1}{2}\eta^2 z_{xx} + (rx - c_x) z_x + \gamma\, z_r + c - rz = z_\tau.$$

This is the Black-Scholes pricing equation except for several simplifying assumptions that they made, namely,

$$\eta = \sigma_x x, \quad \sigma_x \text{ constant},$$

$$\gamma = 0, \quad \text{ie. } r \text{ constant},$$

$$c = 0, \quad \text{the payout on } z \text{ is zero},$$

which give the classic Black-Scholes equation

$$\frac{1}{2}\sigma_x^2 x^2 z_{xx} + (rx - c_x)z_x - rz = z_\tau.$$

In this study I use all of these common assumptions, so that bond option prices are assumed to conform to the differential equation[12]

$$\frac{1}{2}\sigma_B^2 B^2 z_{BB} - (rB - c_B)z_B - rz = z_\tau,$$

where c_B is the instantaneous *dollar* interest accrual of the underlying bond. Note that only one parameter, σ_B, has to be estimated, making this model quite simple when compared to the Brennan-Schwartz model.

As with the Brennan-Schwartz model, boundary conditions for American call and put options are

$$C(B,\tau;K) \geq \max(0, B-K)),$$

$$P(B,\tau;K) \geq \max(0, K-B)),$$

with equality holding at maturity.

[12] The variable denoting the underlying asset has been changed from x to B to stress that the underlying asset is a bond (or treasury bill) in this study.

B.2.2 Arbitrage Portfolios

As was shown in the theory section above, the entire asset pricing theory rests on a 'no arbitrage' rule which is assumed to hold in an efficient market. If we were able to form the 'arbitrage' portfolios mentioned there, we could test whether or not this no arbitrage condition really holds in the market. The problem is that in order to calculate the amount of each asset to hold in the arbitrage portfolio we need a theory to tell us what the subspace S spanned by the columns of $\mathbf{s} = \mathbf{Z}^{-1}\nabla_x\mathbf{z}\,\eta$ is. As a first step to identifying a suitable arbitrage portfolio, note that if we let T be the subspace spanned by the columns of $\mathbf{Z}^{-1}\nabla_x\mathbf{z}$, then

$$\delta'\mathbf{Z}^{-1}\nabla_x\mathbf{z}\,\eta = (\delta'\mathbf{Z}^{-1}\nabla_x\mathbf{z})\eta = 0, \quad \forall\delta \in T^{\perp}$$

which means that $T^{\perp} = S^{\perp}$ and $T = S$. This means that in order to form an arbitrage portfolio we need only find a vector of dollar amounts of each asset, $\delta \in T$. That is, we need only find a δ orthogonal to $Z^{-1}\nabla_x\mathbf{z}$.

For the case where we have n state variables, if we can find an n-vector of assets, $\mathbf{z}_b$, which form a non-singular basis $\mathbf{Z}_b^{-1}\nabla_x\mathbf{z}_b$, then this basis spans the subspace S and we can express $(\nabla_x z_h)/z_h$ of any asset to be hedged as a linear combination of the columns of $\mathbf{Z}_b^{-1}\nabla_x\mathbf{z}_b$. That is, if

$$\mathbf{z} = \begin{pmatrix} z_h \\ \mathbf{z}_b \end{pmatrix}, \quad \delta = \begin{pmatrix} \delta_h \\ \delta_b \end{pmatrix}, \quad \nu = \begin{pmatrix} \nu_h \\ \nu_b \end{pmatrix} = \mathbf{Z}^{-1}\delta = \begin{pmatrix} \delta_h/z_h \\ \mathbf{Z}_b^{-1}\delta_b \end{pmatrix},$$

then

$$\delta'\mathbf{Z}^{-1}\nabla_x\mathbf{z} = \frac{\delta_h}{z_h}\nabla_x z_h + \delta_b'\mathbf{Z}_b^{-1}\nabla_x\mathbf{z}_b = \nu_h\nabla_x z_h + \nu_b'\nabla_x\mathbf{z}_b = 0,$$

or

$$\nu_b' = -\nu_h \nabla_x z_h (\nabla_x \mathbf{z}_b)^{-1},$$

where ν is a vector of the *number of units* of each asset in the arbitrage portfolio.

Therefore, given an n-vector of assets, $\mathbf{z}_b$, which allows us to span S we can hedge away the risk of any other asset. Since the theory requires the resulting riskless portfolio to return the riskless rate, if we combine the asset positions above with an appropriate position in the riskless asset, that is, invest $-\delta'\mathbf{1}$ in the riskless asset, the result is a zero-investment arbitrage portfolio. Because it is a zero-investment portfolio, it should have a zero return.

The test of the no arbitrage rule is therefore a type of market efficiency test. Of course, it is really a joint test of market efficiency, the particular models that I use, and the accuracy of the parameters that I estimate in order to derive theoretical option prices, but this is true of all tests of market efficiency.[13] The procedure that will be followed here is to look for discrepancies between the theoretical and market prices of a bond option and then form a hopefully riskless, zero-investment arbitrage position to take advantage of any mispricings that are found. That is, we choose

$$\nu_h > 0 \quad \text{if} \quad z_h > z_h^M, \quad \text{buy option when undervalued}$$

$$\nu_h < 0 \quad \text{if} \quad z_h < z_h^M, \quad \text{write option when overvalued}$$

where z_h^M is the observed market price of the bond option.

[13] That is, if there appear to be arbitrage possibilities, then either the market is inefficient, the models used in this study are incorrect, or the parameter estimates used to derive theoretical prices were inaccurate.

B.3 DATA DESCRIPTION

The main aim of this study is the comparison of bond option price quotations with theoretical model prices. Naturally, this creates two needs for data. We certainly need to collect the bond option price quotations, but we also need data to help us create the theoretical prices. In the previous section on asset pricing it became clear that before assets can be priced with an asset pricing partial differential equation we must first estimate all the parameters of the underlying stochastic differential equation, and also any parameters involved in the market price of risk function.

In this study, I have chosen to collect data from two non-overlapping periods. The first period is the 'estimation period', and consists of monthly observations running from October 1970 through October 1982. Data from this period is used to estimate any needed model parameters.[14] The estimation period is followed by the 'test period', consisting of daily data covering the period from November 1, 1982 through October 31, 1983. Data from this period is used to test the asset pricing models and perform the arbitrage tests.[15]

---

[14] The longer the estimation period is, the greater the number of data points available and hopefully the better my parameter estimates. On the other hand, the longer the estimation period, the more likely it is that changes occur in the parameters during the period (ie. non-stationarity of parameters). I chose a 12 year estimation period as, hopefully, a good compromise between these two opposing factors.

[15] That is, the arbitrage tests done in this study will be testing whether or not there are arbitrage possibilities given past data series on bond and treasury bill prices. If traders in fact form their expectations based on more sources of information, and markets are efficient, then we would not expect to find any arbitrage profits in the tests done in this study.

B.3.1 Parameter Estimation Data

As shown above, the Brennan-Schwartz pricing model produces the asset pricing equation

$$\frac{1}{2}\sigma_1^2 z_{rr} + \rho\sigma_1\sigma_2 z_{rl} + \frac{1}{2}\sigma_2^2 z_{ll} + (\beta_1 - \sigma_1\lambda_r)z_r + l\,(\sigma_2^2/l^2 + l - r)z_l + c - rz = z_\tau,$$

from the stochastic differential equation in r and l

$$\begin{pmatrix} dr \\ dl \end{pmatrix} = \begin{pmatrix} \beta_1 \\ l\,(\sigma_2^2/l^2 + l - r) + \sigma_2\lambda_l \end{pmatrix} dt + \begin{pmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{pmatrix} \mathbf{G}\, d\mathbf{w}, \quad \mathbf{GG}' = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}.$$

If we were given parameterized forms for β_1, σ_1, σ_2 and λ_l we could use the stochastic differential equation, along with data series for r and l from the estimation period, to estimate these parameters, leaving us with only λ_r still unknown.

Since λ_r is a market price of risk function and, as we noted in the previous section, is the only market risk price that cannot be eliminated from the pricing equation by using an observable asset price, we are forced to estimate this function by comparing theoretical and market prices over the estimation period. That is, we must try various values for the parameters in λ_r and choose those values which result in the best fit between theoretical prices calculated with the asset pricing partial differential equation and actual market prices.

B.3.1.a SHORT RATE SERIES

I mentioned above that estimation of the stochastic process parameters requires a time series for the instantaneous riskless rate of interest r. Of course, there is no such series available and we must instead find an acceptable proxy for r. As my proxy I chose the continuously compounded yield on the outstanding treasury bill which was the closest to having 30 days left to maturity. This data series was readily available for each month-end from the CRSP US government bond tape. I selected monthly proxy values from the estimation period October 1970 through October 1982. This period lies just before the November 1, 1982 start of the testing period containing the bond option price data to be tested.

B.3.1.b CONSOL RATE SERIES

Just as there is no instantaneous riskless rate series available, there is also no series available for the yield on a consol bond: there is no consol bond outstanding in the United States. We are forced once again to find an acceptable proxy for the unavailable series. The proxy that I chose was the yield on a very long maturity bond, which should provide a good approximation to the yield on a consol. As long as the bond's par value repayment is so far in the future that it is discounted almost to zero, the yields on the long term bond and consol should be quite close.

The proxy that I used, therefore, was the continuously compounded yield of the outstanding bond with the longest time left to maturity, under the condition that the bond also be normally taxable.[16] [17] This series was also collected from the CRSP US government bond tape for each month-end over the period of October 1970 through October 1982, a total of 145 observations.[18]

B.3.1.c MARKET PRICE OF SHORT RATE RISK PARAMETERS

Once the parameters of the stochastic differential equation in r and l have been estimated, we can proceed to the estimation of the market determined risk price function λ_r. Since the only way to estimate this function is to actually calculate the theoretical prices of some assets and compare them to the quoted market prices, the asset pricing partial differential equation must be solved repeatedly for different values of the parameters of λ_r until the best fitting parameters are found.

---

[16] The tax treatment of certain bonds, termed 'flower' bonds, is different from the treatment of most bonds, resulting in prices higher than those on normally taxable bonds.

[17] I also chose a second proxy the same as the above but with an additional condition: the bond also had to be trading within a $10 dollar range of par. This condition was added just in case there were tax effects due to differential treatment of coupon payouts and capital gains. As there was no significant difference in the parameter estimates from these two series, the added condition was considered unnecessary and was dropped.

[18] Actually, a total of 289 observations covering the period October 1958 through October 1982 were collected for both the r and l series. The first half of the series from October 1958 to October 1970 were used to test how variable the parameter estimates were from one time period to the next. The parameter estimates from this first half of the estimation period were not otherwise used.

Once again, the best available source of pertinent market data is the CRSP US government bond tape, from which I obtained month-end market bond prices for the estimation period of October 1970 through October 1982.[19] The price used was whatever was available on the CRSP tape, either an actual sale price, bid price or ask price or, if both bid and ask prices were given, the middle of the bid-ask spread.

B.3.2 Test Period Data

The option data that is needed for this study was not available in computer readible format, and had to be collected from quotations published in the Wall Street Journal. The test period follows directly after the parameter estimation period, and runs from November 1, 1982 to October 31, 1983. Data were collected for each trading day in this period.

[19] Only normally taxable bonds and notes were chosen.

B.3.2.a BOND OPTION DATA

Throughout the entire year-long test period there were only options outstanding on five notes and three bonds, with all of the note options listed on the American Stock Exchange (AMEX) and all of the bond options listed on the Chicago Board Options Exchange (CBOE). Because these options mature 9 months after their initial listing[20] and are listed each quarter, there may be as many as three options outstanding which differ only by date of maturity.[21]

Table B.1 shows a summary of the bond and bond option data collected. In all, there were a total of 274 different bond options traded during the one year test period, generating 3793 bond option price observations. As is shown in Table B.2, most of these options traded at prices below \$5.[22] Two further breakdowns show that most of the option trades occurred with less than 5 months left to maturity on the option (Table B.3) and that the options trade close to the money (Table B.4).

There are several technical details that should be mentioned here. First, the note option contracts traded on AMEX are 'small contracts', that is, the underlying principal amount of the note is \$20,000. On the CBOE, bond option contracts

[20] Note and bond options mature and are listed at the end of the third Friday of March, June, September and December.

[21] Exchange rules actually allow listing of options with up to 15 months to maturity, which would allow trading of up to 5 options differing only by maturity date. Neither exchange has listed options with more than 9 months to maturity.

[22] I have standardized all the options so that the underlying bonds have a principal value of \$100.

were either 'small contracts' or 'large contracts', where a 'large contract' has an underlying principal amount of $100,000.[23]

Second, the exercise prices shown in Table B.1 must be adjusted, just as bond price quotations must be adjusted, to take into account accrued interest. For example, the holder of a call option with an exercise price of $102 per $100 of principal value would have to pay $102 plus the accrued interest on the underlying bond in order to exercise his option.

Third, the actual bond option price quotation is shown in the Wall Street Journal as a decimal amount. The decimal portion actually represents 32^{nd}s of a dollar, so that a quotation of $2.10 per $100 principal value is actually a price of $2 10/32.

Lastly, there is a delay of two business days between exercise of a bond option and final settlement. For example, if an option is exercised on Thursday, then the exercise price plus accrued interest on the underlying bond up to and including the *settlement* date must be paid on Monday, the settlement date.

[23] As mentioned in a previous footnote, I ignore this aspect and standardize all options to a contract size of $100.

B.3.2.b *TREASURY BILL OPTION DATA*

The treasury bill option is considerably more complicated than the bond option we have just looked at. Unlike the bond option, the treasury bill option exercise settlement date is not two trading days after exercise, but is instead the Thursday of the week following the week in which the option is exercised.[24] Also, the underlying security which must be supplied on the settlement date is not the same from week to week, as with bond options. The deliverable treasury bill is one which has 13 weeks to maturity as of the settlement date.[25] Naturally the deliverable treasury bill changes every week.

As an example, we can suppose that the writer of a call option has his option called on Monday. In order to lock in the value of his settlement date obligations, he purchases a treasury bill which will have 13 weeks left to maturity on the settlement date. Since the Monday exercise date and Thursday settlement date are 10 days apart, on the exercise date he would buy a treasury bill with 14 weeks and 3 days left to maturity. If the exercise instead took place on a Friday, then the settlement date would be only 6 days away and the writer would purchase a treasury bill with 13 weeks and 6 days left to maturity in order to lock in his settlement date obligations. Therefore, the maturity of the underlying treasury bill is, strictly speaking, not 13 weeks, but varies from 13 weeks and 6 days to 14 weeks and 3 days.

---

[24] Or the next trading day following the Thursday if it is a holiday.

[25] Exchange rules permit options on 26-week treasury bills also, but these have not been listed.

As with bond options, treasury bill options are listed quarterly and initially have 9 months to maturity.[26] The result is that there could be up to 3 options outstanding which differ only by maturity date. Table B.5 shows a summary of the treasury bill option data collected. There were a total of 37 treasury bill options listed during the test period, and a total of 819 treasury bill option price observations were collected.

At this point I must mention the method used to adjust the quoted treasury bill option exercise prices. If, for example, the exercise price quoted is k, then the actual price payable on settlement is $K = 100 - (100 - k)91/360$ dollars per \$100 dollars principal value. For $k = 90$, say, this gives $K = \$97.4722$. The exercise prices used in this study have been converted from the quoted value to the actual dollar amount payable.

Several other details should be mentioned concerning the contract size traded and the method of quoting treasury bill option prices. First, treasury bill options have traded only on AMEX and only in contracts with \$200,000 underlying principal value.[27] The method of quoting treasury bill option prices is also quite different from bond options. The quoted price is given in decimal form and is in fact a decimal number, but must be adjusted as follows. If p is the quoted treasury bill

[26] Treasury bill options mature and are listed at the end of the third Friday in March, June, September and December.

[27] This is a 'small contract' for a 13-week treasury bill option. A 'large contract' for a 13-week treasury bill option would have an underlying principal value of \$1,000,000. The 'small contract' and 'large contract' sizes for the as yet untraded 26-week treasury bill options are \$100,000 and \$500,000 respectively. As usual, I ignore contract size and standardize all options to have an underlying pricipal value of \$100.

option 'premium', then the actual price payable for a 13-week treasury bill option is $P = p\,\frac{13}{52}$ per \$100 principal value.[28] Note that the factor for price adjustment, 13/52, is different from the adjustment factor for exercise prices, 91/360.

Returning once more to the collected data, we see in Table B.6 that most of the treasury bill options traded at prices under 30 cents. Two further breakdowns show that – as with bond options – most of the trades took place with less than 5 months left to maturity on the option (Table B.7) and that the options trade close to the money (Table B.8).

B.3.2.c ARBITRAGE PORTFOLIO DATA

The hedging theory section above showed that we could hedge away the risk of any asset in an n-factor model by combining it with the proper holdings of n other assets forming a basis over the n-dimensional risk space. Therefore, when one is dealing with the Black-Scholes model, since the risk space is one-dimensional, the risk of a bond or treasury bill option can be hedged away by holding the correct amount of the underlying bond or treasury bill.[29]

---

[28] If the option had been one on a 26-week treasury bill we would have had $P = p\,\frac{26}{52}$.

[29] In fact, the theory allows one to hedge away the option risk by using *any* bond or treasury bill, not just the underlying instrument.

When dealing with the Brennan-Schwartz model, the risk space is two-dimensional, so a second hedging asset is needed. I decided to complete the hedge portfolio by using a bond having 5 years to maturity. When combined with the 20 to 30 year maturity underlying bond of a bond option, this 5 year bond would mainly hedge the short rate risk while the position in the underlying bond mainly hedges consol rate risk. With a 13-week treasury bill option the situation is exactly the opposite. The hedge portfolio position in the underlying treasury bill mainly hedges the short rate risk while the 5 year bond position hedges mainly consol rate risk. The 5 year bond is sufficiently different from the underlying bonds and treasury bills to be used in both cases.

So that I would be able to form the required zero-investment arbitrage portfolios, I collected from the Wall Street Journal price data for each underlying bond or treasury bill on all of the days that at least one option on the bond or treasury bill traded. Both bid and ask prices were recorded.[30] For the 5 year maturity bond I chose two bonds which appeared to be heavily traded: their bid-ask spreads were narrow. Over the first part of the test period, from November 1, 1982 to April 30, 1983, I used the 12 5/8% bond maturing November 15, 1987, and for the latter part of the test period, from March 1, 1983, to October 31, 1983 I chose the 9 7/8% bond maturing May 15, 1988.[31] Both bid and ask prices were collected

[30] In the Wall Street Journal bond prices are quoted in decimal format. The decimal portion of the number represents 64^{th}s of a dollar, so that a quotation of \$100.4 is really \$100 4/64 per \$100 principal value. Treasury bill prices are quoted in discount form, that is, if the discount quoted is d, the actual price payable is $P = 100 - d\frac{n}{360}$ where n is the number of days to maturity of the treasury bill. For example, the price of a 13-week treasury bill quoted at a discount of 8.68 would be $P = 100 - 8.68\,\frac{91}{360} = \$97.8059$ per \$100 principal value.

[31] There was a period of overlap in the data collected for these two bonds, since an arbitrage portfolio formed *before* March 1, 1983 would contain the 12 5/8% bond

for these two bonds on each trading day of the periods given above.

The last data required is a time series for the riskless rate r on each day of the test period. This is needed for two purposes. First, we need r in testing the no arbitrage condition. The zero-investment arbitrage portfolio requires an investment in the riskless asset, so we must know what the riskless rate is before we can test for positive expected arbitrage returns. Second, since the numerical solution of bond and option prices will provide prices as a function of r and l, r is needed in order to calculate the theoretical price of an option on any given day.

For this latter reason, I chose the same proxy for r as I chose in the estimation period, namely, the continuously compounded yield on the treasury bill having the closest to 30 days left to maturity. These data were collected from the Wall Street Journal for each trading day of the test period. Both bid and ask prices were collected and transformed from discount form. The yield of the price midway between the bid and ask prices was used to proxy r.

maturing November 1987. If there was no period of overlap, then we would run into trouble if the portfolio was held past March 1, 1983, as we would not be able to provide a price for this bond when the portfolio was liquidated.

The Brennan Schwartz asset pricing model results in the partial differential equation

$$Lz = z_\tau, \quad z = z(r, l, \tau)$$

$$L = a\frac{\partial^2}{\partial r^2} + b\frac{\partial^2}{\partial r \partial l} + c\frac{\partial^2}{\partial l^2} + d\frac{\partial}{\partial r} + e\frac{\partial}{\partial l} + f$$

with a, b, c, d, e and f functions of r and l only.[32] The first step in the solution of this equation is replacing the differential operator L by the finite difference operator $\hat{L}$,[33]

$$\hat{L} = a\frac{\delta_r^2}{\Delta r^2} + b\frac{H_r H_l}{4\Delta r \Delta l} + c\frac{\delta_l^2}{\Delta l^2} + d\frac{H_r}{2\Delta r} + e\frac{H_l}{2\Delta l} + f,$$

where the difference operators δ and H are defined to have the following effects on a function of r and l:

$$\delta_r^2 f(r,l) = f(r+\Delta r, l) - 2f(r,l) + f(r-\Delta r, l),$$

$$\delta_l^2 f(r,l) = f(r, l+\Delta l) - 2f(r,l) + f(r, l-\Delta l),$$

$$H_r f(r,l) = f(r+\Delta r, l) - f(r-\Delta r, l),$$

$$H_l f(r,l) = f(r, l+\Delta l) - f(r, l+\Delta l).$$

The reason for using these particular difference approximations is that they are the ones that follow from the Taylor series expansions of f, so that

$$\hat{L}z = Lz + O(\Delta r^2 + \Delta l^2) = z_\tau + O(\Delta r^2 + \Delta l^2$$

32 Strictly speaking, the partial differential equation also contains a term due to possible asset payouts, such as the coupon payout on a bond. Since I will only be numerically solving the asset princing partial differential equation for the values of assets which do *not* make payouts – discount bonds and options – I ignore this additional term.

33 See Varga (1962) for a detailed outline of the methods used in this section.

The solution, Z, that we would obtain from solving the mixed difference-differential equation, namely,

$$\hat{L}Z = Z_\tau$$

is not equal to the real solution, z. If we let $v = z - Z$ be the difference between the approximate and real solutions, we see from the two equations above that

$$\hat{L}(z - Z) = (z - Z)_\tau + O(\Delta r^2 + \Delta l^2),$$

or

$$\hat{L}v = v_\tau + \epsilon, \quad \epsilon(r, l, 0) = 0, \quad \epsilon(r, l, \tau) = O(\Delta r^2 + \Delta l^2)$$

Since the initial conditions are known at maturity ($\tau = 0$), the initial value for ϵ is zero.

