UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

MULTIFACTOR VALUATION MODELS OF ENERGY FUTURES AND OPTIONS ON FUTURES

A Dissertation

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

degree of

Doctor of Philosophy

By

MARK J BERTUS

Norman, Oklahoma

2003

UMI Number: 3082955

ProQuest Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346

MULTIFACTOR VALUATION MODELS OF ENERGY FUTURES AND OPTIONS ON FUTURES

A Dissertation APPROVED FOR THE

MICHAEL F. PRICE COLLEGE OF BUSINESS

BY

TABLE OF CONTENTS

List of Tables

List of Illustrations

Abstract

The intent of this dissertation is to investigate continuous time pricing models for commodity derivative contracts that consider mean reversion. The motivation for pricing commodity futures and option on futures contracts leads to improved practical risk management techniques in markets where uncertainty is increasing. In the dissertation closed-form solutions to mean reverting one-factor, two-factor, three-factor Brownian motions are developed for futures contracts. These solutions are obtained through risk neutral pricing methods that yield tractable expressions for futures prices, which are linear in the state variables, hence making them attractive for estimation. These functions, however, are expressed in terms of latent variables (i.e. spot prices, convenience yield) which complicate the estimation of the futures pricing equation. To address this complication a discussion on Dynamic factor analysis is given. This procedure documents latent variables using a Kalman filter and illustrations show how this technique may be used for the analysis. In addition, to the futures contracts closed form solutions for two option models are obtained. Solutions to the one- and two-factor models are tailored solutions of the Black-Scholes pricing model. Furthermore, since these contracts are written on the futures contracts, they too are influenced by the same underlying parameters of the state variables used to price the futures contracts. To conclude, the analysis finishes with an investigation of commodity futures options that incorporate random discrete jumps.

Chapter 1

Introduction

Recently, energy markets have exhibited turbulent behavior in energy prices. Crude oil prices tripled over the last three years, natural gas prices increased 350% and heating oil was selling over $100 per barrel. The volatile market conditions have caused an increased interest in the use of financial instruments linked to the price of commodities, such as futures, options on futures and commodity-linked bonds. These instruments are the main vehicles for hedging price risk, and the ability to price these instruments is becoming an increasingly important problem in financial economics. The purpose of this study is to investigate the pricing ability of continuous time models for commodity futures contracts and options on commodity futures.

Early commodity asset pricing models (see Schwartz (1982), and Brennan and Schwartz (1985)) assume all the uncertainty in a commodity's futures prices is summarized by the commodity's spot price. The pricing solution to these one-factor models is the well known cost of carry formula: $J(t,\tau) = e^{(r-\delta)\tau}S(t)$. The formula states the price of a futures/forward contract is a function of r, the spot rate of interest, δ, the spot (marginal) convenience yield, and $S(t)$, the spot price. Empirical evidence (Fama and French (1987), Brennan (1991), and Bessembinder et al (1995)) suggest that modeling the commodity spot price as a geometric Brownian motion is not so unreasonable, whereas assuming a non-stochastic convenience yield is questionable. Gibson and Schwartz (1990), develop a more realistic model for commodity futures contracts by including a second factor, a stochastic convenience yield. Their two-factor model presumes futures prices are a function of a geometric Brownian motion spot price and mean reverting convenience yield. Schwartz (1997) extends the two-factor model by introducing a third stochastic factor, the instantaneous interest rate. Hilliard and Reis (1998) offer a new pricing equation for Schwartz (1997) by introducing jumps in the risk-adjusted spot price of the commodity.

For these three models, closed form solutions are obtained. These closed form solutions greatly simplify the comparative statics, and yield precise specifications ready for estimation. There is, however, one problem with estimating the pricing equations for commodity futures contracts. That is, frequently the factors or state variables are not directly observable. Futures contracts are traded on several exchanges and their prices are easily observed. Commodities do not trade on organized exchanges making it difficult to observe the actual spot prices. In addition, the spot prices in many cases are so uncertain that the corresponding futures contract closest to maturity is used as a proxy. The instantaneous convenience yield, which is not a traded good, is even more difficult to observe. As a result, standard econometric techniques cannot be applied to futures contracts written on commodities.

Dynamic factor analysis is an empirical tool well suited to handle the latent variable estimation problem. This statistical procedure is an expectation maximization algorithm (EM) that uses the state space form so that the Kalman filter component can be used.[1] The Kalman methodology generates conditional forecasts of a latent stochastic process at various points in time. The Kalman output for the unobserved state variable is then used in a maximum likelihood estimation, so the parameters for the model can be acquired. In short, the EM algorithm is a powerful technique capable of documenting the time series behavior of a latent process, as well as, the parameter estimates for a futures pricing equation.