If we now choose a regular two-dimensional grid of r and l points with a spacing of Δr between the I points in the r direction and Δl between the J points in the l direction, we can define the vectors[34]

$$\mathbf{z}(\tau) = \begin{pmatrix} z(r_1, l_1, \tau) \\ \vdots \\ z(r_I, l_1, \tau) \\ \vdots \\ z(r_1, l_J, \tau) \\ \vdots \\ z(r_I, l_J, \tau) \end{pmatrix}, \quad \mathbf{Z}(\tau) = \begin{pmatrix} Z(r_1, l_1, \tau) \\ \vdots \\ Z(r_I, l_1, \tau) \\ \vdots \\ Z(r_1, l_J, \tau) \\ \vdots \\ Z(r_I, l_J, \tau) \end{pmatrix}, \quad \epsilon(\tau) = \begin{pmatrix} \epsilon(r_1, l_1, \tau) \\ \vdots \\ \epsilon(r_I, l_1, \tau) \\ \vdots \\ \epsilon(r_1, l_J, \tau) \\ \vdots \\ \epsilon(r_I, l_J, \tau) \end{pmatrix}$$

[34] The vector $\mathbf{Z}(\tau)$ is defined so that $Z(r_i, l_j, \tau)$, which is the value of $Z(r, l, \tau)$ at the intersection of the ith column and jth row of the grid, is the $i + (j-1)I$st element of the vector $\mathbf{Z}(\tau)$. That is, the first I elements of $\mathbf{Z}(\tau)$ are those in the first row of the grid, namely $Z(r_1, l_1, \tau)$ through $Z(r_I, l_1, \tau)$. The next I elements are from the second row of the grid, and so on until the Jth row.

and replace the difference operator $\hat{L}$ by the matrix $\mathbf{G}$ to give[35]

$$\mathbf{Gz} = \mathbf{z}_r + \mathbf{s} + \epsilon,$$

$$\mathbf{GZ} = \mathbf{Z}_r + \mathbf{s},$$

$$\mathbf{Gv} = \mathbf{v}_r + \epsilon, \quad \mathbf{v} = \mathbf{z} - \mathbf{Z},$$

where the matrix $\mathbf{G}$ is defined as below.[36]

$$\begin{aligned}
\mathbf{G} = \mathbf{A}\frac{(\mathbf{M}_r - 2\mathbf{I} + \mathbf{M}_r')}{\Delta r^2} + \mathbf{B}\frac{(\mathbf{M}_r - \mathbf{M}_r')}{2\Delta r}\frac{(\mathbf{M}_l - \mathbf{M}_l')}{2\Delta l} + \mathbf{C}\frac{(\mathbf{M}_l - 2\mathbf{I} + \mathbf{M}_l')}{\Delta l^2} \\
+ \mathbf{D}\frac{(\mathbf{M}_r - \mathbf{M}_r')}{2\Delta r} + \mathbf{E}\frac{(\mathbf{M}_l - \mathbf{M}_l')}{2\Delta l} + \mathbf{F} + \mathbf{S},
\end{aligned}$$

$\mathbf{M}_r$ = matrix of all zeros except the first

upper diagonal which is all ones,

$\mathbf{M}_l$ = matrix of all zeros except the I^{th}

upper diagonal which is all ones,

$$\mathbf{A} = \text{diag}(a_k), \quad a_{i+(j-1)I} = a(r_i, l_j),$$

$$\mathbf{B} = \text{diag}(b_k), \quad b_{i+(j-1)I} = b(r_i, l_j),$$

$$\mathbf{C} = \text{diag}(c_k), \quad c_{i+(j-1)I} = c(r_i, l_j),$$

$$\mathbf{D} = \text{diag}(d_k), \quad d_{i+(j-1)I} = d(r_i, l_j),$$

$$\mathbf{E} = \text{diag}(e_k), \quad e_{i+(j-1)I} = e(r_i, l_j),$$

$$\mathbf{F} = \text{diag}(f_k), \quad f_{i+(j-1)I} = f(r_i, l_j).$$

For example, $a(r_i, l_j)$ is the $i + (j-1)I$th element along the diagonal of $\mathbf{A}$.

35 The vector $\mathbf{s}$ is the result of imposing boundary conditions on the edges of the (r, l) grid. For simplicity, it will not be explicitly considered here.
36 The matrix $\mathbf{S}$ is also the result of imposing boundary conditions on the edges of the (r, l) grid.

The solutions to these mixed difference-differential systems can be shown to be

$$\mathbf{Z}(\tau) = -\mathbf{G}^{-1}\mathbf{s} + \exp(\tau\mathbf{G})[\mathbf{Z}(0) + \mathbf{G}^{-1}\mathbf{s}],$$

$$\mathbf{v}(\tau) = \int_0^\tau \exp[(\tau - \nu)\mathbf{G}]\epsilon(\nu)\, d\nu.$$

The above equation in $\mathbf{v}(\tau)$ can be used to show that if all of the characteristic values of $\mathbf{G}$ have negative real parts, then $\mathbf{v}(\tau)$ is bounded as τ increases to infinity.[37] If this condition does not hold, then the matrix $\mathbf{G}$ is said to be unstable and we cannot guarantee that $\mathbf{v}$ remains bounded as τ increases to infinity. That is, the finite difference solution, $\mathbf{Z}$, can not be guaranteed to be 'close' to the real solution $\mathbf{s}$ unless the characteristic values of $\mathbf{G}$ all have negative real parts.[38]

B.4.1 The Alternating Direction Method

The alternating direction method is based on an approximation to $\exp(\Delta\tau\mathbf{G})$ in the solution to the mixed difference-differential equation found in the previous section.

$$\mathbf{Z}(\tau + \Delta\tau) = -\mathbf{G}^{-1}\mathbf{s} + \exp(\Delta\tau\mathbf{G})(\mathbf{Z}(\tau) + \mathbf{G}^{-1}\mathbf{s})$$

Before making this approximation, the key of the alternating direction method is first decomposing $\mathbf{G}$ into $\mathbf{G}_r$ and $\mathbf{G}_l$

$$\mathbf{G} = \mathbf{G}_r + \mathbf{G}_l,$$

[37] See Varga (1962).

[38] If all the characteristic values of $\mathbf{G}$ have negative real parts, then the effects of discretization errors dies out exponentially as we solve for larger and larger values of τ. If this condition does not hold, then we may find the effects of discretization errors *increasing* exponentially as we solve for larger and larger values of τ.

where the exact composition of $\mathbf{G}_r$ and $\mathbf{G}_l$ is discussed below.

Once this decomposition has been made, we approximate $\exp(\Delta\tau\mathbf{G})$ by $\mathbf{T}$.

$$\begin{aligned}
\mathbf{T} &= \left(\mathbf{I} - \frac{1}{2}\Delta\tau\mathbf{G}_r\right)^{-1}\left(\mathbf{I} - \frac{1}{2}\Delta\tau\mathbf{G}_l\right)^{-1}\left(\mathbf{I} + \frac{1}{2}\Delta\tau\mathbf{G}_l\right)\left(\mathbf{I} + \frac{1}{2}\Delta\tau\mathbf{G}_r\right) \\
&= \mathbf{I} + \Delta\tau\mathbf{G} + \frac{\Delta\tau^2}{2}\mathbf{G}^2 + O(\Delta\tau^3) \\
&= \exp(\Delta\tau\mathbf{G}) + O(\Delta\tau^3).
\end{aligned}$$

This approximation is consistent,[39] as $\mathbf{T}$ agrees with $\exp(\Delta\tau\mathbf{G})$ up to the linear term $\Delta\tau\mathbf{G}$.

At first glance, the system of equations that result from discretizing the time dimension using the approximation $\mathbf{T}$ seems to be no easier to solve than any other system resulting from a consistent approximation.

$$\begin{aligned}
&\left(\mathbf{I} - \frac{1}{2}\Delta\tau\mathbf{G}_l\right)\left(\mathbf{I} - \frac{1}{2}\Delta\tau\mathbf{G}_r\right)\mathbf{Z}(\tau + \Delta\tau) \\
&\quad = \left(\mathbf{I} + \frac{1}{2}\Delta\tau\mathbf{G}_l\right)\left(\mathbf{I} + \frac{1}{2}\Delta\tau\mathbf{G}_r\right)\mathbf{Z}(\tau) \\
&\qquad + \left[\left(\mathbf{I} + \frac{1}{2}\Delta\tau\mathbf{G}_l\right)\left(\mathbf{I} + \frac{1}{2}\Delta\tau\mathbf{G}_r\right) - \left(\mathbf{I} - \frac{1}{2}\Delta\tau\mathbf{G}_l\right)\left(\mathbf{I} - \frac{1}{2}\Delta\tau\mathbf{G}_r\right)\right]\mathbf{G}^{-1}\mathbf{s} \\
&\quad = \left(\mathbf{I} + \frac{1}{2}\Delta\tau\mathbf{G} + \frac{1}{4}\Delta\tau^2\mathbf{G}_l\mathbf{G}_r\right)\mathbf{Z}(\tau) + \Delta\tau\mathbf{s}
\end{aligned}$$

If, however, we follow the Douglas and Rachford (1956) alternating direction format and define $\mathbf{Z}^*(\tau + \Delta\tau)$ to satisfy

$$\left(\mathbf{I} - \frac{1}{2}\Delta\tau\mathbf{G}_l\right)\mathbf{Z}^*(\tau + \Delta\tau) = \left(\mathbf{I} + \frac{1}{2}\Delta\tau(\mathbf{G} + \mathbf{G}_r)\right)\mathbf{Z}(\tau),$$

[39] Any approximation which agrees with the first two power series terms of $\exp(\Delta\tau\mathbf{G})$, namely, $\mathbf{I} + \Delta\tau\mathbf{G}$, is a consistent approximation.

then the solution $\mathbf{Z}(\tau + \Delta\tau)$ is found from[40]

$$\left(\mathbf{I} - \frac{1}{2}\Delta\tau\mathbf{G}_r\right)\mathbf{Z}(\tau + \Delta\tau) = \mathbf{Z}^*(\tau + \Delta\tau) - \frac{1}{2}\Delta\tau\mathbf{G}_r\mathbf{Z}(\tau) + \Delta\tau\mathbf{s}^*,$$
$$\left(\mathbf{I} - \frac{1}{2}\Delta\tau\mathbf{G}_l\right)\mathbf{s}^* = \mathbf{s}.$$

Once we define exactly what $\mathbf{G}_r$ and $\mathbf{G}_l$ are, it becomes clear exactly why this is called an alternating direction method. Assuming for the moment that the coefficient b of the cross-derivative term bz_{rl} in the partial differential equation is zero, we can decompose $\mathbf{G}$ as follows:

$$\mathbf{G} = \mathbf{G}_r + \mathbf{G}_l,$$
$$\mathbf{G}_r = \left(\frac{\mathbf{A}}{\Delta r^2} + \frac{1}{2}\frac{\mathbf{D}}{\Delta r}\right)\mathbf{M}_r + \left(-2\frac{\mathbf{A}}{\Delta r^2} + \frac{1}{2}\mathbf{F}\right) + \left(\frac{\mathbf{A}}{\Delta r^2} - \frac{1}{2}\frac{\mathbf{D}}{\Delta r}\right)\mathbf{M}'_r + \mathbf{S}_r,$$
$$\mathbf{G}_l = \left(\frac{\mathbf{C}}{\Delta l^2} + \frac{1}{2}\frac{\mathbf{E}}{\Delta l}\right)\mathbf{M}_l + \left(-2\frac{\mathbf{C}}{\Delta l^2} + \frac{1}{2}\mathbf{F}\right) + \left(\frac{\mathbf{C}}{\Delta l^2} - \frac{1}{2}\frac{\mathbf{E}}{\Delta l}\right)\mathbf{M}'_l + \mathbf{S}_l.$$

Since each of these matrices is tridiagonal, solution of the two-factor model is reduced to the solution of I tridiagonal systems of order J plus J tridiagonal systems of order I at each time step, where I and J are the number of grid points in the r and l dimensions. First a set of tridiagonal systems is solved to give $\mathbf{Z}^*(\tau+\Delta\tau)$, and this intermediate value is used in solving the next set of tridiagonal systems, producing $\mathbf{Z}(\tau + \Delta\tau)$. The method is equivalent to alternately solving a one-factor model in the l dimension followed by a second one-factor model in the r dimension. Hence the name alternating direction.

[40] The original papers on the alternating direction method did not address the adjustment of boundary conditions, that is, the use of $\mathbf{s}^*$ instead of $\mathbf{s}$. For a discussion of this topic see Fairweather and Mitchell (1967).

Given a certain form of $\mathbf{G}_r$ and $\mathbf{G}_l$ it is relatively easy to establish the stability of the Douglas-Rachford alternating direction method. If, that is, $\mathbf{G}_r$ and $\mathbf{G}_l$ have only negative diagonal and non-negative off-diagonal elements, then all of their characteristic values, η_{rk} and η_{lk}, have negative real parts.[41] Since the solution after n time steps is

$$\begin{aligned}\mathbf{Z}(n\Delta\tau) + \mathbf{G}^{-1}\mathbf{s} &= \mathbf{T}(\mathbf{Z}((n-1)\Delta\tau) + \mathbf{G}^{-1}\mathbf{s}) \\ &= \mathbf{T}^n(\mathbf{Z}(0) + \mathbf{G}^{-1}\mathbf{s}),\end{aligned}$$

the stability of the system is assured if $||\mathbf{T}||$ is less than 1.

$$\begin{aligned}||\mathbf{T}|| &= \left\| \left(\mathbf{I} - \frac{1}{2}\Delta\tau\mathbf{G}_r\right)^{-1} \left(\mathbf{I} - \frac{1}{2}\Delta\tau\mathbf{G}_l\right)^{-1} \left(\mathbf{I} + \frac{1}{2}\Delta\tau\mathbf{G}_l\right) \left(\mathbf{I} + \frac{1}{2}\Delta\tau\mathbf{G}_r\right) \right\| \\ &\leq \left\| \left(\mathbf{I} - \frac{1}{2}\Delta\tau\mathbf{G}_l\right)^{-1} \left(\mathbf{I} + \frac{1}{2}\Delta\tau\mathbf{G}_l\right) \right\| \left\| \left(\mathbf{I} + \frac{1}{2}\Delta\tau\mathbf{G}_r\right) \left(\mathbf{I} - \frac{1}{2}\Delta\tau\mathbf{G}_r\right)^{-1} \right\| \\ &\leq \left| \max_m \frac{1 + \frac{1}{2}\Delta\tau \mathrm{Re}(\eta_{lm})}{1 - \frac{1}{2}\Delta\tau \mathrm{Re}(\eta_{lm})} \right| \left| \max_n \frac{1 + \frac{1}{2}\Delta\tau \mathrm{Re}(\eta_{rn})}{1 - \frac{1}{2}\Delta\tau \mathrm{Re}(\eta_{rn})} \right|\end{aligned}$$

Since all of the characteristic values of $\mathbf{G}_r$ and $\mathbf{G}_l$ are assumed to have negative real parts, $||\mathbf{T}||$ is in fact less than 1, verifying the stability of the Douglas-Rachford method.

Up to this point the analysis has assumed that the coefficient b of the bz_{rl} term is zero. Without this assumption, we are not able to decompose $\mathbf{G}$ into two tridiagonal matrices. There have been several alternative methods proposed for handling cases such as ours, where b is not zero. The Douglas and Gunn (1964) method extends the original alternating direction idea one step further by decomposing $\mathbf{G}$ into *three* parts

$$\mathbf{G} = \mathbf{G}_r + \mathbf{G}_l + \mathbf{G}_{rl},$$

[41] See Varga (1962) for the properties of positive matrices.

where $\mathbf{G}_r$ and $\mathbf{G}_l$ are the same as defined above, and $\mathbf{G}_{rl}$ is the result of a finite difference approximation to bz_{rl} which is slightly different from the one I have used in previous sections. The approximation to $\exp(\Delta\tau\mathbf{G})$ then becomes

$$\exp(\Delta\tau\mathbf{G}) = \left(\mathbf{I} - \frac{1}{2}\Delta\tau\mathbf{G}_{rl}\right)^{-1}\left(\mathbf{I} - \frac{1}{2}\Delta\tau\mathbf{G}_r\right)^{-1}\left(\mathbf{I} - \frac{1}{2}\Delta\tau\mathbf{G}_l\right)^{-1}$$
$$\times\left(\mathbf{I} + \frac{1}{2}\Delta\tau\mathbf{G}_l\right)\left(\mathbf{I} + \frac{1}{2}\Delta\tau\mathbf{G}_r\right)\left(\mathbf{I} + \frac{1}{2}\Delta\tau\mathbf{G}_{rl}\right) + O(\Delta\tau^3).$$

Unfortunately, this method also requires the solution of three sets of tridiagonal systems at each time step instead of the two sets of systems required in the Douglas-Rachford method.[42]

The necessity of solving three sets of tridiagonal systems at each time step can be avoided, however, as was shown by McKee and Mitchell (1970). This is made possible by using a backward difference approximation to the bz_{rl} term instead of the Crank-Nicolson approximation implied in the Douglas-Gunn procedure. That is, they approximate $\exp(\Delta\tau\mathbf{G})$ by

$$\exp(\Delta\tau\mathbf{G}) = \left(\mathbf{I} - \frac{1}{2}\Delta\tau\mathbf{G}_r\right)^{-1}\left(\mathbf{I} - \frac{1}{2}\Delta\tau\mathbf{G}_l\right)^{-1}$$
$$\times\left[\left(\mathbf{I} + \frac{1}{2}\Delta\tau\mathbf{G}_l\right)\left(\mathbf{I} + \frac{1}{2}\Delta\tau\mathbf{G}_r\right) + \Delta\tau\mathbf{G}_{rl}\right] + O(\Delta\tau^2),$$

where

$$\mathbf{G}_{rl} = \frac{1}{4}\frac{\mathbf{B}}{\Delta r\Delta l}(\mathbf{M}_r - \mathbf{M}_r')(\mathbf{M}_l - \mathbf{M}_l'),$$

and $\mathbf{G}_r$ and $\mathbf{G}_l$ are as in the Douglas-Rachford method. This still produces a consistent approximation.

[42] Actually, if b may be both positive and negative at different points of the grid, the Douglas-Gunn method splits $\mathbf{G}$ into *four* parts, and requires the solution of four sets of tridiagonal systems at each time step. In our problem, however, b is either always positive or always negative, depending on the sign of the correlation coefficient ρ. Because b is either positive or negative, but not both, one of the four splittings becomes zero.

Since making this substitution for $\exp(\Delta\tau\mathbf{G})$ produces the same system of equations as the Douglas-Rachford method, namely,

$$\begin{aligned}\left(\mathbf{I}-\frac{1}{2}\Delta\tau\mathbf{G}_l\right)&\left(\mathbf{I}-\frac{1}{2}\Delta\tau\mathbf{G}_r\right)\mathbf{Z}(\tau+\Delta\tau)\\ &=\left[\left(\mathbf{I}+\frac{1}{2}\Delta\tau\mathbf{G}_l\right)\left(\mathbf{I}+\frac{1}{2}\Delta\tau\mathbf{G}_r\right)+\Delta\tau\mathbf{G}_{rl}\right]\mathbf{Z}(\tau)\\ &\quad+\left[\left(\mathbf{I}+\frac{1}{2}\Delta\tau\mathbf{G}_l\right)\left(\mathbf{I}+\frac{1}{2}\Delta\tau\mathbf{G}_r\right)+\Delta\tau\mathbf{G}_{rl}\right.\\ &\qquad\left.-\left(\mathbf{I}-\frac{1}{2}\Delta\tau\mathbf{G}_l\right)\left(\mathbf{I}-\frac{1}{2}\Delta\tau\mathbf{G}_r\right)\right]\mathbf{G}^{-1}\mathbf{s}\\ &=\left(\mathbf{I}+\frac{1}{2}\Delta\tau\mathbf{G}+\frac{1}{4}\Delta\tau^2\mathbf{G}_l\mathbf{G}_r\right)\mathbf{Z}(\tau)+\Delta\tau\mathbf{s},\end{aligned}$$

using the Douglas-Rachford alternating direction form gives us once again

$$\begin{aligned}&\left(\mathbf{I}-\frac{1}{2}\Delta\tau\mathbf{G}_l\right)\mathbf{Z}^*(\tau+\Delta\tau)=\left(\mathbf{I}+\frac{1}{2}\Delta\tau(\mathbf{G}+\mathbf{G}_r)\right)\mathbf{Z}(\tau),\\ &\left(\mathbf{I}-\frac{1}{2}\Delta\tau\mathbf{G}_r\right)\mathbf{Z}(\tau+\Delta\tau)=\mathbf{Z}^*(\tau+\Delta\tau)-\frac{1}{2}\Delta\tau\mathbf{G}_r\mathbf{Z}(\tau)+\Delta\tau\mathbf{s}^*,\\ &\left(\mathbf{I}-\frac{1}{2}\Delta\tau\mathbf{G}_l\right)\mathbf{s}^*=\mathbf{s}.\end{aligned}$$

The only difference between this and the equivalent systems found using the Douglas-Rachford method is the difference in the decomposition of $\mathbf{G}$.[43]

One of the drawbacks of the McKee-Mitchell method is the difficulty of showing stability. Because of the unusual form of their approximation to $\exp(\Delta\tau\mathbf{G})$, stability cannot in general be shown. Although in general we cannot show stability, McKee and Mitchell were able to show that their method is stable when applied to the problem

$$az_{rr}+bz_{rl}+cz_{ll}=z_\tau,\qquad b^2\le 4ac,\quad a>0,\quad c>0,$$

[43] That is, here we have $\mathbf{G}=\mathbf{G}_r+\mathbf{G}_l+\mathbf{G}_{rl}$, whereas in the Douglas-Rachford case with $b=0$ we had $\mathbf{G}=\mathbf{G}_r+\mathbf{G}_l$.

$$a = a(r,l), \quad b = b(r,l), \quad c = c(r,l),$$

a fact which turns out to be useful below.

B.4.2 Stability of the Solution Method

The Douglas-Rachford alternating direction method is only guaranteed stable if the characteristic values of both $\mathbf{G}_r$ and $\mathbf{G}_l$ all have negative real parts. Unfortunately, the form of $\mathbf{G}$ in our particular problem does not tell us whether this condition holds.

If we look at the $\mathbf{G}$ matrix for the Brennan-Schwartz two-factor model, namely,

$$\begin{aligned}\mathbf{G} = &\left(\frac{\mathbf{A}}{\Delta r^2} + \frac{1}{2}\frac{\mathbf{D}}{\Delta r}\right)\mathbf{M}_r + \left(-2\frac{\mathbf{A}}{\Delta r^2} + \frac{1}{2}\mathbf{F}\right) + \left(\frac{\mathbf{A}}{\Delta r^2} - \frac{1}{2}\frac{\mathbf{D}}{\Delta r}\right)\mathbf{M}_r' \\ &+ \left(\frac{\mathbf{C}}{\Delta l^2} + \frac{1}{2}\frac{\mathbf{E}}{\Delta l}\right)\mathbf{M}_l + \left(-2\frac{\mathbf{C}}{\Delta l^2} + \frac{1}{2}\mathbf{F}\right) + \left(\frac{\mathbf{C}}{\Delta l^2} - \frac{1}{2}\frac{\mathbf{E}}{\Delta l}\right)\mathbf{M}_l' \\ &+ \frac{1}{4}\frac{\mathbf{B}}{\Delta r \Delta l}(\mathbf{M}_r - \mathbf{M}_r')(\mathbf{M}_l - \mathbf{M}_l'),\end{aligned}$$

where

$$\begin{aligned}
\mathbf{A} &= \mathrm{diag}(a_k), \quad a_{i+(j-1)I} = \frac{1}{2}\sigma_1^2(r_i, l_j), \\
\mathbf{B} &= \mathrm{diag}(b_k), \quad b_{i+(j-1)I} = \rho\sigma_1(r_i, l_j)\sigma_2(r_i, l_j), \\
\mathbf{C} &= \mathrm{diag}(c_k), \quad c_{i+(j-1)I} = \frac{1}{2}\sigma_2^2(r_i, l_j), \\
\mathbf{D} &= \mathrm{diag}(d_k), \quad d_{i+(j-1)I} = \beta_1(r_i, l_j) - \sigma_1(r_i, l_j)\lambda_r(r_i, l_j, t), \\
\mathbf{E} &= \mathrm{diag}(e_k), \quad e_{i+(j-1)I} = l_j(\sigma_2^2(r_i, l_j)/l_j^2 + l_j - r_i), \\
\mathbf{F} &= \mathrm{diag}(f_k), \quad f_{i+(j-1)I} = -r_i,
\end{aligned}$$

since a_k and c_k are positive and f_k is negative,

$$\mathbf{A} > 0, \quad \mathbf{C} > 0, \quad \mathbf{F} \leq 0,$$

it is clear that the diagonal of **G** is made up of negative elements. If the off-diagonal elements were all non-negative, then all of the characteristic values of $\mathbf{G}_r$ and $\mathbf{G}_l$ would be guaranteed negative and the stability of our numerical methods would be assured.[44] We see, however, that the bz_{rl} term produces two negative elements per row of **G**. As the partial differential equation is elliptic in its space variables, ie. $b^2 \leq 4ac$, it turns out that these elements do not cause instability.[45]

There are, however, other potentially negative elements. If either of the following conditions holds then **G** will have negative off-diagonal elements.

$$\left|\frac{1}{2}\frac{d_k}{\Delta r}\right| > \left|\frac{a_k}{\Delta r^2}\right| \qquad \text{or} \qquad \left|\frac{1}{2}\frac{e_k}{\Delta l}\right| > \left|\frac{c_k}{\Delta l^2}\right|$$

In using the Brennan-Schwartz model, I assume that the variances of r and l are as below.

$$\sigma_1(r_i, l_j) = \sigma_r r_i, \quad \sigma_r \text{ constant},$$

$$\sigma_2(r_i, l_j) = \sigma_l l_j, \quad \sigma_l \text{ constant}.$$

The variance of changes in the short rate is proportional to the short rate r, so that as r approaches zero so does its variance. Since the drift component of r does not approach zero at the same time, we find that some off-diagonal elements of **G** are negative for small values of r. This raises the possibility of instability.

44 See Varga (1962) for the theory of positive matrices.
45 This can be seen by an analysis similar to the one performed by McKee and Mitchell.

For grid points where either r or l is small, we may find $\mathbf{G}$ containing negative off-diagonal elements.

As a test of the stability of the method, I numerically solved the Brennan-Schwartz model using both the alternating direction and successive overrelaxation (SOR) methods with the same stochastic process forms and parameter values as were found by Brennan and Schwartz (1982). The solution did indeed show instability, occuring when either the r and l grid was not extended to large enough interest rates, or when too large a spacing was used between grid points. For the parameters and grid dimensions used by Brennan and Schwartz, namely, r running from 0.00 to 0.50 with $\Delta r = 0.01$, l running from 0.00 to 0.50 with $\Delta l = 0.01$ and $\Delta \tau = 1/24$ the solution was well-behaved. This stability was also present when various other parameter values were tried. As a result, I decided to use these grid dimensions above when solving the partial differential equation in later parts of the study.[46] [47]

---

[46] The instability of the solution was not the result of using the alternating direction solution method instead of the successive overrelaxation method (SOR) used by Brennan and Schwartz (1982). In a comparison of the alternating direction and SOR methods, I found that the alternating direction method was, in fact, less affected by instability.

[47] One unsettling aspect of instability in a system such as this is the possibility that the 'correct' parameter values remain undiscovered because they lie in a region where the solution method is unstable.

B.5 PARAMETER ESTIMATION

Up until now, the joint stochastic process for r and l has been discussed in quite general terms. In this section we must finally specify particular parameterized forms for the process drift and variance terms, and for the market price of short term risk λ_r.

Brennan and Schwartz (1982) estimated the parameters in parameterized forms of β_1, β_2, σ_1 and σ_2 by using a time series of 30 day treasury bill yields and a series of yields on a very long maturity bond to proxy for r and l, respectively. Once these parameters were estimated, they were left with only the market price of short rate risk λ_r to estimate. This they estimated by finding the value of λ_r which resulted in the best fit between market bond prices and theoretical bond prices. The theoretical bond prices were, of course, found by numerical solution of the asset pricing partial differential equation for various values of λ_r.

In the next section, I first examine the procedure used by Brennan and Schwartz to estimate the parameters of the joint stochastic process for r and l. The conclusion I reach is that for the particular estimates found in this study, no confidence can be placed in the estimates for the parameters of β_1 or β_2. The estimates of the parameters of σ_1 and σ_2 and the correlation ρ between the two Wiener processes can, however, be well estimated. As a result, unlike Brennan and Schwartz, I only estimate σ_1, σ_2 and ρ using the two time series of yields. I then simultaneously estimate λ_r and the parameters of β_1 by finding the values producing the best fitting theoretical bond values.

B.5.1 The Simple Linearization Method

In their first papers on the two-factor model for bond pricing, Brennan and Schwartz (1977, 1979, 1980) used joint stochastic processes for r and l which allowed one to solve the forward equation analytically. This basically limited them to processes which were linear, that is, of the form

$$\begin{pmatrix} dr \\ dl \end{pmatrix} = \left(\mathbf{A} \begin{pmatrix} r \\ l \end{pmatrix} + \mathbf{b} \right) dt + \mathbf{C}\, d\mathbf{w},$$
$$d\mathbf{w} d\mathbf{w}' = \mathbf{I}\, dt,$$

where $\mathbf{A}$ and $\mathbf{C}$ are constant matrices and $\mathbf{b}$ is a constant vector. As shown by Phillips (1972) this system can be solved for Δr and Δl.

Unfortunately, the solution of a linear differential system is a special case. It is not possible to find an analytic solution to a general non-linear differential system, such as is used in this study. As an alternative Ananthanarayanan (1978) and Brennan and Schwartz (1982) proposed what they called a simple linearization method. Instead of directly solving the problem, they suggested the following approximation.

$$\begin{aligned} \int_t^{t+\Delta t} \begin{pmatrix} dr(s) \\ dl(s) \end{pmatrix} &= \int_t^{t+\Delta t} \beta(r(s), l(s), s)\, ds + \int_t^{t+\Delta t} \eta(r(s), l(s), s)\, d\mathbf{w}(s) \\ &\simeq \beta(r(t), l(t), t) \int_t^{t+\Delta t} ds + \eta(r(t), l(t), t) \int_t^{t+\Delta t} d\mathbf{w}(s), \end{aligned}$$

or

$$\eta^{-1}(r(t),l(t),t)\begin{pmatrix}\Delta r(t)\\ \Delta l(t)\end{pmatrix} \simeq \eta^{-1}(r(t),l(t),t)\beta(r(t),l(t),t)\Delta t + \Delta\mathbf{w}(t),$$

$$\Delta\mathbf{w}(t) \sim N\{\mathbf{0},\mathbf{I}\Delta t\}.$$

That is, this approximation assumes that both β and η are approximately constant over the interval Δt.

Clearly, this result can be used in maximizing the likelihood function

$$L = P(r_1,l_1,t_1)\prod_{i=2}^{T} P(r_i,l_i,t_i|r_{i-1},l_{i-1},t_{i-1}),$$

$$P(r_i,l_i,t_i|r_{i-1},l_{i-1},t_{i-1}) = \frac{1}{2\pi\Delta t}\exp\left(-\frac{1}{2}\frac{\Delta\mathbf{w}'_{i-1}\Delta\mathbf{w}_{i-1}}{\Delta t}\right),$$

$$\Delta\mathbf{w}_{i-1} = \eta^{-1}(r_{i-1},l_{i-1},t_{i-1})\left[\begin{pmatrix}\Delta r_{i-1}\\ \Delta l_{i-1}\end{pmatrix} - \beta(r_{i-1},l_{i-1},t_{i-1})\Delta t\right].$$

B.5.1.a MINIMUM DISTANCE ESTIMATOR

The section above ended by suggesting that after making the simple linearization approximation the results could be used to maximize a likelihood function. It would indeed be likelihood maximization if the matrix function η was known, but since η contains unknown parameters the estimates derived from the maximization are only asymptotically maximum likelihood estimators as the number of observations increases to infinity.[48] As is easily seen, maximizing the 'likelihood'

[48] See Malinvaud (1966) and Phillips (1972).

function given above is equivalent to minimizing the distance function

$$D = \sum_{i=2}^{t} \Delta \mathbf{w}'_{i-1} \Delta \mathbf{w}_{i-1},$$

$$\Delta \mathbf{w}_{i-1} = \eta^{-1}(r_{i-1}, l_{i-1}, t_{i-1}) \left[\begin{pmatrix} \Delta r_{i-1} \\ \Delta l_{i-1} \end{pmatrix} - \beta(r_{i-1}, l_{i-1}, t_{i-1}) \Delta t \right],$$

except for the factor $P(r_1, l_1, t_1)$ which I have ignored in this study.

B.5.1.b A ONE-DIMENSIONAL EXAMPLE OF SIMPLE LINEARIZATION

The first test of the simple linearization method proved quite successful. In his thesis Ananthanarayanan (1978) showed that the simple linearization parameter estimates were almost identical to the estimates one would find by analytic solution of a particular one-dimensional stochastic process for the instantaneous riskless rate r. The process examined was

$$dr = m(\mu - r)\,dt + r^{\alpha}\sigma\,dw, \quad dw \sim N\{0, dt\},$$

where m, μ, α and σ are constants.

We can better analyze the situation if we change to a process which is homoscedastic. By Ito's Lemma,

$$dx(r) = x_r dr + \frac{1}{2} x_{rr} (dr)^2$$
$$= \left[\frac{1}{2} x_{rr} r^{2\alpha} \sigma^2 + x_r m(\mu - r) \right] dt + x_r r^{\alpha} \sigma\, dw.$$

If we choose $x = r^{1-\alpha}/\sigma(1-\alpha)$, then $x_r r^\alpha \sigma = 1$ and we have

$$dx = \left[\frac{m}{\sigma}(\mu - r)r^{-\alpha} - \frac{1}{2}\alpha\sigma r^{\alpha-1}\right] dt + dw.$$

Now, if we make the simple linearization assumption that the drift term changes very little over the interval Δt, we have

$$\Delta x_i \simeq \left[\frac{m}{\sigma}(\mu - r_i)r_i^{-\alpha} - \frac{1}{2}\alpha\sigma r_i^{\alpha-1}\right] \Delta t + \Delta w_i, \quad \Delta w_i \sim N\{0, \Delta t\}.$$

As can be seen from Tables B.9 and B.10, the assumption that the drift term changes very little over the span of a month is a good one given the specific parameter values found by Ananthanarayanan. For a range of beginning of month values of r ranging from 0.05 to 0.20 we see that the end of period values are practically unchanged.[49] In fact, for the second set of parameter values shown in Table B.10, the end of period range is too small to show up with only three decimal digits.

Actually, the way I have presented the simple linearization method is slightly different from the way that it has actually been used by Ananthanarayanan and Brennan and Schwartz. In my example, I first changed variables from r to x in order to end up with a homoscedastic process. The simple linearization procedure used by Ananthanarayanan (1978) and Brennan and Schwartz (1982) would instead use

$$\frac{\Delta r_i}{r_i^\alpha \sigma} \simeq \frac{m(\mu - r_i)}{r_i^\alpha \sigma} \Delta t + \Delta w_i.$$

[49] I calculated these values using the simple linearization method. The end of month range of r given in the last column is a range of two standard deviations around the mean – again computed using simple linearization.

The agreement between my procedure and the original should be good since we have shown that r_i^α changes little over Δt.

B.5.2 Brennan-Schwartz Parameter Estimates

The simple linearization procedure has been shown to work well in the one particular case studied by Ananthanarayanan, but this is not justification for its use in any other case. Justification would only come from showing that the assumptions of the simple linearization method are valid.

In the Brennan-Schwartz scenario that I consider here, the first step in verifying these assumptions is to transform the process

$$\begin{pmatrix} dr \\ dl \end{pmatrix} = \beta(r,l)\,dt + \eta(r,l)\,d\mathbf{w}(t), \quad d\mathbf{w}d\mathbf{w}' = \mathbf{I}\,dt,$$

to an equivalent homoscedastic one:

$$d\mathbf{x} = \gamma(\mathbf{x})\,dt + d\mathbf{w}(t).$$

Since Ito's Lemma gives us

$$\begin{aligned} d\mathbf{x}(r,l) &= \nabla_{rl}\mathbf{x} + \frac{1}{2}\mathrm{tr}(\eta\eta'\nabla_{rl}\nabla'_{rl})\mathbf{x}\,dt \\ &= \left[\frac{1}{2}\mathrm{tr}(\eta\eta'\nabla_{rl}\nabla'_{rl})\mathbf{x} + \nabla_{rl}\mathbf{x}\beta\right]dt + \nabla_{rl}\mathbf{x}\eta\,d\mathbf{w}, \end{aligned}$$

if we choose $\mathbf{x}(r,l)$ such that $\nabla_{rl}\mathbf{x} = \eta^{-1}$ the result will be the homoscedastic stochastic differential system

$$d\mathbf{x} = \eta^{-1}\left[\beta + \frac{1}{2}\eta\,\mathrm{tr}(\eta\eta'\nabla_{rl}\nabla'_{rl})\mathbf{x}\right]dt + d\mathbf{w}.$$

The particular Brennan-Schwartz process used in this study was[50]

$$\begin{pmatrix} dr \\ dl \end{pmatrix} = \beta\, dt + \eta\, d\mathbf{w},$$

$$\beta = \begin{pmatrix} \alpha(l-r) \\ l\,(\sigma_l^2 + \sigma_l \lambda_l + l - r) \end{pmatrix}, \quad \eta = \begin{pmatrix} \sigma_r r & 0 \\ 0 & \sigma_l l \end{pmatrix} \mathbf{G}, \quad \mathbf{G}\mathbf{G}' = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix},$$

where σ_r and σ_l are constant. The process for r is a mean reverting process, reverting to a changing mean: the consol rate l. The drift term for the l process was derived in the theory section using the fact that the consol bond value is a known function of the consol yield. In addition, I assume that λ_l is constant.[51]

As mentioned in the previous section, if we want to transform this process to an equivalent homoscedastic form, we need to choose alternate variables $\mathbf{x}$ which

[50] Other processes were tried before this one was decided on. I had to reject using

$$\beta = \begin{pmatrix} a_1 + b_1(l-r) \\ l\,(a_2 + b_2 l + c_2 r) \end{pmatrix}$$

which is the drift of the process used by Brennan and Schwartz. There was almost perfect correlation between the estimates of a_1 and b_1 and between a_2, b_2 and c_2. This high degree of correlation makes the parameter estimates meaningless, regardless of their standard errors. In order to avoid this correlation between parameter estimates it was necessary to adopt a process such as the one used in this study.

[51] As noted in the previous footnote, assuming λ_l is a linear combination of r and l – which is the next level of sophistication – would result in a problem of almost perfect correlation between parameter estimates.

satisfy $\nabla_{rl}\mathbf{x} = \eta^{-1}$. The desired alternate variables are

$$\mathbf{x} = \mathbf{G}^{-1}\begin{pmatrix}\sigma_r & 0\\ 0 & \sigma_l\end{pmatrix}^{-1}\begin{pmatrix}\ln r\\ \ln l\end{pmatrix},$$

$$\nabla_{rl}\mathbf{x} = \mathbf{G}^{-1}\begin{pmatrix}\sigma_r & 0\\ 0 & \sigma_l\end{pmatrix}^{-1}\begin{pmatrix}1/r\\ 1/l\end{pmatrix} = \eta^{-1},$$

which lead to the desired homoscedastic form

$$\begin{aligned}
\frac{1}{2}\eta\mathrm{tr}(\eta\eta'\nabla_{rl}\nabla_{rl}')\mathbf{x} &= \frac{1}{2}\eta\left(\sigma_r^2 r^2\frac{\partial^2}{\partial r^2} + 2\rho\sigma_r\sigma_l rl\frac{\partial^2}{\partial r\partial l} + \sigma_l^2 l^2\frac{\partial^2}{\partial l^2}\right)\begin{pmatrix}\ln r\\ \ln l\end{pmatrix}\\
&= \frac{1}{2}\eta\mathbf{G}^{-1}\begin{pmatrix}\sigma_r & 0\\ 0 & \sigma_l\end{pmatrix}^{-1}\begin{pmatrix}-\sigma_r^2\\ -\sigma_l^2\end{pmatrix}\\
&= -\frac{1}{2}\begin{pmatrix}\sigma_r^2 r\\ \sigma_l^2 l\end{pmatrix},
\end{aligned}$$

$$\begin{aligned}
d\mathbf{x} &= \eta^{-1}\left[\beta - \frac{1}{2}\begin{pmatrix}\sigma_r^2 r\\ \sigma_l^2 l\end{pmatrix}\right]dt + d\mathbf{w}\\
&= \mathbf{G}^{-1}\begin{pmatrix}\beta_1/\sigma_r r - \frac{1}{2}\sigma_r\\ \beta_2/\sigma_l l - \frac{1}{2}\sigma_l\end{pmatrix}dt + d\mathbf{w}\\
&= \mathbf{G}^{-1}\begin{pmatrix}\frac{\alpha}{\sigma_r}\left(\frac{l}{r} - 1\right) - \frac{1}{2}\sigma_r\\ \frac{1}{2}\sigma_l + \lambda_l\frac{1}{\sigma_l}(l - r)\end{pmatrix}dt + d\mathbf{w}.
\end{aligned}$$

To be complete, of course, we must note that r and l in the above stochastic process for $\mathbf{x}$ are to be treated as functions of $\mathbf{x}$.

$$\begin{pmatrix}r\\ l\end{pmatrix} = \exp\left(\begin{pmatrix}\sigma_r & 0\\ 0 & \sigma_l\end{pmatrix}\mathbf{G}\mathbf{x}\right)$$

Also note that I have not uniquely determined $\mathbf{x}$. There are many $\mathbf{G}$ matrices which satisfy

$$\mathbf{G}\mathbf{G}' = \begin{pmatrix}1 & \rho\\ \rho & 1\end{pmatrix},$$

but they differ only by a rotation.

B.5.2.b BRENNAN-SCHWARTZ AND SIMPLE LINEARIZATION

The simple linearization assumption of a locally constant stochastic process drift term was shown by Ananthanarayanan to be a good assumption for the particular case that he was examining. This does not mean that it is a good assumption for all models that might be used. To be sure that the assumption is also good for the Brennan-Schwartz model that we are using in this study, we should redo the calculations that we performed when examining Ananthanarayanan's stochastic process.

If we assume for the moment that the simple linearization assumption is valid for the Brennan-Schwartz model used here, then evolution of the homoscedastic system derived in the previous section can be approximated by the following discrete difference equation.

$$\int_t^{t+\Delta t} d\mathbf{x}(s)$$

$$\simeq \int_t^{t+\Delta t} \mathbf{G}^{-1} \begin{pmatrix} \frac{\alpha}{\sigma_r}\left(\frac{l(\mathbf{x}(t))}{r(\mathbf{x}(t))} - 1\right) - \frac{1}{2}\sigma_r \\ \frac{1}{2}\sigma_l + \lambda_l + \frac{1}{\sigma_l}(l(\mathbf{x}(t)) - r(\mathbf{x}(t))) \end{pmatrix} ds + \int_t^{t+\Delta t} d\mathbf{w}(s),$$

$$\Delta\mathbf{x}(t) \simeq \mathbf{G}^{-1} \begin{pmatrix} \frac{\alpha}{\sigma_r}\left(\frac{l(\mathbf{x}(t))}{r(\mathbf{x}(t))} - 1\right) - \frac{1}{2}\sigma_r \\ \frac{1}{2}\sigma_l + \lambda_l + \frac{1}{\sigma_l}(l(\mathbf{x}(t)) - r(\mathbf{x}(t))) \end{pmatrix} \Delta t + \Delta\mathbf{w}(t),$$

$$\Delta\mathbf{w}(t) \sim N\{\mathbf{0}, \mathbf{I}\Delta t\}.$$

If we use this approximation and the minimum distance function to estimate the parameters of the process – σ_r, σ_l, ρ, α and $\sigma_l^2 + \sigma_l\lambda_l$ – the results are the parameters shown in Table B.11.

But is the simple linearization assumption valid? We can make a quick check by taking the parameters estimated for the October 1970 to October 1982 period and repeating the exercise we followed in the one-dimensional example above. There will be a slight difference, however, since we are dealing here with a two-dimensional case. If we start at the point $r = 0.08$ and $l = 0.10$, our first task is identifying reasonable end of period points.

If we let δ be

$$\delta = \mathbf{G}\Delta\mathbf{x}(t) - \begin{pmatrix} \frac{\alpha}{\sigma_r}\left(\frac{l(\mathbf{x}(t))}{r(\mathbf{x}(t))} - 1\right) - \frac{1}{2}\sigma_r \\ \frac{1}{2}\sigma_l + \lambda_l + \frac{1}{\sigma_l}(l(\mathbf{x}(t)) - r(\mathbf{x}(t))) \end{pmatrix} \Delta t,$$

$$\delta \sim N\left\{\mathbf{0}, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} \Delta t\right\},$$

then the Mahalanobis distance of any end of period point can be computed by calculating

$$D_M = \sqrt{\frac{1}{\Delta t}\delta' \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} \delta}.$$

Table B.12 shows the results of this exercise for various points which have a Mahalanobis distance of 2 from the expected end of period point (calculated using the simple linearization assumption).

We see that the beginning and end of period drifts are radically different. For example, the first row tells us that the point $r = 0.0626$, $l = 0.0964$ is not an unreasonable end of period point. At that point, however, the end of period drift for the first component of $\mathbf{x}$ is 3 times what it is at the beginning of the period. For the ending point $r = 0.10$, $l = 0.1083$ it is *negative* 20% of the beginning of period drift value.

Apparently, the simple linearization assumption does not hold for the Brennan-Schwartz model parameters in Table B.11. In order to investigate this further, the parameters were reestimated using an assumption similar to the simple linearization assumption. If the process drift is in fact close to constant over the period Δt, we should see no difference in parameter estimates whether we use the beginning of period drift, end of period drift or some convex combination of the two in our estimation procedure.

$$\Delta\mathbf{x}(t) \simeq \mathbf{G}^{-1} \begin{pmatrix} \frac{\alpha}{\sigma_r}\left(\frac{l^*}{r^*} - 1\right) - \frac{1}{2}\sigma_r \\ \frac{1}{2}\sigma_l + \lambda_l + \frac{1}{\sigma_l}(l^* - r^*) \end{pmatrix} \Delta t + \Delta\mathbf{w}(t)$$

$$l^* = (1-f)l(x(t)) + fl(x(t+\Delta t)), \quad r^* = (1-f)r(x(t)) + fr(x(t+\Delta t))$$

That is, we should find the same parameter estimates no matter what value of f between 0 and 1 we use in the estimation procedure above.

As might be expected from our analysis, the parameter estimates do in fact change as f is varied. Table B.13 shows, however, that only α and $\sigma_l^2 + \sigma_l \lambda_l$ are affected. The values of σ_r, σ_l and ρ remain practically the same for the three values of f – zero, one-half and one – that were used. Since the estimation of α is not reliable when using the simple linearization method, I decided to estimate it in the next stage, along with the market price of short rate risk, λ_r. I did, however, decide that the simple linearization estimates of σ_r, σ_l and ρ were stable enough to justify their use in later stages of the study. The parameter values for the intermediate case, $f = 0.5$ were used in the rest of the study.

B.5.3 Estimation of the Price of Risk and Reversion Coefficient α

As was illustrated in the previous section, it is not possible to estimate the reversion coefficient α of the stochastic process for the short rate of interest r by analyzing a short and long rate time series. The standard deviations and correlation – σ_r, σ_l and ρ – for the joint process in r and l appear, however, to be stably estimable from these time series. For these reasons, I decided to estimate the reversion coefficient at the same time as the market price of short rate risk λ_r, which for simplicity I assume to be constant.

B.5.3.a MINIMUM DISTANCE ESTIMATOR

The only thing preventing numerical solution of the asset pricing partial differential equation is the lack of values for λ_r[52] and α. All other parameter values have been estimated. The procedure in this section, therefore, is to try various values of λ_r and α with the aim of finding the 'best-fitting' pair of values. For each pair of values tried the partial differential equation must be numerically solved for discount bond values, and theoretical coupon bond values calculated as portfolios of discount bonds. These theoretical prices are then compared to actual market prices, giving a large vector of bond pricing errors, ϵ_t, for each month of the estimation period from October 1970 to October 1982.

Assuming multivariate normality and lack of serial correlation in bond pricing errors

$$\mathrm{E}[\epsilon_t] = 0, \quad \mathrm{E}[\epsilon_s \epsilon_t'] = \begin{cases} \mathbf{S}, & s = t, \\ \mathbf{0}, & s \neq t, \end{cases}$$

we would choose the pair of λ_r and α which give the best fit between theoretical and actual bond prices by minimizing the distance function[53]

$$D = \sum_{t=1}^{T} \epsilon_t' \mathbf{S} \epsilon_t.$$

The 'best' values of λ_r and α are those resulting in the minimum value for this function.[54]

---

52 Assumed constant, as mentioned above.

53 As mentioned previously, when $\mathbf{S}$ is unknown and is estimated from the sample the distance function is asymptotically maximum likelihood. See Malinvaud (1966) and Phillips (1972).

54 I found that the estimates of λ_r and α were different if serial correlation was or

B.5.3.a.i Portfolio Formation Schemes

We cannot, however, estimate the above distance function directly by using individual bond price errors. There are two problems with such an approach. First, since the maturity of each bond decreases as we follow its time series of prices, presumably the variance of its pricing residuals also changes. This would lead to non-stationarity in the covariance matrix **S**. Second, individual bonds do not have prices spanning the entire sample period, resulting in a missing data problem. Both these problems are ameliorated by combining the bonds into portfolios according to maturity, and using these portfolios to estimate the unknown parameters, λ_r and α.

In this study three different equally-weighted portfolio formation schemes were compared. Scheme number 1 places all bonds with maturities between 0 and 1 years into portfolio 1, between 1 and 2 years into portfolio 2, and so on until the

---

was not deleted from bond pricing errors. As the main purpose of this study is the pricing of options on bonds, not simply the pricing of bonds, I decided not to deleting the serial correlation from bond pricing errors. The reason is that deletion of serial correlation from the pricing errors when estimating bond prices logically requires a similar correction of theoretical bond prices in later stages of the study. For example, if we were testing trading strategies based on differences between theoretical and market prices, as our theoretical price we would use this period's model price corrected for serial correlation by using the ***previous period's pricing error*** and an estimated correlation coefficient. In this study, however, we use the theoretical bond prices to form boundary conditions when solving the partial differential equation for option prices. In this context there is no previous period pricing error that can be used to correct the model prices for serial correlation. Since there is no way that this correction can be applied when pricing these options, I decided to ignore the correction when pricing bonds also. As a result, the model bond prices which I obtain are as close as possible to the actual market bond prices. Hopefully this improves the accuracy of the option prices obtained in later stages of the study.

tenth portfolio which contains bonds maturing in 9 to 10 years. This is the scheme used by Brennan and Schwartz and, as can be seen in Table B.14, there are some problems with scarcity of data in portfolios 8, 9 and 10. In addition, bonds with maturities greater than 10 years are totally ignored.

It is probably not a good idea to ignore the higher maturity bonds in this study, as these are typically exactly the bonds which options are written on. In order to address this problem an additional portfolio containing bonds of maturities between 10 and 20 years was added in scheme number 2. This still does not correct the scarcity of data in the 7 to 10 year maturity range, however, so in portfolio formation scheme number 3, the three portfolios previously containing the 7 through 10 year maturity bonds were combined into a single portfolio. In addition, because of a relative abundance of data in the 10 to 20 year portfolio in scheme 2, in scheme number 3 it was split into two portfolios.

For the reasons given above, I feel that the third portfolio formation scheme is the most desirable. The other two have been included here to examine how sensitive the estimates of λ_r and α are to the particular portfolio scheme used.

B.5.3.a.ii Covariance Matrix Assumptions

The formation of bond portfolios avoids the problem of having a non-stationary covariance matrix, but still leaves the problem of actually estimating the covariance matrix. In order to test the sensitivity of the estimates of λ_r and α, once again three different approaches were taken. The first approach corresponds to an ordinary least-squares regression (OLS) of theoretical portfolio prices on observed prices and is accomplished by assuming the covariance matrix is the identity matrix.

$$\mathbf{S} = \mathbf{I}, \quad \text{(OLS)}$$

The second approach applies an *ad hoc* heteroscedacity adjustment with the assumption that the variance of a portfolio's pricing errors is proportional to the average maturity of the portfolio.

$$\mathbf{S} = \mathbf{H}, \quad \mathbf{H} = \mathrm{diag}(h_i), \quad h_i = \text{average maturity of portfolio } i. \text{ (HETERO)}$$

Finally, the third approach applies generalized least-squares methodology, approximating the covariance matrix by using the portfolio pricing errors.

$$\mathbf{S} = \frac{1}{T}\sum_{t=1}^{T} \epsilon_t \epsilon_t', \quad \text{(GLS)}$$

This last method is asymptotically maximum likelihood.

B.5.3.b PARAMETER ESTIMATES

Table B.15 shows a selection of the values of values of λ_r and α tested in the two-dimensional non-linear search for the minimums of the 9 different distance functions. Since there were only two unknowns, the search procedure was begun by selecting a large number of initial (λ_r, α) points with the intention of sketching out the overall shape of the surface that was being searched.

The first thing to note from Table B.15 is that the market price of short term risk is positive, contrary to the expected negative value found by Brennan and Schwartz. This is independent of the portfolio formation scheme and the covariance matrix assumption.[55] This is unexpected, but not impossible.

The theory section tells us that the expected return μ of a discount bond $\delta(r, l, \tau)$ in the Brennan-Schwartz model satisfies

$$\mu - r = \mathbf{s}\,\lambda = \frac{1}{\delta}[\delta_r \sigma_r r \lambda_r + \delta_l \sigma_l l \lambda_l].$$

We would expect higher risk to command a higher return, and since the choice of the form of the Brennan-Schwartz model used here links higher interest rates with higher interest rate variances and hence higher risk, we would expect higher returns at higher levels of interest rates. Since we would expect the price of a discount bond to decrease with increasing r, that is, $\delta_r < 0$, the only way to have μ increase as r increases is to have λ_r negative.

This argument is quite persuasive, but the numerical solution points out that it is also quite incorrect. The price of a discount bond does not always decrease as we increase the short rate of interest r. In some areas of the r and l grid we find

[55] Also independent of whether the serial correlation of bond pricing errors is taken into account or not.

$\delta_r > 0$. The key to understanding why the numerical solution is correct and the argument incorrect is the fact that while r is a yield on a discount bond, l is not. It is a yield on an asset which pays a perpetual continuous coupon. When we take the partial derivative of the discount bond price with respect to r, we are doing so while holding l constant. This is *not* the same as holding the long rate of interest constant, but should be imagined instead as keeping constant a weighted average of *all* rates, from the shortest to the longest.

As a result, when we increase the short rate the only way to keep this weighted average constant is to decrease some other rate or rates. When we increase the short rate the value of consol coupon payments relatively close to the present decrease, and the value of other payments further in the future must increase to keep the value of the consol constant. But, this means that as we increase the short rate r, some longer term discount bond must increase in value. That is, for shorter maturity discount bonds we, of course, find δ_r negative, but as we increase the time to maturity we eventually find δ_r positive. As is clear, if try to use the argument above to establish the sign of λ_r, this ambiguity in the sign of δ_r leads to an ambiguity in the sign of λ_r.

Once the initial sketch of the surface was finished, the most promising areas were investigated for a minimum by assuming that the surface could be treated as locally parabolic.[56] The method consisted of finding the minimum point of this fit-

[56] The exact form used was

$$z(r,l) \simeq a_{00} + a_{10}r + a_{01}l + a_{20}r^2 + a_{11}rl + a_{02}l^2$$

where z is the height of the surface.

ted parabolic surface, evaluating the distance function at this new point, refitting the parabolic surface and so on until convergence was reached. It turned out that the minimum was in a long, very narrow trough running through the surface.[57] As a result, the distance function value for the final parameter estimates differs only slightly from its value farther along the trough for different parameter values.[58] Reporting the standard errors of these estimates would, therefore, be misleading, as the trough indicates a high degree of correlation between the parameter estimates.

What is more promising is a comparison of the optimal values of λ_r and α for the three different portfolio formation schemes and three different covariance matrix assumptions. As shown in Table B.16 the estimates appear relatively insensitive to these different assumptions. The optimal parameter values for all nine of these treatments lie within a factor of two of each other. For the rest of the study, I decided to use the estimates derived from using the third portfolio scheme and the GLS covariance matrix assumption, namely, $\lambda_r = 0.260$ and $\alpha = 0.558$.

[57] This is apparently not uncommon. See Marquardt (1963).

[58] This is not the reason for λ_r being positive. The trough lies almost completely in the quadrant where both λ_r and α are positive.

B.6 PRICING MODEL ERRORS

The numerical solution of the asset pricing partial differential equation only supplies asset prices for the actual grid points used. For points lying between these grid points some sort of interpolation must be used. When interpolating bond prices, I used the simplest interpolation method: linear interpolation. This is the same method as was used successfully by Brennan and Schwartz, and is successful because of the small curvatures present in bond prices as a function of r and l. When options are valued, however, the curvatures are much greater, and linear interpolation may produce a poor fit. To investigate this possibility, I compared three different methods of interpolation when pricing options, linear interpolation, cubic spline interpolation and a form of quadratic interpolation.

B.6.1 Linear Interpolation

By far the easiest interpolation method is the one used by Brennan and Schwartz to interpolate bond prices: linear interpolation. Quite simply, if we have prices at the four points (r_m, l_n), $m = i, i+1$, $n = j, j+1$, $r_{i+1} - r_i = \Delta r$, $l_{j+1} - l_j = \Delta l$, then we can linearly interpolate the price at any point inside this rectangle as

$$P(r,l) = (1-f)(1-g)P_{ij} + (1-f)gP_{i,j+1} + f(1-g)P_{i+1,j} + fgP_{i+1,j+1},$$

$$r_i \le r \le r_{i+1}, \quad l_j \le l \le l_{j+1},$$

$$P_{ij} = P(r_i, l_j),$$

$$0 \le f = \frac{r-(r_i)}{\Delta r} \le 1, \quad 0 \le g = \frac{l-(l_j)}{\Delta l} \le 1.$$

This is the same as using a bilinear surface[59] to interpolate over the rectangle. In order to estimate the first partial derivatives at the point (r,l), we need several other data points.[60]

$$\begin{aligned} P_r(r,l) &\simeq \frac{H_r}{2\Delta r}[(1-f)(1-g)P_{ij} + (1-f)gP_{i,j+1} + f(1-g)P_{i+1,j} + fgP_{i+1,j+1}] \\ &= \frac{1}{2\Delta r}[(1-f)(1-g)(P_{i+1,j} - P_{i-1,j}) + (1-f)g(P_{i+1,j+1} - P_{i-1,j+1}) \\ &\qquad + f(1-g)(P_{i+2,j} - P_{i,j}) + fg(P_{i+2,j+1} - P_{i,j+1})], \end{aligned}$$

which simplifies to the desired expression at the nodes (r_i, l_j)

$$P_r(r_i, l_j) \simeq \frac{H_r}{2\Delta r}P_{ij} = \frac{P_{i+1,j} - P_{i-1,j}}{2\Delta r}.$$

The partial derivative in the l direction is approximated similarly. This is the simplest method available, and it will be successful if the curvature of the surface being interpolated is not too great.

---

59 A bilinear surface satisfies the equation $z(x,y) = a_{00} + a_{10}x + a_{11}xy + a_{01}y$ and is so named because a cross-section of this surface in either the x or y direction is a straight line.

60 The points in the r direction are assumed to be separated by a distance Δr, and those in the l direction by Δl.

B.6.2 Cubic Spline Interpolation

Probably the most popular form of curve interpolation is interpolation using cubic splines. Basically, if we start with a one-dimensional grid of points, r_i, and prices at these points, P_i, the interpolation is done by fitting a separate cubic function between each pair of points. That is, between r_i and r_{i+1} we would fit the cubic function

$$P(r) = a_i + b_i r + c_i r^2 + d_i r^3, \quad r_i \le r \le r_{i+1}.$$

Actually, the function that is fitted is the cubic spline

$$\begin{aligned} S_i(r) &= s_i \left(\frac{r_{i+1} - r}{\Delta r}\right)^3 + s_{i+1} \left(\frac{r - r_i}{\Delta r}\right)^3 \\ &\quad + (P_i - s_i)\left(\frac{r_{i+1} - r}{\Delta r}\right) + (P_{i+1} - s_{i+1})\left(\frac{r - r_i}{\Delta r}\right), \end{aligned}$$

$$s_i = \frac{\Delta r^2}{6} S_i''(r_i), \quad r_i \le r \le r_{i+1}.$$

By 'fitting' a spline between adjacent points, I mean that the first and second derivatives of the entire curve are made continuous.

$$P_i = S_i(r_i) = S_{i-1}(r_i), \quad S_i'(r_i) = S_{i-1}'(r_i), \quad S_i''(r_i) = S_{i-1}''(r_i).$$

This is accomplished by finding the appropriate values for the second derivatives, S_i'', at the grid points.[61]

This method is perfect for interpolating between a large number of points on a line, but of course the computation required increases as we increase the number

[61] See Vemuri and Karplus (1981) regarding the use of cubic splines for interpolation.

of points. We could use a two dimensional version of the cubic spline interpolation method and compute the interpolating splines for the entire r, l grid at each time step, but since we need to interpolate at most one point at each time step this would be an expensive proposition. Instead, I decided to use one-dimensional cubic spline interpolation in the r direction followed by one-dimensional cubic spline interpolation in the l dimension. For each interpolation, in order to reduce computational requirements, I would use only a four by four grid of data points.

B.6.3 Quadratic Interpolation

As an alternative to cubic spline interpolation I wanted to try interpolation with a quadratic curve. In keeping with the cubic method, I also wanted the method to only use a four by four grid of data points.

The quadratic interpolating polynomial used was

$$Q(x) = a_0 + a_1(x - r) + a_2(x - r)^2,$$

where what is desired is the interpolation function value and first derivative at $x = r$, that is, $P(r) = Q(r) = a_0$ and $P'(r) = Q'(r) = a_1$. I also wanted the quadratic to fit perfectly at the points r_i and r_{i+1}.

$$P_i = Q(r_i) = a_0 + a_1(r_i - r) + a_2(r_i - r)^2 = a_0 - a_1 f\Delta r + a_2 f^2 \Delta r^2$$

$$\begin{aligned} P_{i+1} &= Q(r_{i+1}) = a_0 + a_1(r_{i+1} - r) + a_2(r_{i+1} - r)^2 \\ &= a_0 - a_1(1-f)\Delta r + a_2(1-f)^2\Delta r^2 \end{aligned}$$

$$f = \frac{r - r_i}{\Delta r}$$

These two equations can be solved for the two needed unknowns $P(r) = a_0$ and $P'(r) = a_1$

$$P(r) = a_0 = P_i(1-f) + P_{i+1}f - a_2 f(1-f)\Delta r^2,$$

$$P'(r) = a_1 = \frac{P_{i+1} - P_i}{\Delta r} - a_2(1-2f)\Delta r,$$

leaving us with only a_2 to estimate.

If the prices at the four points r_{i-1} through r_{i+2} really were described by a quadratic function, then the value of $a_2 = \frac{1}{2}P''$ would be exactly $a_2 = (P_{i-1} - P_i - P_{i+1} + P_{i+2})/4\Delta r^2$. The closer the curve is to being quadratic (ie. having a constant second derivative), the better this approximation will be. Using this estimate of the second derivative gives us[62]

$$P(r) = P_i(1-f) + P_{i+1}f - \frac{1}{4}f(1-f)(P_{i-1} - P_i - P_{i+1} + P_{i+2}),$$

$$P'(r) = \frac{1}{\Delta r}\left[P_{i+1} - P_i - \frac{1}{4}(1-2f)(P_{i-1} - P_i - P_{i+1} + P_{i+2})\right].$$

If the curve is linear in this area, then $P_{i-1} - P_i - P_{i+1} + P_{i+2}$ is zero and the quadratic interpolation value is the same as the linear one.

This one-dimensional interpolation method was extended to two dimensions us-

[62] Once again notice that the linear interpolation functions $P(r) = P_i(1-f) + P_{i+1}f$ and $P'(r) = (P_{i+1} - P_i)/\Delta r$ are contained in these quadratic approximations. As in the cubic spline case, the quadratic interpolation value is the linear interpolation value plus a 'correction' term due to the second derivative.

ing a four by four grid of data points in the same manner used for cubic spline interpolation.

B.6.4 Bond Option Pricing

As mentioned in the pricing theory section, any asset value which is a function of r, l and maturity only will be described by the same asset pricing partial differential equation. The differences between the prices of these assets arises from their different boundary conditions. Since we will be valuing American calls and puts, the boundary conditions that need to be imposed are[63]

$$C(r, l, \tau; K) \geq \max(0, B(r, l, \tau + \tau_B; c) - K),$$

with equality at maturity.

Of course, in our numerical solution method we cannot impose this boundary condition at all points in time, simply because we are only solving for option values at discrete time points. The best that we can do is to approximate the option holder's right of exercise at any time by the right to exercise at certain discrete points in time, namely, the points where we calculate the partial differential equation solution.

[63] I only consider a call option on a specific bond here as the treatment of put options is analogous. There are no 'fixed maturity' bond options traded, only 'fixed maturity' treasury bill options.

At each time step, therefore, we first solve the partial differential equation ignoring the right to exercise. This gives us a preliminary option value $\mathbf{C}^{+}(\tau_K)$. We then allow the option holder to exercise his option, which gives us the option value we desire.

$$C(r_i, l_j, \tau; K) = \max[C^{+}(r_i, l_j, \tau; K), B(r_i, l_j, \tau + \tau_B; c) - K]$$

This means, of course, that we need a complete grid of theoretical bond prices at each time step of the solution.

Aside from the interpolation method to be used there are several other details that needed to be addressed. When close to maturity, option price functions have a zone of high curvature in the region where the bond price equals the exercise price (the region where the option trades at the money). In order to hopefully avoid problems arising from too large a spacing between grid points, I decided to reduce the time step size to one day as compared to the two weeks that I used when valuing bonds. In addition, since the option values at extreme points on the grid should be close to either zero or the value of the option when exercised, $B(r_i, l_j, \tau + \tau_B; c) - K$, I tested to see whether the grid's range of interest rates could be reduced somewhat. While the numerical solution was unaffected when I reduced the maximum l value from 0.50 to 0.25, I was not able to do this with the r dimension. The resulting grid of r from 0.00 to 0.50 and l from 0.00 to 0.25 was used in pricing all options.

B.6.4.a *CHOICE OF INTERPOLATION METHOD*

How then do we judge whether one interpolation method is better than another? One test would be a comparison of pricing errors for the three different methods, such as is presented in Table B.17. As expected, the linear interpolation method performs poorest, with a root mean square error of $1.70. Both the cubic and quadratic methods do substantially better, ending up with root mean square errors of $0.67 and $0.66, respectively.

Actually, the improvement is even better than it would seem from Table B.17, since all three methods should produce essentially the same result when interpolating only a small distance from available data points. For example, if we have data points at $r = 0.10$ and 0.11, then we would expect the methods to differ more when interpolating to the point point halfway between these points, $r = 0.105$, than when interpolating to $r = 0.101$. There is a great potential for improvement in some of the pricing errors, and practically no possibility of improvement in others. Including the cases with little possibility of improvement decreases the apparent reduction in size of interpolation errors.[64] Judging from Table B.17 alone, there does not seem to be any reason to prefer cubic over quadratic interpolation.

We have been mentioning all along, however, that we also need to be concerned with the first partial derivatives that come from our numerical solution of asset prices. These are needed in order to form arbitrage portfolios. Table B.18 shows

[64] For example, imagine the much simplified situation where half of the theoretical prices are not interpolated, but still give an error of $0.50, and the other half are interpolated and result in errors of $1.00 when linear interpolation is used. The average error is then $0.75. The $0.50 errors are not affected by a change from linear to quadratic interpolation, but say that the $1.00 errors drop to $0.70. This is really a reduction of $0.30, but the average error only drops $0.15 to $0.60.

that the partial derivative of option prices with respect to r is not likely to cause much difficulty for interpolation methods. Only the results from the linear interpolation method are given, since the other two methods produce essentially the same results. As the last line of Table B.18 shows, the average change in bond option price from one r grid to the next – a separation of 0.01 – results in an average option price drop of $0.021. The maximum price drop was $0.71 and the maximum rise $0.60. Calls and puts have not been separated, as the results are similar for both.

Notice that the derivative can be either positive or negative for both calls and puts whereas we might at first expect only negative slopes for calls and positive slopes for puts. This follows from the discussion above about the change in bond prices when r is changed while holding l constant.[65] As we saw, bond prices could either rise or fall as r increased. Consequently, the same can be said of option values.

These figures contrast starkly with those in Tables B.19 and B.20, showing the partial derivative of bond call and put options with respect to l for the three different interpolation methods. The average drop in call option price for a 0.01 increase in l is $4.30 when linear interpolation is used – not anywhere near the $0.021 change that we saw in the r direction – with a maximum drop of $10.80 and a minimum drop of $0.10. The figures are similar for put options, except that prices now rise when l increases. The average rise is $4.50 for a 0.01 increase in l, with a maximum rise of $10.50 and a minimum of $0.20. In general, all three

[65] See section B.5.3.b.

methods still produce results that differ only slightly.

There is one problem, however, which shows up in the first row of both tables. Unlike the effect of changes in r on bond prices, an increase in l while holding r constant should unambiguously cause a decrease in bond prices. This should cause a decrease in call values and an increase in put values. Unfortunately, we see in the first row of Table B.19 that both the cubic and quadratic methods produce an *increase* in some call option prices when l increases. The first row of Table B.20 shows the same story for put options. Both the cubic and quadratic methods produce a *decrease* in some values. As this occurs only when interpolating options with one month or less to maturity, we can guess that the problem lies with the high curvature existing in the at the money region of the option price function. As the maturity of the option decreases, this curvature becomes more and more localized and abrupt, until at maturity we find a discontinuity in the first derivative. Apparently, the abruptness of this at the money region for very short maturity options causes trouble for the cubic and quadratic interpolation methods.

In an attempt to correct this problem with the l partial derivative, I decided that if cubic or quadratic interpolation resulted in a option value below or above the two bracketing option values, then linear interpolation would be used. Similarly, if the signs of the partial differential with respect to l differed at the two bracketing points, once again linear interpolation would be used. The results of this adjustment are shown in Tables B.21 and B.22. As can be seen, the adjustment works for quadratic interpolation but not for cubic spline interpolation. In fact, the adjusted partial derivatives for the cubic method are worse than the unadjusted values. For this reason, I decided to use the quadratic interpolation method in this study.

B.6.4.b BOND OPTION PRICING ERRORS

The net result of the entire process is displayed in Table B.23. As can be seen, on average the Brennan-Schwartz model leads to an average overpricing of \$0.33 for the call option sample and \$0.30 for the puts. There may be a slight rise in root mean square errors as time to maturity of the options increases, but there does not appear to be any pronounced pattern to the errors. In Table B.24 the same data are shown by the ratio of bond price to option exercise price (in, at or out of the money). The pricing errors seem to be slightly greater for in the money options versus out of the money options, but once again, the pattern is not very pronounced.

Since we are pricing options on long term bonds, assuming a constant short term interest rate – as in the Black-Scholes model – may not have much of an effect on option values. At least, we might believe this before looking at the Black-Scholes pricing errors in Tables B.25 and B.26.[66] [67] Making this assumption increases the average call option pricing error from \$0.33 to \$0.57, almost double what it was using the Brennan-Schwartz model.[68] Similarly, the average put option pricing error rises to \$0.55.

As mentioned in the section on parameter estimation, the parameters that we have estimated may be non-stationary. If this is the case, it may be that we are using incorrect parameter estimates to compute theoretical prices. As has been shown many times with options on stocks, option prices are quite sensitive to the variance estimate used. To test this possibility, the variance and covariance parameters σ_r, σ_l and ρ were estimated for the option testing period, October 1982 to October 1983, and the options repriced with the new estimates. These 'in sample' variance estimates were $\sigma_r = 0.215$, $\sigma_l = 0.132$ and $\rho = 0.193$. Notice that while both σ_r and ρ change from their 'out of sample' estimates of 0.448 and 0.512, respectively, σ_l remains about the same.[69] As a result, we see in Tables B.27 and B.28 that the average Brennan-Schartz pricing errors decrease by half to \$0.15 for calls and \$0.14 for puts. Therefore, a good deal of the pricing error seems to be due to the use of 'out of sample' parameter estimates, not the solution method.

---

66 The variance rate of the consol rate, σ_l, was used as the required Black-Scholes variance estimate as in Brennan and Schwartz (1983a). Theoretically, the correct variance to use would be the variance of the price of the underlying bond. Since bond options are typically written on recently issued bonds, however, in general there is no past time series of prices to use in computing the variance.

67 The method used in this study to solve the Black-Scholes model solves over a grid of r versus bond price.

68 When using a two sample t-test, this difference is highly significant.

69 Since σ_l was used as the Black-Scholes variance estimate, no significant change occurred in the Black-Scholes pricing errors when using the 'in sample' variance estimates.

B.6.5 Treasury Bill Option Pricing

The procedure for pricing treasury bill options is similar to the procedure used in pricing discount bonds and bond options. Once again, however, the boundary conditions are different. The underlying security of traded treasury bill options is not fixed, as it is for bond options. There, an option was written on a specific bond outstanding, say, the 12% bond maturing august 15, 2013. With a treasury bill option, the security which must be delivered upon exercise is always a 13-week treasury bill.[70] The boundary conditions imposed, therefore, are those where the underlying security is a discount bond with a fixed time to maturity, $\overline{\tau}$,

$$C(r, l, \tau; K) \geq \max[0, \delta(r, l, \overline{\tau}) - K],$$

with equality at maturity. As with the bond options, we only impose this boundary condition at certain discrete points in time, namely, the points where we calculate the partial differential equation solution.

The above boundary conditions are appropriate as far as theoretical option pricing is concerned, but in order to conform to reality we must make a small adjustment. When a call option is actually exercised, the parties of the contract do not have to settle up until several days after the exercise day. When we were pricing bond options, the slight delay of two business days could be ignored with little impact on option pricing. The reason is that the exercise price of the option includes accrued interest on the bond up to and including the settlement date. Therefore, any interest gathered by putting the exercise price amount in the bank for those

[70] This is discussed in section B.3.2.b.

few days must be used to pay the additional accrued interest on the bond.[71] The situation is different with treasury bill options, however, as is illustrated in the next section.

B.6.5.a TREASURY BILL OPTION SETTLEMENT ADJUSTMENT

We were able to ignore the delay between exercise and settlement dates when pricing bond options because the option exercise price included accrued interest on the underlying bond. Since interest does not accrue to the holder of a treasury bill, however, we find that a small adjustment must be made to the pricing method when pricing treasury bill options. Consider, for example, what happens when a treasury bill call option is exercised. In order to lock in his settlement date obligations the writer of the option must immediately buy a treasury bill which will have 13 weeks to maturity on the settlement date. His immediate cost at exercise, therefore, is the current cost of the underlying treasury bill, just as it would be if settlement occurred on the same day as exercise.

On the other hand, the person exercising the option can place the exercise amount in the bank for those few days and earn the riskless rate on it. The immediate cost to him, therefore, is the discounted exercise amount, discounted by the riskless

[71] The only benefit that would result would be from differences in the bank rate and the bond accrual rate. Since options are typically written only on recently issued bonds, this difference sould be minimal.

rate of interest r. Since the exercise and settlement dates are between 6 and 10 days apart, the discounted exercise price may be significantly different from the undiscounted price.

The market is well aware of this detail, as can be seen from market option prices on the date of maturity. At maturity, we expect the option price to be either zero or the treasury bill price minus exercise price. For example, our sample contains a treasury bill option price on the maturity date of December 17, 1982. The call exercise price was 92, which translates to \$97.9778 per \$100 principle value, and the option premium was 0.20, which corresponds to a price of \$0.05 per \$100.00 principle value. On the same day, the underlying treasury bill was quoted at a bid discount of 7.81 (price of \$97.89564) and an ask discount of 7.71 (\$97.92258). Using that day's yield on a 27 day treasury bill as a proxy for the riskless rate gives us a bid yield of 0.076256 (bid discount 7.50) and an ask yield of 0.075033 (discount 7.38). There are 6 days between the exercise and settlement dates.

We can now calculate the return to the following two arbitrage strategies: (a) Buy one call option, short one underlying treasury bill and exercise the option. Place the *discounted* exercise amount in an account bearing the riskless rate of interest. On the settlement date, pay the exercise amount and close out the short position in the treasury bill. (b) Write one call option. Since the option has positive market value, assume that it is exercised, buy one underlying treasury bill and take out a loan at the riskless rate for the discounted exercise amount. On the settlement date, deliver the underlying treasury bill and pay off the loan.

The first strategy results in a loss of

$-$(call price) $-$ (disc. exercise price) $+$ (underlying bond price)

$$= -0.05 - 97.97778 \times \exp(-0.075033 \times 6/365) + 97.89564$$

$$= -0.01137.$$

Similarly, the second strategy results in a loss of

$+$(call price) $+$ (disc. exercise price) $-$ (underlying bond price)

$$= +0.05 + 97.97778 \times \exp(-0.076256 \times 6/365) + 97.92258$$

$$= -0.01754.$$

If we had not discounted the exercise price there would apparently have been arbitrage profits to be made.

B.6.5.b TREASURY BILL OPTION PRICING ERRORS

Brennan-Schwartz treasury bill pricing errors are shown in Tables B.29 and B.30 by time to maturity and the ratio of treasury bill price to option exercise price, respectively. The model overprices treasury bill options by an average of $0.033 for calls and $0.044 for puts. Once again, there does not seem to be any pattern to the errors. As was mentioned above, it is unfair to use the Black-Scholes method for pricing options on treasury bills. The assumption of a constant short rate of interest might have had empirical validity when pricing options on long term bonds, but is totally incorrect when pricing treasury bills.

When we reprice our sample of treasury bill options using 'in sample' variance estimates, as shown in Tables B.31 and B.32, the model no longer overprices options. Instead, it now underprices call options by \$0.025 on average and puts by \$0.015.

B.7 ARBITRAGE TESTS

In the pricing theory section we showed that the same model used to price assets also tells us how to form an arbitrage portfolio.[72] If our modelling of reality is correct, then these theoretical arbitrage portfolios should allow us to take advantage of any arbitrage opportunities – assuming any exist. If our modelling is incorrect, then what appear to be arbitrage opportunities will in reality be due to theoretical mispricings, and should not lead to trading profits.

The arbitrage procedures followed in this study consist of initially buying or writing one option contract. The option is bought if it is 'underpriced' by the market (i.e. market price less than theoretical price), and written if it is 'overpriced'. That is, if ν_1 is the number of option contracts bought, $\nu_1 = +1$ if the option is 'underpriced' and $\nu_1 = -1$ if 'overpriced'. We then hedge this position in the option by going long or short in two other assets. The most natural choice for one of these assets is, of course, the option's underlying asset. The other asset chosen was a 5 year bond.[73] According to the formula derived in section B.2.2, the quantity of assets that should be bought to hedge the option position is given

[72] See section B.2.2.

[73] The second asset cannot be too close in characteristics to the first asset. If it is too close, then the hedging procedure may choose to take extreme offsetting positions in the two hedging assets. The five year maturity bond is sufficiently different from both the long term bonds underlying bond options (typically of 20 to 30 year maturity) and the treasury bills underlying treasury bill options so that it can be used in all the hedging portfolios formed in this study.

by

$$\begin{pmatrix} \nu_2 \\ \nu_3 \end{pmatrix} = -\nu_1 \begin{pmatrix} \frac{\partial z_2}{\partial r} & \frac{\partial z_3}{\partial r} \\ \frac{\partial z_2}{\partial l} & \frac{\partial z_3}{\partial l} \end{pmatrix}^{-1} \begin{pmatrix} \frac{\partial z_1}{\partial r} \\ \frac{\partial z_1}{\partial l} \end{pmatrix},$$

where ν_1 and z_1 are the quantity bought and (theoretical) price of the option, ν_2 and z_2 are quantity bought and (market) price of the underlying asset and ν_3 and z_3 are the quantity bought and (market) price of the 5 year bond.[74] The zero-investment arbitrage portfolio is completed by investing $-(\nu_1 z_1 + \nu_2 z_2 + \nu_3 z_3)$ in the riskless asset.

Two different arbitrage trading strategies were used. In strategy number 1, each of the 248 options is treated separately. The arbitrage portfolio is formed at each available option price observation and the dollar arbitrage returns cumulated for each option. These cumulative dollar arbitrage returns are then averaged over the 248 options. In the second trading strategy returns are not cumulated. The average of the roughly 3500 dollar arbitrage returns is given.

[74] The partial derivatives were calculated by fitting a quadratic curve to the price grids given by the numerical solution procedure (see section B.6.3). When pricing bond options, I used the value for l that is given by assuming both that the proxy for r is correct and that the theory correctly prices the underlying bond. The grid of theoretical bond prices is then examined for the value of l which is consistent with both of these assumptions. A similar procedure was used when pricing treasury bill options, except that the l proxy was assumed correct and the r value was found by examining the grid of theoretical treasury bill prices.

B.7.1 Bond Option Arbitrage Tests

As explained above, trading strategy number 1 calculates cumulative dollar arbitrage profits, cumulated over the life of each option. When applying this strategy, trading was free of any commissions on trading, and bonds were bought and sold at the middle of the bid-ask spread.[75] Over the 248 options in the sample, the average cumulative dollar arbitrage return was \$1.297 when the Brennan-Schwartz model was used and \$0.790 for the Black-Scholes model.[76] Both of these average returns are significant, with t-statistics of 9.8 and 6.3, respectively.[77] The difference between the two models, \$0.51, is also significant, with a t-statistic of 2.8.

In Table B.33, however, we see that these apparent arbitrage profits are not large enough to cover reasonable transactions costs which would be incurred in implementing the strategy. The first line of the table shows that simply buying and selling bonds at the bid and ask prices instead of at the middle of the bid-ask spread reduces these profits from \$1.297 to \$0.771 (Brennan-Schwartz model).

[75] In order to weed out any incorrect data, I ignored an option price observation if the ratio of the theoretical and market prices differed by a factor of 10. The ratio of 10 was chosen as it seemed unlikely that correct data would result in market and theoretical prices differing by a factor of 10. Only a few observations were eliminated as a result of this test.

[76] As the difference between results for calls and puts is slight, only the aggregated results are provided here.

[77] To check for outliers, scatter plots of the arbitrage gain versus days to maturity and holding period were made. The plots showed no apparent outliers and also indicated that there was little, if any, relation between the size of the arbitrage return and the number of days to maturity. There was a slight relation between the size of arbitrage return and the length of the holding period, though doubling the holding period did not appear to double the return.

Adding bond commissions of 1/4% results in very significant losses from the arbitrage strategy.[78]

The picture does not change if we use our 'in sample' variance estimates. Without charging commissions, and buying and selling bonds at the middle of the bid-ask spread we find an apparent cumulative dollar arbitrage return of $1.678, with a t-statistic of 12.5. Once again, however, Table B.34 shows us that this is not large enough to cover reasonable transaction costs.

When we turn to the second trading strategy – where returns are not cumulated – we see in Table B.35 that once more there appear to be arbitrage opportunities when no transaction costs are charged. The average return in the first line of the table is $0.092, with a highly significant t-statistic of 11.1. In computing the next lines of the table, we put restrictions on the formation of the arbitrage portfolio. Instead of arbitraging whenever we had an option price observation, we instead only formed the arbitrage portfolio when the difference between the theoretical and market prices of the option differed by at least the amount of the filter.

This approach recognizes that no matter what model is used, some level of mispricing is inevitable. Because of this, it is quite possible that small differences between theoretical and market prices do not indicate real arbitrage opportunities, only theoretical mispricings. By imposing a filter, hopefully we improve the

[78] The arbitrage profits are more sensitive to bond commissions than option commissions because the arbitrage trading strategy results in much more bond trading than option trading when portfolio rebalancing is needed.

chances of recognizing real arbitrage opportunities.

As we see from Table B.35, as the filter increases so does the average arbitrage return. The significance level of the returns stays approximately constant, however, due to the decrease of sample size. At all the filter levels the average Brennan-Schwartz return is significantly larger than the Black-Scholes return. Once again, using 'in sample' variances increases the Brennan-Schwartz arbitrage returns, as shown in Table B.36. In all three of these models – Brennan-Schwartz, Black-Scholes and Brennan-Schwartz ('in sample' variance) – Tables B.37 and B.38 show that these apparent arbitrage profits disappear when we buy and sell bonds at their bid and ask prices and charge reasonable commissions.[79]

B.7.2 Treasury Bill Option Arbitrage Tests

The same tests were applied to the treasury bill option sample as were used in the previous section. Because of the more limited sample size, however, the results were much less significant. For example, the cumulative dollar return for trading strategy 1 averaged over the 31 available treasury bill options (no commissions charged, treasury bills and bonds bought and sold at the middle of the bid-ask spread) is \$0.2070 with a t-statistic of 1.3 when using the Brennan-Schwartz model,[80] increasing to \$0.2753 with a t-statistic of 1.4 when 'in sample'

79. A filter of zero is used in Tables B.37 and B.38.

80 As mentioned above, the Black-Scholes model is not suited to valuing treasury

variances are used. Tables B.39 and B.40 show, however, that these apparent arbitrage returns are not large enough to cover the transaction costs incurred by the arbitrage strategy.

We see the same pattern when looking at the results of trading strategy number 2 (returns not cumulated). Tables B.41 and B.42 show that using a filter to restrict formation of the arbitrage portfolio does increase the apparent arbitrage profits. The inclusion of transaction costs in Tables B.43 and B.44, however, shows that these apparent arbitrage profits disappear when reasonable costs must be borne.[81]

---

bill options.

[81] The values in Tables B.43 and B.44 are computed using a filter of zero.

B.8 SUGGESTIONS FOR FURTHER RESEARCH

As the involved procedure of the Brennan-Schwartz model was being performed, I noted down possible improvements in the method that might bring about an increase in pricing accuracy. These suggestions cover all aspects of the procedure.

B.8.1 Analytic Solution of a Schaefer-Schwartz Stochastic Process

This and the next several suggestions deal with the choice of a joint stochastic process similar in form to the process used by Schaefer and Schwartz (1983). In their paper, Schaefer and Schwartz note that the long rate is empirically uncorrelated with the *spread* between the short and long rates. Because of this, they propose using a joint process which is in terms of the spread, $s = r - l$, and the consol yield, l.

$$ds = \beta_1\, dt + \sigma_1\, dw_1$$

$$dl = \beta_2\, dt + \sigma_2\, dw_2$$

$$\mathrm{E}[dw_1] = \mathrm{E}[dw_2] = 0, \quad dw_1 dw_2 = 0$$

The advantage of this formulation is that there is no correlation between the processes for s and l, which simplifies the asset pricing partial differential equation by eliminating the cross-derivative term in z_{sl}.

The specific joint process form proposed here

$$ds = \alpha(\mu - s)\,dt + \sigma_s\,dw_1,$$

$$dl = l(\sigma_l^2 + \sigma_l\lambda_l - s)\,dt + \sigma_l l\,dw_2,$$

is slightly different from the one used by Schaefer and Schwartz and allows an analytic solution for the transition probability function. If we make a transformation of variables to $x = s/\sigma_s$ and $y = \ln(l)/\sigma_l$, the result is the homoscedastic process

$$dx = \alpha(\frac{\mu}{\sigma_s} - x)\,dt + dw_1,$$
$$dy = (\sigma_l + \lambda_l - \frac{\sigma_s}{\sigma_l}x)\,dt + dw_2.$$

Notice that the stochastic differential equation for x does not contain any reference to the variable y, and can therefore easily be solved.

$$x(t) = x(0) + [1 - \exp(-t\alpha)]\left(\frac{\mu}{\sigma_s} - x(0)\right) + \int_0^t \exp[-(t-s)\alpha]\,dw_1(s)$$

Using this expression for x, we can solve for y.[82]

$$y(t) = y(0) + (\sigma_l + \lambda_l)t - \frac{\sigma_s}{\sigma_l}\int_0^t x(s)\,ds + \int_0^t dw_2(s)$$

We see from the process in x that

$$\begin{aligned}
\int_0^t x(s)\,ds &= \int_0^t \left[\frac{\mu}{\sigma_s}\,ds - \frac{1}{\alpha}\,dx(s) + \frac{1}{\alpha}\,dw_1(s)\right] \\
&= \frac{\mu}{\sigma_s}t - \frac{1}{\alpha}[x(t) - x(0)] + \frac{1}{\alpha}\int_0^t dw_1(s) \\
&= \frac{\mu}{\sigma_s}t - \frac{1}{\alpha}[1 - \exp(-t\alpha)]\left(\frac{\mu}{\sigma_s} - x(0)\right) \\
&\quad + \frac{1}{\alpha}\int_0^t [1 - \exp(-(t-s)\alpha)]\,dw_1(s),
\end{aligned}$$

82 I assume that the market price of long term risk function λ_l is constant. The analysis is similar if λ_l is a linear function of x and y.

which allows us to finish our solution for y.

$$y(t) = y(0) + \left(\sigma_l + \lambda_l - \frac{\mu}{\sigma_l}\right) t + \frac{\sigma_s}{\alpha\sigma_l}[1 - \exp(-t\alpha)] \left(\frac{\mu}{\sigma_s} - x(0)\right)$$
$$- \frac{\sigma_s}{\alpha\sigma_l} \int_0^t [1 - \exp(-(t-s)\alpha)]\, dw_1(s) + \int_0^t dw_2(s)$$

We see, then, that the expectations and variances of x and y are analytically solvable.

$$\mathrm{E}[x(t)] = x(0) + [1 - \exp(-t\alpha)] \left(\frac{\mu}{\sigma_s} - x(0)\right)$$
$$\mathrm{E}[y(t)] = y(0) + \left(\sigma_l + \lambda_l - \frac{\mu}{\sigma_l}\right) t + \frac{\sigma_s}{\alpha\sigma_l}(\mathrm{E}[x(t)] - x(0))$$
$$\begin{aligned}\mathrm{var}[x(t)] &= \int_0^t \exp[-2(t-s)\alpha]\, ds \\ &= \frac{1}{2\alpha}[1 - \exp(-2t\alpha)]\end{aligned}$$
$$\begin{aligned}\mathrm{cov}[x(t), y(t)] &= -\frac{\sigma_s}{\alpha\sigma_l} \int_0^t \exp(-(t-s)\alpha)[1 - \exp(-(t-s)\alpha)]\, ds \\ &= \frac{\sigma_s}{2\sigma_l\alpha^2}[1 - \exp(-t\alpha)]^2\end{aligned}$$
$$\begin{aligned}\mathrm{var}[y(t)] &= \left(\frac{\sigma_s}{\alpha\sigma_l}\right)^2 \int_0^t [1 - \exp(-(t-s)\alpha)]^2\, ds + \int_0^t ds \\ &= \left[1 + \left(\frac{\sigma_s}{\alpha\sigma_l}\right)^2\right] t - \frac{1}{2\alpha}\left(\frac{\sigma_s}{\alpha\sigma_l}\right)^2 [1 - \exp(-t\alpha)][3 - \exp(-t\alpha)]\end{aligned}$$

With observations spaced Δt apart, we have

$$\mathrm{E}[s_i] = s_{i-1} + (1-\delta)(\mu - s_{i-1}),$$
$$\mathrm{E}[\ln(l_i)] = \ln(l_{i-1}) + (\sigma_l^2 + \sigma_l\lambda_l - \mu)\Delta t + \frac{1-\delta}{\alpha}(\mu - s_{i-1}),$$
$$\mathrm{var}[\Delta s] = \frac{\sigma_s^2}{2\alpha}(1 - \delta^2),$$
$$\mathrm{cov}[\Delta s, \Delta \ln l] = \frac{\sigma_s^2}{2\alpha^2}(1-\delta)^2,$$
$$\mathrm{var}[\Delta \ln l] = \left(\sigma_l^2 + \frac{\sigma_s^2}{\alpha^2}\right)\Delta t - \frac{\sigma_s^2}{2\alpha^3}(1-\delta)(3-\delta),$$

where

$$\delta = \exp(-\Delta t \alpha).$$

These expressions can be used to maximize the likelihood function

$$L(\alpha, \mu, \sigma_s, \sigma_l) = \prod_{i=2}^{T} P(\epsilon_i),$$

$$P(\epsilon_i) = \frac{1}{2\pi|\Sigma|} \exp\left(-\frac{1}{2}\epsilon_i' \Sigma^{-1} \epsilon_i\right),$$

$$\epsilon_i = \begin{pmatrix} s_i - \mathrm{E}[s_i] \\ \ln(l_i) - \mathrm{E}[\ln(l_i)] \end{pmatrix}, \quad \Sigma = \begin{pmatrix} \mathrm{var}[\Delta s] & \mathrm{cov}[\Delta s, \Delta \ln l] \\ \mathrm{cov}[\Delta s, \Delta \ln l] & \mathrm{var}[\Delta \ln l] \end{pmatrix},$$

to find the 'best-fitting' parameters.

B.8.2 Simultaneous Solution of First Partial Derivatives

If in addition to the partial differential equation which we solve for the price we had two other partial differential equations which we solved for the two first partial derivatives, then perhaps accuracy of the interpolation methods used would increase. That is, say that we are using a Schaefer-Schwartz stochastic differential equation in s and l such as the one discussed above. Since we assume that there is zero correlation between the processes for s and l, the cross-derivative term in z_{sl} does not appear in the partial differential equation for z. The partial differential equation, therefore, takes the form

$$Hz = z_\tau, \quad H = a\frac{\partial^2}{\partial s^2} + b\frac{\partial^2}{\partial l^2} + c\frac{\partial}{\partial s} + d\frac{\partial}{\partial l} - (s + l).$$

Taking the partial derivative of both sides of this partial differential equation with respect to s or l gives us two more partial differential equations, this time in terms of z_s and z_l.

$$\frac{\partial}{\partial s}Hz = H_s z + Hz_s = (z_s)_\tau, \quad H_s = a_s\frac{\partial^2}{\partial s^2} + b_s\frac{\partial^2}{\partial l^2} + c_s\frac{\partial}{\partial s} + d_s\frac{\partial}{\partial l} - 1$$

$$\frac{\partial}{\partial l}Hz = H_l z + Hz_l = (z_l)_\tau, \quad H_l = a_l\frac{\partial^2}{\partial s^2} + b_l\frac{\partial^2}{\partial l^2} + c_l\frac{\partial}{\partial s} + d_l\frac{\partial}{\partial l} - 1$$

These three partial differential equations form a system of equations

$$\mathbf{Hz} \;\; \mathbf{z}_\tau, \quad \mathbf{H} = \begin{pmatrix} H & 0 & 0 \\ H_s & H & 0 \\ H_l & 0 & H \end{pmatrix}, \quad \mathbf{z} = \begin{pmatrix} z \\ z_s \\ z_l \end{pmatrix},$$

with the solution

$$\mathbf{z}(r, l, \tau) = \exp(\tau\mathbf{H})\mathbf{z}(r, l, 0).$$

If we split the operator $\mathbf{H}$ in a manner reminiscent of the Douglas-Rachford procedure

$$\mathbf{H} = \mathbf{S} + \mathbf{L}, \quad \mathbf{S} = \begin{pmatrix} S & 0 & 0 \\ S_s & S & 0 \\ S_l & 0 & S \end{pmatrix}, \quad \mathbf{L} = \begin{pmatrix} L & 0 & 0 \\ L_s & L & 0 \\ L_l & 0 & L \end{pmatrix},$$

with

$$S = a\frac{\partial^2}{\partial s^2} + c\frac{\partial}{\partial s} - \frac{1}{2}(s+l), \quad L = b\frac{\partial^2}{\partial l^2} + d\frac{\partial}{\partial l} - \frac{1}{2}(s+l),$$

and similar definitions for S_s, L_s, S_l and L_l, we can replace $\exp(\Delta\tau\mathbf{H})$ with

$$\exp(\Delta\tau\mathbf{H}) = \left(\mathbf{I} - \frac{1}{2}\Delta\tau\mathbf{S}\right)^{-1}\left(\mathbf{I} - \frac{1}{2}\Delta\tau\mathbf{L}\right)^{-1}\left(\mathbf{I} + \frac{1}{2}\Delta\tau\mathbf{L}\right)\left(\mathbf{I} + \frac{1}{2}\Delta\tau\mathbf{S}\right)$$
$$+ O(\Delta\tau^3).$$

This means that we can approximate the solution $\mathbf{z}$ by

$$\left(\mathbf{I} - \frac{1}{2}\Delta\tau\mathbf{L}\right)\left(\mathbf{I} - \frac{1}{2}\Delta\tau\mathbf{S}\right)\mathbf{z}(r, l, \tau + \Delta\tau) = \left(\mathbf{I} + \frac{1}{2}\Delta\tau\mathbf{L}\right)\left(\mathbf{I} + \frac{1}{2}\Delta\tau\mathbf{S}\right)\mathbf{z}(r, l, \tau),$$

which can be expressed in Douglas-Rachford alternating direction form as

$$\left(\mathbf{I}-\frac{1}{2}\Delta\tau\mathbf{L}\right)\mathbf{z}^*(r,l,\tau+\Delta\tau)=\left(\mathbf{I}+\frac{1}{2}\Delta\tau(\mathbf{H}+\mathbf{S})\right)\mathbf{z}(r,l,\tau)=\alpha(r,l,\tau),$$

$$\left(\mathbf{I}-\frac{1}{2}\Delta\tau\mathbf{S}\right)\mathbf{z}(r,l,\tau+\Delta\tau)=\mathbf{z}^*(r,l,\tau+\Delta\tau)-\frac{1}{2}\Delta\tau\mathbf{S}\mathbf{z}(r,l,\tau)=\beta(r,l,\tau).$$

With the system in this form, we can make a final simplification to eliminate some of the partial derivatives in the operators above. For example, in the solution for $\mathbf{z}^*(r,l,\tau+\Delta\tau)$ it can be shown that

$$\begin{aligned}&\left(\mathbf{I}-\frac{1}{2}\Delta\tau\mathbf{L}\right)\mathbf{z}^*(r,l,\tau+\Delta\tau)\\ &=\phi^{-1}\begin{pmatrix}1&0&2\theta\left(b\frac{\partial}{\partial l}+d\right)\\ \theta&1&-2\theta\left(b\frac{\partial^2}{\partial s\partial l}+d\frac{\partial}{\partial s}+b_s\frac{\partial}{\partial l}+d_s\right)\\ \theta&0&1-2\theta\left(b\frac{\partial^2}{\partial l^2}+(b_l+d)\frac{\partial}{\partial l}+d_l\right)\end{pmatrix}\begin{pmatrix}z^*\\ z_s^*\\ z_l^*\end{pmatrix}\\ &=\alpha(r,l,\tau),\end{aligned}$$

where

$$\phi=\frac{1}{1+r\Delta t/4},$$
$$\theta=\frac{\Delta t}{4}\phi.$$

This can easily be converted to the more convenient upper triangular form

$$\begin{aligned}&\begin{pmatrix}1&0&0\\ -\theta&1&0\\ -\theta&0&1\end{pmatrix}\left(\mathbf{I}-\frac{1}{2}\Delta\tau\mathbf{L}\right)\mathbf{z}^*\\ &\qquad=\phi^{-1}\begin{pmatrix}1&0&-2\theta\left(b\frac{\partial}{\partial l}+d\right)\\ 0&1&-2\theta u_{23}\\ 0&0&1-2\theta u_{33}\end{pmatrix}\begin{pmatrix}z^*\\ z_s^*\\ z_l^*\end{pmatrix}\\ &\qquad=\begin{pmatrix}\alpha_1\\ \alpha_2-\theta\alpha_1\\ \alpha_3-\theta\alpha_1\end{pmatrix},\end{aligned}$$

where

$$u_{23} = b\frac{\partial^2}{\partial s \partial l} + d\frac{\partial}{\partial s} + (b_s - \theta b)\frac{\partial}{\partial l} + (d_s - \theta d)$$

$$u_{33} = b\frac{\partial^2}{\partial l^2} + (b_l + d - \theta b)\frac{\partial}{\partial l} + (d_l - \theta d).$$

This reduces the entire solution of $\mathbf{z}^*$ to the solution of a differential equation in z_l^*.

$$(1 - 2\theta u_{33})z_l^* = \phi(\alpha_3 - \theta\alpha_1)$$

Once the tridiagonal system resulting from the finite difference approximation to the above equation in z_l^* has been solved, the values of z_s^* and z^* are simply calculated as

$$z^* = \phi\alpha_1 + 2\theta\left(b\frac{\partial}{\partial l} + d\right)z_l^*$$

$$z_s^* = \phi(\alpha_2 - \theta\alpha_1) + 2\theta u_{23} z_l^*$$

Given the solution for $\mathbf{z}^*$, we can proceed similarly to solve for $\mathbf{z}$ using $(\mathbf{I} - \frac{1}{2}\Delta\tau\mathbf{L})\mathbf{z} = \beta$. This procedure may result in better estimates of the first partial derivatives of the solution, and improve interpolation accuracy.

B.8.3 Asset Pricing by Risk-Adjusted Expectation

The amount of memory required by the alternating direction method for a large grid size is quite large, and may exceed that available on, for example, a microcomputer. There is an approximation that can be made to simplify the solution procedure and reduce the amount of memory needed. The question, of course, is whether or not the approximation results in a poor numerical solution to the partial differential equation.

As Cox and Ross (1976) have shown, the value of an asset described by the models we are dealing with can be expressed as the expectation of a discounted value

$$z(s,l,\tau) = \mathrm{E}\left[\exp\left(-\int_{\tau}^{\tau-\Delta\tau} r(\nu)\,d\nu\right) z(s,l,\tau-\Delta\tau)\right],$$

where the expectation is taken with respect to the *risk-adjusted* stochastic process for s and l. That is, if we were using the stochastic process

$$ds = \alpha(\mu - s)\,dt + \sigma_s\,dw_1,$$

$$dl = l(\sigma_l^2 + \sigma_l\lambda_l - s)\,dt + \sigma_l l\,dw_2,$$

the risk-adjusted process would be

$$ds = [\alpha(\mu - s) - \sigma_s\lambda_s]\,dt + \sigma_s\,dw_1,$$

$$dl = l(\sigma_l^2 - s)\,dt + \sigma_l l\,dw_2.$$

As was shown in a previous section, the transition probability function for this risk-adjusted process can be solved analytically. If we change over to a homoscedastic system by using the new variable $\ln(l)$ instead of l, then the elements $\mathrm{var}[\Delta s]$, $\mathrm{cov}[\Delta s, \Delta \ln l]$ and $\mathrm{var}[\Delta \ln l]$ will be independent of s, $\ln(l)$ and τ. Then, making the approximation

$$z(s,l,\tau) \simeq \exp\left(-\frac{1}{2}\Delta\tau\,[r(\tau) + \mathrm{E}(r(\tau-\Delta\tau))]\right)\mathrm{E}[z(s,l,\tau-\Delta\tau)]$$

allows us to compute the discount factors

$$\exp\left(-\frac{1}{2}\Delta\tau\,[r(\tau) + \mathrm{E}(r(\tau-\Delta\tau))]\right),$$

only once for each grid point.

The calculation of the expectation $E[z(s, l, \tau - \Delta\tau)]$ can be considerably simplified, along the lines of Brennan and Schwartz (1978). Basically, that paper illustrated how the transition probability function resulting from a continuous time stochastic process could be approximated by a set of jump probabilities where the jumps are restricted to be to the grid points available. That is, if we had a mean-zero normal process, $x \sim N\{0, \sigma_x^2\}$, instead of computing averages and variances as

$$E[f(x)] = \frac{1}{2\pi} \int_{-\infty}^{\infty} f(x) \exp\left(-\frac{1}{2}\frac{x^2}{\sigma_x^2}\right) dx,$$

$$\text{var}[f(x)] = \frac{1}{2\pi} \int_{-\infty}^{\infty} [f(x) - E(f(x))]^2 \exp\left(-\frac{1}{2}\frac{x^2}{\sigma_x^2}\right) dx,$$

we use the discrete jump probabilities $p_{xi} = p_x q_x^{|i|}$ where p_{xi} is the probability of jumping i grid points, that is from 0 to $i\Delta x$.

$$E[f(x)] = \sum_{i=-\infty}^{\infty} f(x + i\Delta x) p_{xi}$$

$$\text{var}[f(x)] = \sum_{i=-\infty}^{\infty} [f(x + i\Delta x) - E(f(x))]^2 p_{xi}$$

In order to be interpreted as probabilities, the sum of the p_{xi} must equal 1. In addition, we must ensure that the variance of x is, in fact, equal to σ_x^2. These two restrictions uniquely determine the two unknowns p_x and q_x.

$$\sum_{i=-\infty}^{\infty} p_{xi} = p_x \sum_{i=-\infty}^{\infty} q_x^{|i|} = p_x \left(1 + 2\sum_{i=1}^{\infty} q_x^i\right)$$

$$= p_x \left(\frac{1+q_x}{1-q_x}\right) = 1$$

$$p_x = \left(\frac{1-q_x}{1+q_x}\right)$$

$$\text{var}[x] = \sum_{i=-\infty}^{\infty} [i\Delta x]^2 p_{xi} = 2\Delta x^2 \left(\frac{1-q_x}{1+q_x}\right) \sum_{i=1}^{\infty} i^2 q_x^i$$

$$= 2\Delta x^2 \frac{q_x}{(1-q_x)^2} = \sigma_x^2$$

$$q_x^2 - 2\left(1 + \frac{\Delta x^2}{\sigma_x^2}\right) q_x + 1 = 0$$

$$q_x = 1 + \frac{\Delta x^2}{\sigma_x^2} \pm \sqrt{\left(1 + \frac{\Delta x^2}{\sigma_x^2}\right)^2 - 1}$$

$$= 1 + \frac{\Delta x^2}{\sigma_x^2} \left(1 \pm \sqrt{1 + 2\frac{\sigma_x^2}{\Delta x^2}}\right)$$

Since we want q_x to be between 0 and 1, we choose

$$q_x = 1 - \frac{\Delta x^2}{\sigma_x^2} \left(\sqrt{1 + 2\frac{\sigma_x^2}{\Delta x^2}} - 1\right).$$

When we have two independent mean-zero variables

$$\begin{pmatrix} x \\ y \end{pmatrix} \sim \mathrm{N}\{\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_x^2 & 0 \\ 0 & \sigma_y^2 \end{pmatrix}\},$$

we can approximate the expectation of $z(x, y)$ as

$$\mathrm{E}[z(x,y)] = \sum_{i=-\infty}^{+\infty} \sum_{j=-\infty}^{+\infty} p_{xi} p_{yj} z(x + i\Delta x, y + j\Delta y),$$

where $p_{yj} = p_y q_y^{|j|}$ and p_y and q_y are defined similarly to p_x and q_x. If we define the sums

$$\alpha(x,y) = \alpha^+(x,y) - z(x,y) + \alpha^-(x,y),$$
$$\alpha^+(x,y) = \sum_{j=0}^{+\infty} q_y^j z(x, y + j\Delta y), \quad \alpha^-(x,y) = \sum_{j=0}^{+\infty} q_y^j z(x, y - j\Delta y),$$
$$\beta(x,y) = \beta^+(x,y) - \alpha(x,y) + \beta^-(x,y),$$
$$\beta^+(x,y) = \sum_{i=0}^{+\infty} q_x^i \alpha(x + i\Delta x, y), \quad \beta^-(x,y) = \sum_{i=0}^{+\infty} q_x^i \alpha(x - i\Delta x, y),$$

then we can express $\mathrm{E}[z(x,y)]$ as

$$\mathrm{E}[z(x,y)] = p_x p_y \beta(x,y).$$

In addition, there are simple recursion formulae for computing α and β.

$$\alpha^+(x,y) = z(x,y) + q_y \alpha^+(x, y + \Delta y)$$

$$\alpha^-(x,y) = z(x,y) + q_y \alpha^-(x, y - \Delta y)$$

$$\beta^+(x,y) = \alpha(x,y) + q_x \beta^+(x + \Delta x, y)$$

$$\beta^-(x,y) = \alpha(x,y) + q_x \beta^-(x - \Delta x, y)$$

This general idea might be used for our problem. The key lies in separately handling the expected drift and variance terms in the stochastic process. That is, we approximate

$$\mathrm{E}[z(x(\tau), y(\tau), \tau] \simeq \widehat{\mathrm{E}}\left[z\left(\mathrm{E}[x(\tau - \Delta\tau)], \mathrm{E}[x(\tau - \Delta\tau)], \tau - \Delta\tau\right)\right],$$

where both E and $\widehat{\mathrm{E}}$ are expectations with respect to a homoscedastic stochastic process, but $\widehat{\mathrm{E}}$ is an expectation ignoring the drift of the process. That is, we are assuming that the variance and drift terms of x and y are independent and can be considered one after another. First we take account of the drift, acting as if the variance was zero, and then we take account of the variance as if the drift were zero.

Because of the simple recursion relations defining α and β above, computational requirements of this method are of the same order as numerical solution using the alternating direction method, namely, $O(IJ)$ for an $I \times J$ grid. Similarly, memory requirements are also $O(IJ)$. This is, of course, quite desirable, but even more desirable is the fact that the memory requirements are small when compared with the alternating direction method. This approach requires only four $I \times J$ grids, one for the solution at a particular point in time, another for calculating α and β and two others for storing the expected drifts in the x and y directions. This method should be compared with a standard partial differential equation solution method to see whether the its approximations affect the solution values to any marked degree.

B.8.4 Testing for Instability

In this study, I assumed that my numerical solutions were relatively unaffected by instability, even though the conditions for stability could not be shown to hold. One way to test for possible instability in the numerical solution of, say, a bond call option, $C(r,l,\tau)$, would be to also solve for the difference between the bond price and option value, namely, $X(r,l,\tau) = B(r,l,\tau+\tau_B;c) - C(r,l,\tau;K)$. If the values of C from these two approaches differ too greatly, then it would have to be due to some instability in the system.

Alternatively, a more traditional approach could be used. The more common method is to first numerically solve for $C(r,l,\tau+\Delta\tau)$ using the values from the previous step, $C(r,l,\tau)$, and a single time step of size $\Delta\tau$. Then, we once more solve for $C(r,l,\tau+\Delta\tau)$, but this time using two time steps of size $\frac{1}{2}\Delta\tau$. If the difference between the two values exceeds an allowable error tolerance, the step size is decreased until the tolerance is acceptable.

If properly done, these two methods should require the same amount of memory. The traditional method, however, requires an additional numerical solution per time step.

B.8.5 Grid Spacing and Boundary Conditions

The numerical methods used in this study were all developed assuming that the underlying r,l grid had regular spacings of Δr between points in the r direction and Δl in the l direction. The accuracy of our procedures could probably be greatly increased by using an irregularly spaced grid instead, placing more grid points in the areas of high curvature of the solution (the at the money region) and allowing larger spacings between points elsewhere. An attempt in this direction, however, may lead to difficulties. If the spacings are made too large in some parts of the grid in order to allow closer spacing elsewhere we may find instability showing up in the solutions.

Perhaps a related problem could be addressed at the same time, namely, the fact that we are not interested in what values the numerical solution takes on in large areas of the grid. As our interest rate data show, r and l are seldom very far apart, which means that only a relatively small region surrounding the diagonal $r = l$ is of interest to us. If we transformed variables in order to exclude some of the uninteresting areas, presumably we would be able to move our grid points closer together. For example, we could use the Schaefer-Schwartz joint process for s and l instead of the Brennan-Schwartz process for r and l. This automatically places the solution grid along the diagonal, $r = l$.

An alternative to this would be to use the variables which transform the stochastic process used in this study to a homoscedastic form,[83] namely,

$$\mathbf{x} = \mathbf{G}^{-1} \begin{pmatrix} \ln(r)/\sigma_r \\ \ln(l)/\sigma_l \end{pmatrix}, \quad \mathbf{G}\mathbf{G}' = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}.$$

Since $\mathbf{G}$ is only determined to within a rotation, we can choose a $\mathbf{G}$ such that the $\mathbf{x}$ grid excludes most of the off-diagonal area included in the r, l grid.[84]

[83] See section B.5.2.a.

[84] In this study, on the boundaries where r and l are at their largest values the boundary conditions $z_{rr} = 0$ and $z_{ll} = 0$, respectively, were used. If these bound-

B.9 SUMMARY AND CONCLUSIONS

In this study the Brennan and Schwartz two-factor model, which has previously been used to value bonds, was applied to the valuation of options on US government bonds and treasury bills. The results were promising, and suggest that the model prices for options are accurate enough for practical purposes.

Several innovations on the Brennan-Schwartz methodology were introduced in this study. One of these was the due to an examination of the 'simple linearization method' used by Brennan-Schwartz (1983b) and Ananthanarayanan (1978) in estimating the parameters of the joint process in r and l. For the particular data and process forms used in this study, this method was found to be unsuitable for the estimation of stochastic drift parameters, although estimates of the variance and covariance parameters appeared to be well estimated. This problem was addressed by estimating the drift parameter required, α, at the same time as the market price of risk parameter, λ_r. That is, a two-dimensional search was performed to find the pair of values, (α, λ_r), which minimized an appropriate distance function.

The theoretical bond values required at each iteration of the non-linear search procedure for finding the best-fitting pair, (α, λ_r), were calculated by numerically

---

aries are moved in closer to the diagonal, however, it may be that these boundary conditions become inappropriate. If we are solving a system of equations for the price plus its first partial derivatives, as in section B.8.2, we could easily impose the boundary conditions $z_{rrr} = 0$, $z_{rrl} = 0$, $z_{rll} = 0$ and $z_{lll} = 0$, instead.

solving the asset partial differential equation derived in section B.2. A comparison was made between discount function values derived numerically by the successive overrelaxation method (SOR) used by Brennan and Schwartz and the alternating direction method used by Schaefer and Schwartz. For the range of parameter values tested, the SOR and alternating direction methods produced almost identical results. When the results differed, it was the result of instability of the solution, an occurrence which appeared more often when using the SOR method. Because of its apparent greater stability in the problem at hand and its lower computational cost, the alternating direction method was used in the rest of the study.

The estimation procedure for (α, λ_r) consisted of forming portfolios of bonds to represent various maturities for each time point available, and then using an estimated covariance matrix of portfolio pricing errors in forming a weighted sum of squared bond pricing errors. Such a distance function was minimized for three different portfolio formation schemes and three different covariance matrix assumptions. The values of (α, λ_r) were found to be relatively similar for the nine distance functions that were minimized.

Once all needed parameters had been estimated, the numerical solution for bond and treasury bill option values was straightforward. In order to compare the numerical option values with actual option values, however, some form of interpolation was required. Because most of the option price observations collected were for options trading close to the money, the high curvature of the theoretical solution in this region resulted in poor correspondence between actual and theoretical values when using linear interpolation. The other two methods tested

– cubic spline and quadratic interpolation – also had some difficulty in the high curvature region, but the quadratic method in general gave reasonable values.

Comparison of market and theoretical option prices showed that there were differences between theoretical and market values. When 'in-sample' variance estimates were used, the size of these errors decreased considerably suggesting that further research into the optimal length of the parameter estimation period and possible use of a moving estimation period could result in better pricing results.

One possible explanation for the remaining pricing errors is incorrectness of the market data used, and the arbitrage tests at the end of this study were designed to examine this possibility. If arbitrage profits had resulted from the tests, it would not necessarily have indicated that the market data were incorrect due to market inefficiency, although this is one possibility. Another possible interpretation would have been that there were apparent – but not necessarily realizable – market inefficiencies due to such factors as transcription errors in the data, non-simultaneity of bond and option quotations, or use of quotations at which trading could not have occurred due to thin trading. The close correspondence between theoretical and market prices, such that returns from arbitrage, while positive, were insufficient to cover reasonable transactions costs, suggests that the model prices are accurate enough to justify practical use of the model, and that realizable arbitrage possibilities promising positive returns after transactions costs would presumably be detectable.

B.10 REFERENCES

Ananthanarayanan, A.L., 1978, A stochastic specification of short term interest rate process and the pricing of extendible and retractable bonds, unpublished PhD dissertation, University of British Columbia.

Ball, C.B. and W.N. Torous, 1983, Bond price dynamics and options, Journal of Financial and Quantitative Analysis 18, 517-531.

Black, F. and M. Scholes, 1973, The pricing of options and corporate liabilities, Journal of Political Economy 8, 637-654.

Brennan, M.J. and E.S. Schwartz, 1977, Savings bonds, retractable bonds and callable bonds, Journal of Financial Economics 5, 67-88.

Brennan, M.J. and E.S. Schwartz, 1978, Finite difference methods and jump processes arising in the pricing of contingent claims: a synthesis, Journal of Financial and Quantitative Analysis, September, 461-474.

Brennan, M.J. and E.S. Schwartz, 1979, A continuous time approach to the pricing of bonds, Journal of Banking and Finance 3, July, 133-155.

Brennan, M.J. and E.S. Schwartz, 1980, Conditional predictions of bond prices and returns, Journal of Finance 35, No. 2, 405-417.

Brennan, M.J. and E.S. Schwartz, 1982, An equilibrium model of bond pricing and a test of market efficiency, Journal of Financial and Quantitative Analysis 17, No. 3, 301-329.

Brennan, M.J. and E.S. Schwartz, 1983a, Alternate methods for valuing debt options, Finance 4, No. 2, 119-137.

Brennan, M.J. and E.S. Schwartz, 1983b, Duration, bond pricing and portfolio management, in: G. Kaufman, G. Biervag and A. Toevs, eds., Innovations in portfolio management: Duration analysis and immunization (JAI Press).

Courtadon, G., 1982, The pricing of options on default free bonds, Journal of Financial and Quantitative Analysis 17, 75-100.

Cox, J.C., J.E. Ingersoll and S.A. Ross, 1978, A theory of the term structure of interest rates, Research paper No. 468 (Graduate School of Business, Stanford University, Stanford, CA).

Cox, J.C. and S.A. Ross, 1976, The valuation of options for alternative stochastic processes, Journal of Financial Economics 3, No. 1/2, 145-166.

Debreu, G., 1959, Theory of Value (New York: John Wiley).

Douglas, J. and J.E. Gunn, 1964, A general formulation of alternating direction methods. Part 1. Parabolic and hyperbolic problems, Num. Math. 6, 428-453.

Douglas, J. and H.H. Rachford, 1956, On the numerical solution of heat conduction problems in two and three space variables, Trans. Am. Math. Soc. 82, 421-439.

Fairweather, G. and A.R. Mitchell, 1967, A new computational procedure for A.D.I. methods, SIAM Journal 4, No. 2, 163-170.

Maliaris, A.G. and W.A. Brock, 1982, Stochastic methods in economics and finance (Amsterdam: North-Holland).

Malinvaud, E., 1966, Statistical methods of econometrics (Amsterdam: North-Holland).

Marquardt, D.W., 1963, An algorithm for least-squares estimation of nonlinear parameters, J. Soc. Indust. Appl. Math. 11, No. 2, 431-441.

McKee, S. and A.R. Mitchell, 1970, Alternating direction methods for parabolic equations in two space dimensions with a mixed derivative, The Computer Journal 13, No. 1, 81-86.

Merton, R.C., 1973, Theory of rational option pricing, Journal of Finance 29, 449-470.

Phillips, P.C.B., 1972, The structural estimation of a stochastic differential equation system, Econometrica 40, 1021-1041.

Richard, S., 1978, An arbitrage model of the term structure of interest rates, Journal of Financial Economics 6, No. 1, 33-57.

Ross, S.A., 1976, The arbitrage theory of capital asset pricing, Journal of Economic Theory 13, 341-360.

Schaefer, S.M. and E.S. Schwartz, 1983, A two factor model of the term structure: an approximate solution, unpublished paper.

Varga, R.S., 1962, Matrix iterative analysis (Englewood Cliffs, N.J: Prentice-Hall).

Vasicek, O., 1977, An equilibrium characterization of the term structure, Journal of Financial Economics 5, 177-178.

Vemuri, V. and W.J. Karplus, 1981, Digital computer treatment of partial differential equations (Englewood Cliffs, NJ: Prentice-Hall).

Table B.1.

Bond and bond option data collected from the Wall Street Journal over the period of November 1, 1982 to October 31, 1983.

	B[b]		calls[c]				puts					
				K				K			total	total
bond/note[a]	min	max	n	min	max	N	n	min	max	N	options	obs
$13\frac{3}{4}$*s*, 1992 May 15n	$109\frac{2}{32}$	$119\frac{3}{32}$	13	112	120	265	11	112	120	198	24	463
$10\frac{1}{2}$*s*, 1992 Nov 15n	$91\frac{14}{32}$	$102\frac{11}{32}$	11	96	104	190	12	92	104	117	23	307
$10\frac{7}{8}$*s*, 1993 Feb 15n	$94\frac{16}{32}$	$104\frac{16}{32}$	8	92	108	84	5	96	104	24	13	108
$10\frac{1}{8}$*s*,1993 May 15n	$89\frac{2}{32}$	$100\frac{2}{32}$	13	90	102	77	10	88	100	22	23	99
$11\frac{7}{8}$*s*, 1993 Aug 15n	$101\frac{2}{32}$	$103\frac{13}{32}$	3	100	104	14	4	98	104	9	7	23
14*s* , 2011 Nov 15	113	$130\frac{12}{32}$	47	112	132	1003	47	112	132	515	94	1518
$10\frac{3}{8}$*s*, 2012 Nov 15	$86\frac{17}{32}$	$101\frac{6}{32}$	38	86	104	780	34	86	104	375	72	1155
12*s* , 2013 Aug 15	$99\frac{6}{32}$	$105\frac{19}{32}$	12	98	106	101	6	98	104	19	18	120
total options			145				129				274	
total observations						2514				1279		3793

[a] Bonds followed by the letter 'n' are actually treasury notes.

[b] Bond price in dollars per $100 principal value.

[c] n denotes the number of options that were written over the period on each bond.
K is the option exercise price per $100 principal value.
N is the number of option price observations collected for the bond.

Table B.2.

Number of bond option observations for various bond option price levels.

p^a	N^b	
	calls	puts
0-1	795	269
1-2	718	309
2-3	432	275
3-4	272	149
4-5	140	86
5-6	97	72
6-7	40	38
7-8	15	28
8-15	5	53

[a] Range of bond option price in dollars per $100 principal value.
[b] Number of observations.

Table B.3.

Bond option data by number of months to maturity.

	calls[b]			puts		
m^a	N	$\bar{p}$	σ_p	N	$\bar{p}$	σ_p
0-1	464	1.29	1.42	278	1.72	1.90
1-2	529	1.73	1.59	271	2.82	2.81
2-3	555	2.06	1.55	249	2.23	1.72
3-4	466	2.17	1.39	217	3.16	1.87
4-5	252	2.30	1.42	139	3.78	2.88
5-6	120	2.22	1.34	60	3.30	1.97
6-7	72	2.74	1.76	24	3.94	2.43
7-8	43	3.33	1.93	34	4.63	2.63
8-9	13	4.00	1.37	7	3.18	1.56
0-9	2514	1.95	1.56	1279	2.72	2.35

[a] Number of 'months' to maturity of the option, where a 'month' is actually 30 days. eg. the first row summarizes all observations with 1 to 30 days to maturity.

[b] N denotes number of option observations.
$\bar{p}$ is the average option price in dollars.
σ_p is the sample standard deviation of option prices.

Table B.4.

Bond option data for in and out of the money options.

B/K[a]	calls[b]			puts		
	N	$\bar{p}$	σ_p	N	$\bar{p}$	σ_p
0.85-0.98	704	0.73	0.66	270	6.08	2.55
0.98-1.00	687	1.36	0.82	293	2.60	1.11
1.00-1.02	637	2.19	0.97	331	1.91	1.14
1.02-1.15	486	4.24	1.30	385	1.16	0.93

[a] Bond price, B, divided by option exercise price, K.

[b] N denotes number of option observations.
$\bar{p}$ is the average option price in dollars.
σ_p is the sample standard deviation of option prices.

Table B.5.

Treasury bill and treasury bill option data collected from the Wall Street Journal over the period of November 1, 1982 to October 31, 1983.

B^a		calls[b]				puts					
			K				K			total	total
min	max	n	min	max	N	n	min	max	N	options	obs
97.3095	97.9845	20	97.4722	98.2306	395	17	97.4722	98.2306	424	37	819

[a] Treasury bill price in dollars per $100 principal value.

[b] n denotes the number of options that were written over the period on each bond. K is the option exercise price per $100 principal value. N is the number of option price observations collected for the bond.

Table B.6.

Number of treasury bill option observations for various treasury bill option price levels.

p^a	N^b	
	calls	puts
0.00-0.10	238	187
0.10-0.20	83	99
0.20-0.30	56	66
0.30-0.40	17	46
0.40-0.50	1	16
0.50-0.60	0	10

[a] Range of treasury bill option price in dollars per \$100 principal value.
[b] Number of observations.

Table B.7.

Treasury bill option data by number of months to maturity.

m^a	calls[b]			puts		
	N	$\bar{p}$	σ_p	N	$\bar{p}$	σ_p
0-1	62	0.11	0.10	93	0.16	0.14
1-2	79	0.09	0.08	77	0.12	0.10
2-3	80	0.10	0.09	77	0.13	0.11
3-4	84	0.12	0.10	86	0.18	0.14
4-5	39	0.12	0.08	30	0.22	0.13
5-6	16	0.14	0.12	22	0.21	0.12
6-7	11	0.18	0.11	15	0.28	0.15
7-8	16	0.14	0.08	14	0.17	0.11
8-9	8	0.17	0.08	10	0.21	0.17
0-9	395	0.11	0.09	424	0.16	0.13

[a] Number of 'months' to maturity of the option, where a 'month' is actually 30 days. eg. the first row summarizes all observations with 1 to 30 days to maturity.

[b] N denotes number of option observations.
$\bar{p}$ is the average option price in dollars.
σ_p is the sample standard deviation of option prices.

Table B.8.

Treasury bill option data for in and out of the money options.

B/K[a]	calls[b]			puts		
	N	$\overline{p}$	σ_p	N	$\overline{p}$	σ_p
0.9915-0.9965	77	0.03	0.03	123	0.32	0.10
0.9965-0.9980	100	0.06	0.04	102	0.14	0.07
0.9980-0.9995	133	0.11	0.04	119	0.10	0.08
0.9995-1.0025	85	0.25	0.06	80	0.05	0.04

[a] Treasury bill price, B, divided by option exercise price, K.

[b] N denotes number of option observations.
$\overline{p}$ is the average option price in dollars.
σ_p is the sample standard deviation of option prices.

Table B.9.

Analysis of the one-factor stochastic process examined by Ananthanarayanan, namely, $dr = m(\mu - r)dt + \sigma r^{\alpha} dw$, $dw \sim N\{0, dt\}$. Parameter values are $\alpha = 0.5$, $\mu = 0.09517$, $m = 0.007162$, $\sigma = 0.008856$, $\Delta t = 1/12$.

r_t [a]	x_t	$\overline{x}_{t+\Delta t}$ [b]	$r^-_{t+\Delta t}$	$r^+_{t+\Delta t}$
0.050	50.50	50.51	0.049	0.051
0.100	71.42	71.42	0.098	0.102
0.150	87.47	87.46	0.148	0.152
0.200	101.00	100.98	0.198	0.202

[a] r_t is the beginning of period instantaneous riskless rate.
x_t is the corresponding beginning of period value for the homoscedastic process

$$dx = \left[\frac{m}{\sigma}(\mu - r)r^{-\alpha} - \frac{1}{2}\alpha\sigma r^{\alpha-1}\right] dt + dw,$$

that is, $x_t = r_t^{1-\alpha}/\sigma(1-\alpha)$.

[b] $\overline{x}_{t+\Delta t}$ is the (simple linearization model) expected value of x at the end of the period,

$$\overline{x}_{t+\Delta t} = x_t + \left[\frac{m}{\sigma}(\mu - r)r_t^{-\alpha} - \frac{1}{2}\alpha\sigma r_t^{\alpha-1}\right] \Delta t.$$

$r^-_{t+\Delta t}$ is the short rate corresponding to $\overline{x}_{t+\Delta t} - 2\sqrt{\Delta t}$, that is, the short rate corresponding to the x which is two (simple linearization model) standard deviations below the mean.
$r^+_{t+\Delta t}$ is the short rate corresponding to $\overline{x}_{t+\Delta t} + 2\sqrt{\Delta t}$.

Table B.10.

Analysis of the one-factor stochastic process examined by Ananthanarayanan, namely, $dr = m(\mu - r)dt + \sigma r^{\alpha} dw$, $dw \sim N\{0, dt\}$. Parameter values are $\alpha = 0.5$, $\mu = 0.0012934$, $m = 0.0025221$, $\sigma = 0.00083096$, $\Delta t = 1/12$.

r_t [a]	x_t	$\overline{x}_{t+\Delta t}$ [b]	$r^-_{t+\Delta t}$	$r^+_{t+\Delta t}$
0.050	538.19	538.13	0.050	0.050
0.100	761.11	761.03	0.100	0.100
0.150	932.17	932.07	0.150	0.150
0.200	1076.38	1076.27	0.200	0.200

[a] r_t is the beginning of period instantaneous riskless rate.
x_t is the corresponding beginning of period value for the homoscedastic process

$$dx = \left[\frac{m}{\sigma}(\mu - r)r^{-\alpha} - \frac{1}{2}\alpha\sigma r^{\alpha-1}\right] dt + dw,$$

that is, $x_t = r_t^{1-\alpha}/\sigma(1-\alpha)$.

[b] $\overline{x}_{t+\Delta t}$ is the (simple linearization model) expected value of x at the end of the period,

$$\overline{x}_{t+\Delta t} = x_t + \left[\frac{m}{\sigma}(\mu - r)r_t^{-\alpha} - \frac{1}{2}\alpha\sigma r_t^{\alpha-1}\right] \Delta t.$$

$r^-_{t+\Delta t}$ is the short rate corresponding to $\overline{x}_{t+\Delta t} - 2\sqrt{\Delta t}$, that is, the short rate corresponding to the x which is two (simple linearization model) standard deviations below the mean.
$r^+_{t+\Delta t}$ is the short rate corresponding to $\overline{x}_{t+\Delta t} + 2\sqrt{\Delta t}$.

Table B.11.

Simple linearization method estimates for the Brennan-Schwartz interest rate process parameters.

period	α^a	$\sigma_l^2 + \sigma_l \lambda_l$	σ_r	σ_l	ρ
total period (Oct 70 - Oct 82)	0.876 (0.266)	0.0546 (0.0373)	0.442	0.142	0.530
first half (Oct 70 - Oct 76)	0.461 (0.298)	0.0276 (0.0554)	0.381	0.141	0.425
second half (Oct 76 - Oct 82)	1.644 (0.467)	0.0792 (0.0481)	0.495	0.142	0.631

[a] Standard errors in parentheses.

Table B.12.

Analysis of the two-factor Brennan-Schwartz process. Parameter values are $\alpha = 0.876$, $\sigma_l^2 + \sigma_l \lambda_l = 0.0546$, $\sigma_r = 0.442$, $\sigma_l = 0.142$, $\rho = 0.530$, $\Delta t = 1/12$.

$r_{t+\Delta t}$ [a]	$l_{t+\Delta t}$	$\beta_{1t+\Delta t}/\beta_{1t}$ [b]	$\beta_{2t+\Delta t}/\beta_{2t}$ [b]
0.0626	0.0964	3.1	1.2
0.0650	0.0934	2.3	1.1
0.0750	0.0929	0.9	1.0
0.0900	0.0962	-0.3	0.8
0.1000	0.1003	-0.8	0.7
0.1044	0.1050	-0.8	0.7
0.1000	0.1083	-0.2	0.8
0.0900	0.1091	0.7	1.0
0.0750	0.1061	2.2	1.2
0.0650	0.1005	3.1	1.2

[a] $(r_{t+\Delta t}, l_{t+\Delta t})$ is an end of period point with a (simple linearization model) Mahalanobis distance of 2 from the (simple linearization model) expected end of period point. The expected end point and the Mahalanobis distance were calculated using the homoscedastic transformation of the process in r and l. The starting point was $r = 0.08$, $l = 0.10$.

[b] β_{1t} and $\beta_{1t+\Delta t}$ are the beginning and end of period drifts in the first component of the homoscedastic system

$$\beta_{1t} = \frac{\alpha}{\sigma_r}(l_t - r_t) - \frac{1}{2}\sigma_r.$$

β_{2t} and $\beta_{2t+\Delta t}$ are similar values for the second component:

$$\beta_{2t} = \frac{1}{2}\sigma_l + \lambda_l + \frac{(l-r)}{\sigma_l}.$$

These columns show the ratios of the end of period to beginning of period drifts.

Table B.13.

Simple linearization method estimates of the Brennan-Schwartz interest rate process parameters for three different drift assumptions.

f^a	α^b	$\sigma_l^2 + \sigma_l \lambda_l$	σ_r	σ_l	ρ
0.0	0.876 (0.266)	0.0546 (0.0373)	0.442	0.142	0.530
0.5	0.317 (0.276)	0.0259 (0.0377)	0.448	0.142	0.512
1.0	-0.257 (0.269)	-0.0021 (0.0378)	0.444	0.143	0.508

[a] Indicates the simple linearization drift used: $f = 0$ for the beginning of period drift, $f = 1$ for end of period drift, and $f = 0.5$ for an average of the two.

[b] Standard errors in parentheses.

Table B.14.

Patterns of missing data for 3 different portfolio formation schemes.

	scheme #1[b]			scheme #2			scheme #3		
ptf[a]	*y*	*n*	*N/n*	*y*	*n*	*N/n*	*y*	*n*	*N/n*
1	0-1	144	16.1	0-1	144	16.1	0-1	144	16.1
2	1-2	144	16.6	1-2	144	16.6	1-2	144	16.6
3	2-3	144	9.5	2-3	144	9.5	2-3	144	9.5
4	3-4	144	7.9	3-4	144	7.9	3-4	144	7.9
5	4-5	144	5.1	4-5	144	5.1	4-5	144	5.1
6	5-6	144	3.1	5-6	144	3.1	5-6	144	3.1
7	6-7	142	2.9	6-7	142	2.9	6-7	142	2.9
8	7-8	88	1.8	7-8	88	1.8	7-10	136	4.1
9	8-9	96	1.8	8-9	96	1.8	10-15	132	5.3
10	9-10	108	2.1	9-10	108	2.1	15-20	117	3.2
11	n/a	n/a	n/a	10-20	132	8.1	n/a	n/a	n/a

[a] Portfolio number in the given portfolio formation scheme.

[b] Portfolio formation scheme.
y denotes the number of years to maturity, ie. '0-1' indicates that all bonds with maturities from 0 to 1 year are put into the portfolio.
n is the number of non-missing portfolio observations, ie. months in which there was at least one bond observation in the portfolio. $n = 144$ indicates no missing portfolio observations.
N is the total number of bond observations for this portfolio over the entire test period. Therefore, N/n is the average number of bond observations per each portfolio observation.

Table B.15.

Distance function values for various values of the market price of short rate risk, λ_r, and the reversion coefficient, α, from the stochastic process for the short rate, r. Estimated using month-end data from the period of October 1970 to October 1982.

parameters		scheme #1[a]			scheme #2			scheme #3		
λ_r	α	OLS	HET	GLS	OLS	HET	GLS	OLS	HET	GLS
-0.250	0.800	39,100	248.0	26,300	54,400	291.0	47,700	56,900	275.0	21,100
0.000	0.800	9680	59.9	5710	16,000	77.4	11,100	19,400	80.6	9750
0.250	0.800	2940[b]	27.6	2430	5040[b]	33.5	3550	7420	38.8	4350
0.450	0.800	8970	81.1	3860	9920	83.7	4180	10,500	82.5	3170
0.000	0.600	7550	49.4	4800	13,100	64.8	9680	16,600	69.5	9650
0.000	0.500	6570	46.4	4300	11,600	60.3	8890	15,200	65.8	8890
0.000	0.400	6050	48.3	3790	10,300	60.3	8240	13,900	66.5	7480
0.000	0.300	7040	62.5	4290	10,600	72.3	122,000	14,100	78.7	4930
0.260	0.558	7560	68.9	3200	8660	71.9	3900	9730	72.5	3030[b]
0.200	0.667	3530	33.5	2110[b]	5520	39.0	2840[b]	7670	43.6	3350
0.240	0.840	3000	26.7[b]	2160	5320	33.2[b]	3080	7700	38.2[b]	3770
0.300	0.926	3070	28.4	2140	5180	34.3	2910	7400[b]	39.0	3520

[a] Three different portfolio formation schemes. OLS, HET(=HETERO) and GLS are the three covariance matrix assumptions.

[b] Minimum value of the distance function.

Table B.16.

Minimum distance estimates of λ_r and α for three portfolio formation schemes and three covariance matrix assumptions.

	scheme #1[a]			scheme #2			scheme #3		
parameters	OLS	HET	GLS	OLS	HET	GLS	OLS	HET	GLS
λ_r	0.250	0.240	0.200	0.250	0.240	0.200	0.300	0.240	0.260
α	0.800	0.840	0.667	0.800	0.840	0.667	0.926	0.840	0.558

[a] Three different portfolio formation schemes. OLS, HET(=HETERO) and GLS are the three covariance matrix assumptions.

Table B.17.

Bond option pricing errors (calculated for the three different interpolation methods) by number of months to maturity.

m^a	N	linear[b]		cubic		quadratic	
		$\bar{e}^c$	e_{RMS}	$\bar{e}$	e_{RMS}	$\bar{e}$	e_{RMS}
0-1	742	0.47	1.72	0.10	0.53	0.07	0.53
1-2	800	0.86	1.84	0.27	0.56	0.25	0.55
2-3	804	0.69	1.73	0.41	0.60	0.39	0.58
3-4	683	0.61	1.53	0.40	0.74	0.38	0.73
4-5	391	0.74	1.77	0.41	0.87	0.39	0.86
5-6	180	0.72	1.51	0.48	0.70	0.46	0.69
6-7	96	0.46	1.47	0.45	1.05	0.44	1.04
7-8	77	0.49	1.61	0.05	0.90	0.04	0.90
8-9	20	0.32	1.90	0.94	1.34	0.92	1.34
0-9	3793	0.66	1.70	0.32	0.67	0.30	0.66

[a] m is the number of 'months' to maturity of the option, where a 'month' is actually 30 days, eg. the first row summarizes all observations with 1 to 30 days to maturity.
N is the number of option observations.

[b] The three different interpolation methods.

[c] $\bar{e}$ is the average pricing error.
e_{RMS} is the root mean square pricing error.

Table B.18.

First derivative of bond option prices with respect to r (calculated using linear interpolation) by number of months to maturity.

m[a]	N	linear[b]		
		$V_r^{\min}$	$\overline{V}_r$	$V_r^{\max}$
0-1	742	-67	-1.5	60
1-2	800	-71	-3.0	45
2-3	804	-62	-3.1	33
3-4	683	-60	-1.6	28
4-5	391	-45	-2.5	19
5-6	180	-32	-0.6	14
6-7	96	-38	2.5	16
7-8	77	-43	-3.4	16
8-9	20	-32	2.4	19
0-9	3793	-71	-2.1	60

[a] m is the number of 'months' to maturity of the option, where a 'month' is actually 30 days, eg. the first row summarizes all observations with 1 to 30 days to maturity.
N is the number of option observations.

[b] Linear interpolation method used.

[c] $V_r^{\min}$, $\overline{V}_r$ and $V_r^{\max}$ are the minimum, average and maximum values, respectively, of the partial derivative of option price with respect to r.

Table B.19.

First derivative of bond call option price with respect to l (for the three different interpolation methods) by number of months to maturity.

		linear[b]			cubic			quadratic		
m^a	N	C_l^{min} [c]	$\overline{C}_l$	C_l^{max}	C_l^{min}	$\overline{C}_l$	C_l^{max}	C_l^{min}	$\overline{C}_l$	C_l^{max}
0-1	464	-1040	-440	-10	-1160	-420	60	-1140	-420	50
1-2	529	-1080	-450	-20	-1120	-430	-7	-1110	-430	-2
2-3	555	-1000	-440	-20	-1040	-430	-20	-1030	-430	-20
3-4	466	-870	-430	-50	-890	-410	-30	-890	-420	-30
4-5	252	-990	-410	-80	-1000	-390	-70	-1000	-390	-70
5-6	120	-810	-390	-140	-810	-380	-110	-810	-380	-120
6-7	72	-850	-420	-110	-860	-400	-100	-850	-400	-110
7-8	43	-830	-410	-150	-830	-400	-130	-830	-400	-130
8-9	13	-730	-550	-320	-730	-540	-300	-710	-530	-310
0-9	2514	-1080	-430	-10	-1160	-420	60	-1140	-420	50

[a] m is the number of 'months' to maturity of the option, where a 'month' is actually 30 days, eg. the first row summarizes all observations with 1 to 30 days to maturity.
N is the number of option observations.

[b] The three different interpolation methods.

[c] C_l^{min}, $\overline{C}_l$ and C_l^{max} are the minimum, average and maximum values, respectively, of the partial derivative of call option price with respect to l.

Table B.20.

First derivative of bond put option price with respect to l (for the three different interpolation methods) by number of months to maturity.

m[a]	N	linear[b]			cubic			quadratic		
		$P_l^{\min}$[c]	$\overline{P}_l$	$P_l^{\max}$	$P_l^{\min}$	$\overline{P}_l$	$P_l^{\max}$	$P_l^{\min}$	$\overline{P}_l$	$P_l^{\max}$
0-1	278	20	460	1050	-10	470	1180	-50	470	1150
1-2	271	70	470	1000	30	480	1010	40	480	1010
2-3	249	90	420	960	80	430	990	70	430	980
3-4	217	130	460	930	120	470	970	130	470	960
4-5	139	120	460	930	110	470	950	110	470	950
5-6	60	140	410	780	130	420	810	130	420	800
6-7	24	160	430	810	160	440	830	150	440	830
7-8	34	220	430	800	220	430	810	220	430	820
8-9	7	220	440	530	220	450	550	220	440	540
0-9	1279	20	450	1050	-10	460	1180	-50	460	1150

[a] m is the number of 'months' to maturity of the option, where a 'month' is actually 30 days, eg. the first row summarizes all observations with 1 to 30 days to maturity.
N is the number of option observations.

[b] The three different interpolation methods.

[c] $P_l^{\min}$, $\overline{P}_l$ and $P_l^{\max}$ are the minimum, average and maximum values, respectively, of the partial derivative of put option price with respect to l.

Table B.21.

Adjusted first derivative of bond call option price with respect to l (for the three different interpolation methods) by number of months to maturity.

		linear[b]			cubic			quadratic		
m[a]	N	$C_l^{\min}$[c]	$\overline{C}_l$	$C_l^{\max}$	$C_l^{\min}$	$\overline{C}_l$	$C_l^{\max}$	$C_l^{\min}$	$\overline{C}_l$	$C_l^{\max}$
0-1	464	-1040	-440	-10	-1150	-410	0.7	-1140	-420	-1.6
1-2	529	-1080	-450	-20	-1120	-420	1.1	-1110	-430	-10
2-3	555	-1000	-440	-20	-1040	-430	-20	-1030	-430	-20
3-4	466	-870	-430	-50	-890	-410	-30	-890	-420	-30
4-5	252	-990	-410	-80	-1000	-390	-70	-1000	-390	-70
5-6	120	-810	-390	-140	-810	-380	-110	-810	-380	-120
6-7	72	-850	-420	-110	-860	-400	-100	-850	-400	-110
7-8	43	-830	-410	-150	-830	-400	-130	-830	-400	-130
8-9	13	-730	-550	-320	-730	-540	-300	-710	-530	-310
0-9	2514	-1080	-430	-10	-1150	-410	1.1	-1140	-420	-1.6

[a] m is the number of 'months' to maturity of the option, where a 'month' is actually 30 days, eg. the first row summarizes all observations with 1 to 30 days to maturity.
N is the number of option observations.

[b] The three different interpolation methods.

[c] $C_l^{\min}$, $\overline{C}_l$ and $C_l^{\max}$ are the minimum, average and maximum values, respectively, of the partial derivative of call option price with respect to l.

Table B.22.

Adjusted first derivative of bond put option price with respect to l (for the three different interpolation methods) by number of months to maturity.

m^a	N	linear[b]			cubic			quadratic		
		$P_l^{\min}$ [c]	$\overline{P}_l$	$P_l^{\max}$	$P_l^{\min}$	$\overline{P}_l$	$P_l^{\max}$	$P_l^{\min}$	$\overline{P}_l$	$P_l^{\max}$
0-1	278	20	460	1050	-3.7	360	1010	10	470	1110
1-2	271	70	470	1000	-10	400	1010	40	480	1010
2-3	90	420	960	80	-10	410	860	70	430	980
3-4	130	460	930	120	-20	470	970	130	470	960
4-5	139	120	460	930	-30	440	890	110	470	940
5-6	60	140	410	780	130	420	810	130	420	800
6-7	24	160	430	810	160	440	830	150	440	830
7-8	34	220	430	800	-30	410	810	220	430	820
8-9	7	220	440	530	220	450	550	220	440	540
0-9	1279	20	450	1050	-30	410	1010	10	460	1010

[a] m is the number of 'months' to maturity of the option, where a 'month' is actually 30 days, eg. the first row summarizes all observations with 1 to 30 days to maturity.
N is the number of option observations.

[b] The three different interpolation methods.

[c] $P_l^{\min}$, $\overline{P}_l$ and $P_l^{\max}$ are the minimum, average and maximum values, respectively, of the partial derivative of put option price with respect to l.

Table B.23.

Brennan-Schwartz bond option pricing errors by number of months to maturity.

	calls[b]			puts		
m^a	N	$\bar{e}$	e_{RMS}	N	$\bar{e}$	e_{RMS}
0-1	464	0.09	0.49	278	0.14	0.60
1-2	529	0.31	0.57	271	0.20	0.54
2-3	555	0.43	0.63	249	0.36	0.52
3-4	466	0.40	0.75	217	0.40	0.73
4-5	252	0.41	0.81	139	0.42	0.98
5-6	120	0.51	0.71	60	0.41	0.68
6-7	72	0.38	1.08	24	0.66	0.95
7-8	43	0.00	0.90	34	0.11	0.90
8-9	13	0.71	0.94	7	1.34	1.93
0-9	2514	0.33	0.67	1279	0.30	0.68

[a] Number of 'months' to maturity of the option. where a 'month' is actually 30 days. eg. The first row summarizes all observations with 1 to 30 days to maturity.

[b] N denotes the number of option price observations.
$\bar{e}$ is the average pricing error in dollars.
e_{RMS} is the root mean square pricing error.

Table B.24.

Brennan-Schwartz bond option pricing errors for in and out of the money options.

B/K^a	calls[b]			puts		
	N	$\bar{e}$	e_{RMS}	N	$\bar{e}$	e_{RMS}
0.85-0.98	704	0.18	0.55	270	0.38	0.79
0.98-1.00	687	0.36	0.64	293	0.36	0.68
1.00-1.02	637	0.44	0.77	331	0.31	0.69
1.02-1.15	486	0.38	0.71	385	0.18	0.59

[a] Bond price, B, divided by option exercise price, K.
[b] N denotes the number of option price observations.
$\bar{e}$ is the average pricing error in dollars.
e_{RMS} is the root mean square pricing error.

Table B.25.

Black-Scholes bond option pricing errors by number of months to maturity.

	calls[b]			puts		
m^a	N	$\bar{e}$	e_{RMS}	N	$\bar{e}$	e_{RMS}
0-1	464	0.15	0.50	278	0.16	0.60
1-2	529	0.51	0.74	271	0.37	0.66
2-3	555	0.71	0.88	249	0.59	0.72
3-4	466	0.70	1.19	217	0.75	0.96
4-5	252	0.75	1.06	139	0.83	1.19
5-6	120	0.90	1.05	60	0.94	1.14
6-7	72	0.68	1.27	24	1.21	1.33
7-8	43	0.43	0.95	34	0.88	1.28
8-9	13	0.90	1.12	7	1.72	2.07
0-9	2514	0.57	0.86	1279	0.55	0.86

[a] Number of 'months' to maturity of the option, where a 'month' is actually 30 days. eg. The first row summarizes all observations with 1 to 30 days to maturity.

[b] N denotes the number of option price observations.
$\bar{e}$ is the average pricing error in dollars.
e_{RMS} is the root mean square pricing error.

Table B.26.

Black-Scholes bond option pricing errors for in and out of the money options.

	calls[b]			puts		
B/K^a	N	$\bar{e}$	e_{RMS}	N	$\bar{e}$	e_{RMS}
0.85-0.98	704	0.37	0.69	270	0.53	0.89
0.98-1.00	687	0.60	0.84	293	0.60	0.87
1.00-1.02	637	0.75	1.00	331	0.60	0.88
1.02-1.15	486	0.59	0.88	385	0.48	0.82

[a] Bond price, B, divided by option exercise price, K.

[b] N denotes the number of option price observations.
$\bar{e}$ is the average pricing error in dollars.
e_{RMS} is the root mean square pricing error.

Table B.27.

Brennan-Schwartz bond option pricing errors (using 'in-sample' variance estimates) by number of months to maturity.

	calls[b]			puts		
m[a]	N	$\bar{e}$	e_{RMS}	N	$\bar{e}$	e_{RMS}
0-1	464	0.02	0.48	278	0.08	0.59
1-2	529	0.16	0.49	271	0.07	0.50
2-3	555	0.23	0.50	249	0.19	0.42
3-4	466	0.17	0.65	217	0.19	0.65
4-5	252	0.15	0.71	139	0.18	0.91
5-6	120	0.23	0.56	60	0.15	0.57
6-7	72	0.11	0.99	24	0.39	0.80
7-8	43	-0.31	1.00	34	-0.20	0.91
8-9	13	0.39	0.72	7	1.05	1.75
0-9	2514	0.15	0.58	1279	0.14	0.62

[a] Number of 'months' to maturity of the option, where a 'month' is actually 30 days. eg. The first row summarizes all observations with 1 to 30 days to maturity.

[b] N denotes the number of option price observations.
$\bar{e}$ is the average pricing error in dollars.
e_{RMS} is the root mean square pricing error.

Table B.28.

Brennan-Schwartz bond option pricing errors (using 'in-sample' variance estimates) for in and out of the money options.

B/K[a]	calls[b]			puts		
	N	$\bar{e}$	e_{RMS}	N	$\bar{e}$	e_{RMS}
0.85-0.98	704	0.03	0.50	270	0.26	0.72
0.98-1.00	687	0.16	0.54	293	0.20	0.60
1.00-1.02	637	0.22	0.67	331	0.13	0.63
1.02-1.15	486	0.20	0.62	385	0.01	0.56

[a] Bond price, B, divided by option exercise price, K.

[b] N denotes the number of option price observations.
$\bar{e}$ is the average pricing error in dollars.
e_{RMS} is the root mean square pricing error.

Table B.29.

Brennan-Schwartz treasury bill option pricing errors by number of months to maturity.

	calls[b]			puts		
m^a	N	$\bar{e}$	e_{RMS}	N	$\bar{e}$	e_{RMS}
0-1	62	0.004	0.022	93	0.013	0.023
1-2	79	0.029	0.044	77	0.037	0.049
2-3	80	0.050	0.068	77	0.042	0.053
3-4	84	0.051	0.063	86	0.033	0.061
4-5	39	0.072	0.086	30	0.024	0.061
5-6	16	0.078	0.097	22	0.038	0.058
6-7	11	0.047	0.057	15	0.001	0.091
7-8	16	0.072	0.085	14	0.054	0.082
8-9	8	0.065	0.089	10	0.057	0.065
0-9	395	0.043	0.062	424	0.031	0.053

[a] Number of 'months' to maturity of the option, where a 'month' is actually 30 days. eg. The first row summarizes all observations with 1 to 30 days to maturity.

[b] N denotes the number of option price observations.
$\bar{e}$ is the average pricing error in dollars.
e_{RMS} is the root mean square pricing error.

Table B.30.

Brennan-Schwartz treasury bill option pricing errors for in and out of the money options.

B/K^a	calls[b]			puts		
	N	$\bar{e}$	e_{RMS}	N	$\bar{e}$	e_{RMS}
0.9915-0.9965	77	0.024	0.041	123	0.009	0.042
0.9965-0.9980	100	0.057	0.072	102	0.040	0.057
0.9980-0.9995	133	0.052	0.071	119	0.037	0.059
0.9995-1.0025	85	0.033	0.050	80	0.044	0.054

[a] Treasury bill price, B, divided by option exercise price, K.

[b] N denotes the number of option price observations.
$\bar{e}$ is the average pricing error in dollars.
e_{RMS} is the root mean square pricing error.

Table B.31.

Brennan-Schwartz treasury bill option pricing errors (using 'in-sample' variance estimates) by number of months to maturity.

	calls[b]			puts		
m^a	N	$\bar{e}$	e_{RMS}	N	$\bar{e}$	e_{RMS}
0-1	62	-0.020	0.027	93	-0.002	0.015
1-2	79	-0.019	0.030	77	0.000	0.033
2-3	80	-0.020	0.040	77	-0.010	0.032
3-4	84	-0.024	0.037	86	-0.021	0.052
4-5	39	-0.028	0.044	30	-0.040	0.065
5-6	16	-0.025	0.047	22	-0.031	0.053
6-7	11	-0.065	0.073	15	-0.074	0.113
7-8	16	-0.052	0.067	14	-0.035	0.068
8-9	8	-0.067	0.087	10	-0.029	0.034
0-9	395	-0.025	0.041	424	-0.015	0.045

[a] Number of 'months' to maturity of the option, where a 'month' is actually 30 days. eg. The first row summarizes all observations with 1 to 30 days to maturity.

[b] N denotes the number of option price observations.
$\bar{e}$ is the average pricing error in dollars.
e_{RMS} is the root mean square pricing error.

Table B.32.

Brennan-Schwartz treasury bill option pricing errors (using 'in-sample' variance estimates) for in and out of the money options.

	calls[b]			puts		
B/K^a	N	$\bar{e}$	e_{RMS}	N	$\bar{e}$	e_{RMS}
0.9915-0.9965	77	-0.024	0.036	123	-0.009	0.045
0.9965-0.9980	100	-0.028	0.037	102	-0.009	0.042
0.9980-0.9995	133	-0.028	0.049	119	-0.023	0.054
0.9995-1.0025	85	-0.026	0.034	80	-0.022	0.033

[a] Treasury bill price, B, divided by option exercise price, K.

[b] N denotes the number of option price observations.
$\bar{e}$ is the average pricing error in dollars.
e_{RMS} is the root mean square pricing error.

Table B.33.

Bond option arbitrage strategy number 1. Comparison of Brennan-Schwartz and Black-Scholes average cumulative dollar returns for different levels of commissions.

commission[a]		Brennan-Schwartz[b]			Black-Scholes			difference[c]	
option	bond	n	$\bar{d}_c$	t_d	n	$\bar{d}_c$	t_d	$\bar{d}_c$	t_d
0%	0%	248	0.771	6.6	248	0.447	3.9	0.32	2.0
3	0	248	0.266	2.5	248	0.022	0.2	0.24	1.6
0	1/4	248	-0.578	-5.2	248	-0.409	-3.9	-0.17	-1.1
3	1/4	248	-1.083	-8.8	248	-0.834	-7.8	-0.25	-1.5

[a] Commissions charged for trading in options and bonds to rebalance the zero-investment arbitrage portfolio.
Note that bonds are bought at the ask price quoted in the Wall Street Journal and sold at the bid price.

[b] n denotes the number of options.
$\bar{d}_c$ is the average ***cumulative*** dollar arbitrage return.
t_d is the t-statistic of $\bar{d}_c$.

[c] Difference between Brennan-Schwartz and Black-Scholes cumulative returns.

Table B.34.

Bond option arbitrage strategy number 1. Brennan-Schwartz (using 'in-sample' variances) average cumulative dollar returns for different levels of commissions.

commission[a]		Brennan-Schwartz[b] ('in-sample' variance)		
option	bond	n	$\bar{d}_c$	t_d
0%	0%	248	1.113	9.5
3	0	248	0.577	5.5
0	1/4	248	-0.356	-0.3
3	1/4	248	-0.892	-7.3

[a] Commissions charged for trading in options and bonds to rebalance the zero-investment arbitrage portfolio.
Note that bonds are bought at the ask price quoted in the Wall Street Journal and sold at the bid price.

[b] n denotes the number of options.
$\bar{d}_c$ is the average *cumulative* dollar arbitrage return.
t_d is the t-statistic of $\bar{d}_c$.

Table B.35.

Bond option arbitrage strategy number 2. Comparison of Brennan-Schwartz and Black-Scholes average (non-cumulative) dollar returns for various filter levels.

	Brennan-Schwartz[b]			Black-Scholes			difference[c]	
f^a	N	$\bar{d}$	t_d	N	$\bar{d}$	t_d	$\bar{d}$	t_d
0.00	3510	0.092	11.1	3504	0.056	6.7	0.04	3.0
0.30	2140	0.139	12.2	2552	0.080	7.8	0.06	3.9
0.50	1467	0.181	11.9	2015	0.093	8.0	0.09	4.6
0.70	959	0.237	11.3	1550	0.103	7.3	0.13	4.1
1.00	425	0.379	9.4	928	0.154	7.6	0.22	5.0

[a] Filter level in dollars. The zero-investment arbitrage portfolio was only formed if theoretical and market option prices differed by at least the filter amount. Note that no commissions are charged and bonds are bought and sold at the middle of the bid-ask spread quoted in the Wall Street Journal.

[b] N denotes the number of observations.
$\bar{d}$ is the average (non-cumulative) dollar arbitrage return.
t_d is the t-statistic for $\bar{d}$.

[c] Difference between the Brennan-Schwartz and Black-Scholes returns.

Table B.36.

Bond option arbitrage strategy number 2. Brennan-Schwartz (using 'in-sample' variances) average (non-cumulative) dollar returns for various filter levels.

	Brennan-Schwartz[b] ('in-sample' variance)		
f^a	N	$\bar{d}$	t_d
0.00	3511	0.119	14.6
0.30	1826	0.191	14.9
0.50	1159	0.240	13.2
0.70	695	0.342	12.5
1.00	285	0.565	10.2

[a] Filter level in dollars. The zero-investment arbitrage portfolio was only formed if theoretical and market option prices differed by at least the filter amount. Note that no commissions are charged and bonds are bought and sold at the middle of the bid-ask spread quoted in the Wall Street Journal.

[b] N denotes the number of observations.
$\bar{d}$ is the average (non-cumulative) dollar arbitrage return.
t_d is the t-statistic for $\bar{d}$.

Table B.37.

Bond option arbitrage strategy number 2. Comparison of Brennan-Schwartz and Black-Scholes average (non-cumulative) dollar returns for different levels of commissions.

commission[a]		Brennan-Schwartz[b]			Black-Scholes			difference[c]	
option	bond	N	$\bar{d}$	t_d	N	$\bar{d}$	t_d	$\bar{d}$	t_d
0%	0%	3510	0.059	7.5	3504	0.035	4.4	0.02	2.1
3	0	3510	0.030	4.0	3504	0.012	1.5	0.02	1.7
0	1/4	3510	-0.022	-2.9	3504	-0.016	-2.0	-0.01	-0.6
3	1/4	3510	-0.051	-6.6	3504	-0.039	-5.1	-0.01	-1.1

[a] Commissions charged for trading in options and bonds to rebalance the zero-investment arbitrage portfolio.
Note that bonds are bought at the ask price quoted in the Wall Street Journal and sold at the bid price.

[b] N denotes the number of observations.
$\bar{d}$ is the average (non-cumulative) dollar arbitrage return.
t_d is the t-statistic for $\bar{d}$.

[c] Difference between the Brennan-Schwartz and Black-Scholes returns.

Table B.38.

Bond option arbitrage strategy number 2. Brennan-Schwartz (using 'in-sample' variances) average (non-cumulative) dollar returns for different levels of commissions.

commission[a]		Brennan-Schwartz[b] ('in-sample' variance)		
option	bond	N	$\bar{d}$	t_d
0%	0%	3511	0.083	10.7
3	0	3511	0.052	7.0
0	1/4	3511	-0.007	-0.9
3	1/4	3511	-0.038	-5.0

[a] Commissions charged for trading in options and bonds to rebalance the zero-investment arbitrage portfolio.
Note that bonds are bought at the ask price quoted in the Wall Street Journal and sold at the bid price.

[b] N denotes the number of observations.
$\bar{d}$ is the average (non-cumulative) dollar arbitrage return.
t_d is the t-statistic for $\bar{d}$.

Table B.39.

Treasury bill option arbitrage strategy number 1. Brennan-Schwartz average cumulative dollar returns for different levels of commissions.

commission[a]		Brennan-Schwartz[b]		
option	tbill	n	$\bar{d}_c$	t_d
0%	0%	31	0.1521	1.0
3	0	31	0.1103	0.7
0	1/4	31	-0.8370	-3.4
3	1/4	31	-0.8789	-3.4

[a] Commissions charged for trading in options and treasury bills to rebalance the zero-investment arbitrage portfolio.
Note that treasury bills are bought at the ask price quoted in the Wall Street Journal and sold at the bid price.

[b] n denotes the number of options.
$\bar{d}_c$ is the average *cumulative* dollar arbitrage return.
t_d is the t-statistic of $\bar{d}_c$.

Table B.40.

Treasury bill option arbitrage strategy number 1. Brennan-Schwartz (using 'in-sample' variances) average cumulative dollar returns for different levels of commissions.

commission[a]		Brennan-Schwartz[b] ('in-sample' variance)		
option	tbill	n	$\bar{d}_c$	t_d
0%	0%	30	0.2027	1.1
3	0	30	0.1572	0.8
0	1/4	30	-1.1399	-3.7
3	1/4	30	-1.1854	-3.8

[a] Commissions charged for trading in options and treasury bills to rebalance the zero-investment arbitrage portfolio.
Note that treasury bills are bought at the ask price quoted in the Wall Street Journal and sold at the bid price.

[b] n denotes the number of options.
$\bar{d}_c$ is the average *cumulative* dollar arbitrage return.
t_d is the t-statistic of $\bar{d}_c$.

Table B.41.

Treasury bill option arbitrage strategy number 2. Brennan-Schwartz average (non-cumulative) dollar returns for various filter levels.

f^a	Brennan-Schwartz[b]		
	N	$\overline{d}$	t_d
0.00	777	0.0083	2.1
0.03	477	0.0166	3.0
0.05	301	0.0281	3.5
0.07	176	0.0465	4.1

[a] Filter level in dollars. The zero-investment arbitrage portfolio was only formed if theoretical and market option prices differed by at least the filter amount. Note that no commissions are charged and treasury bills and bonds are bought and sold at the middle of the bid-ask spread quoted in the Wall Street Journal.

[b] N denotes the number of observations.
$\overline{d}$ is the average (non-cumulative) dollar arbitrage return.
t_d is the t-statistic for $\overline{d}$.

Table B.42.

Treasury bill option arbitrage strategy number 2. Brennan-Schwartz ('in-sample' variance) average (non-cumulative) dollar returns for various filter levels.

f^a	Brennan-Schwartz[b] ('in-sample' variance)		
	N	$\overline{d}$	t_d
0.00	738	0.0112	2.4
0.03	274	0.0220	2.2
0.05	122	0.0485	2.4
0.07	79	0.0788	2.7

[a] Filter level in dollars. The zero-investment arbitrage portfolio was only formed if theoretical and market option prices differed by at least the filter amount. Note that no commissions are charged and treasury bills and bonds are bought and sold at the middle of the bid-ask spread quoted in the Wall Street Journal.

[b] N denotes the number of observations.
$\overline{d}$ is the average (non-cumulative) dollar arbitrage return.
t_d is the t-statistic for $\overline{d}$.

Table B.43.

Treasury bill option arbitrage strategy number 2. Brennan-Schwartz average (non-cumulative) dollar returns for different levels of commissions.

commisions[a]		Brennan-Schwartz[b]		
option	tbill	N	$\overline{d}$	t_d
0%	0%	777	0.0063	1.6
3	0	777	0.0048	1.2
0	1/4	777	-0.0297	-6.3
3	1/4	777	-0.0311	-6.5

[a] Commissions charged for trading in options and treasury bills to rebalance the zero-investment arbitrage portfolio.
Note that treasury bills are bought at the ask price quoted in the Wall Street Journal and sold at the bid price.

[b] N denotes the number of observations.
$\overline{d}$ is the average (non-cumulative) dollar arbitrage return.
t_d is the t-statistic for $\overline{d}$.

Table B.44.

Treasury bill option arbitrage strategy number 2. Brennan-Schwartz ('in-sample' variance) average (non-cumulative) dollar returns for different levels of commissions.

commisions[a]		Brennan-Schwartz[b] ('in-sample' variance)		
option	tbill	N	$\bar{d}$	t_d
0%	0%	738	0.0085	1.8
3	0	738	0.0069	1.5
0	1/4	738	-0.0422	-7.5
3	1/4	738	-0.0438	-7.6

[a] Commissions charged for trading in options and treasury bills to rebalance the zero-investment arbitrage portfolio.
Note that treasury bills are bought at the ask price quoted in the Wall Street Journal and sold at the bid price.

[b] N denotes the number of observations.
$\bar{d}$ is the average (non-cumulative) dollar arbitrage return.
t_d is the t-statistic for $\bar{d}$.