In addition to commodity futures, market participants also use options on futures contracts as hedging instruments, and thereby, have a growing interest in pricing them. It is natural to ask why people choose to trade options on commodity futures rather than options on the underlying commodity. The main reason appears to be that a futures contract is, in many circumstances, more liquid and easier to trade than the underlying commodity. Furthermore, a futures price is known immediately from trading on the futures exchange, whereas the spot price of the commodity is not so readily available.

In the finance literature, models for commodity futures options are natural extensions of the original work of Black and Scholes (1973) and Merton (1973). The earliest work for pricing options on futures is presented by Black (1976). Black presumes commodity prices behave identically with a stock which pays a constant dividend yield, and his model is therefore analogous to the Black-Scholes model. The problem with Black's work is that we know from current futures pricing models that economists find the assumption of a non-stochastic convenience yield to be too restrictive. In light of this existing research, Hilliard and Reis (1998) and Miltersen and Schwartz (1998) develop option models that allow for a stochastic convenience yield. Miltersen and Schwartz's approach uses the methodology of the Heath-Jarrow-Morton (1991), whereas, Hilliard and Reis use a traditional equilibrium approach. Furthermore, Hilliard and Reis (1998) also investigate a commodity futures option model that allows for discrete random jumps. The authors observe that commodity prices may not fit the presumed dynamics of the Black-Scholes model and may therefore not be an accurate model for pricing options on futures contracts. Their model subsumes the work of Merton (1976) and Bates (1988, 1991, and 1996) and applies it to commodity futures options. In this dissertation, we investigate and develop pricing models for futures options given the existing futures models discussed in the literature.

The analysis is organized as follows. In chapter 2, the market fundamental for oil, natural gas and heating oil are described. Derivations for the proposed futures pricing models are presented in chapter 3. Chapter 4 develops the methodology for pricing options written on commodity futures contracts. Chapter 5 presents simulations of the theoretical prices for both the futures and option models developed in the dissertation and chapter 6 concludes the analysis.

Endnotes

[1] The Kalman Filter technique is well treated in the control literature and the interested reader is referred to Appendix A for an in depth discussion. In addition, other references for the Kalman Filter can be found in Anderson and Moore (1979), Harvey (1989), or Hamilton (1991) for details.

Chapter 2

World Crude Oil Market

A fundamental tenet in economics posits that spot prices are the outcome of the intersection between market supply and demand schedules. Over the last six years, idiosyncratic supply and demand shocks from individual energy markets have created turbulent behavior in their respective spot prices. Noted in chapter 1, crude oil prices tripled over the last three years, natural gas prices increased 350% and heating oil was selling over $100 per barrel. These volatile market conditions have caused an increased interest in the use of financial instruments linked to the price of commodities, such as futures, options on futures and commodity-linked bonds. These instruments are the main vehicles for hedging price risk, and the ability to price these instruments is becoming an increasingly important problem in financial economics. The intent of this chapter is to better understand the time-series behavior of energy spot prices as it pertains to valuing contingent claims written on the energy commodities. In particular, I detail the economic characteristics for the crude oil spot market.[1]

2.1 Crude Oil

In a general it can be said that crude oil, a nonrenewable natural resource, is the world's most consumed energy resource. Individuals who rely highly on energy for transportation, power and heat are naturally the largest suitors for crude oil. Globally, these consumers are the industrialized countries, or the OECD countries. In addition, individuals who have a natural endowment of oil produce oil. Crude producers are located all over the globe and the volume of production from region to region is conditional on the logistics of a particular region. In the following analysis, we detail the supply, inventory, demand, and their impact on the world's crude oil spot prices

2.2 *Supply*

Prior to the 1800's, crude oil production was only a fraction of today's supply. This was due impart to the available sources of oil at that time. That is, the world's crude oil supply originated only from small isolated pools, where oil seeped to the Earth's surface from underground reservoirs. With limited technology, gathering and distributing oil from these reservoirs was an arduous and timely process, which restricted a producer's ability to supply oil. During the 1800's, however, the world's crude oil supply underwent dramatic changes. It started in 1859, when Edwin L. Drake successfully drilled the first underground oil well. The discovery marked the first time an individual had produced oil from an underground oil reserve. Soon after the discovery, these producers realized that these underground reserves were not only greater in volume than existing above ground reserves but located around the globe as well. As a result, more and more reserves were tapped and the supply of world crude oil flourished.

Today, the production of crude oil originates from several regions around the globe. Oil production from these regions depends on their geological attributes. For example, identifying and extricating oil from land-based reservoirs is easier and less expensive than a identifying and extricating oil from a reserve found at the bottom of the sea. Additionally, underground oil formations, where the oil is concentrated in pools rather than diffused throughout the rock formation, is also easier to find and produce. Therefore the exploration and extraction costs for land-based formations, of concentrated pools, are in general considerably lower than sea based reserves with diffused oil. Consequently, some regions will have a natural production advantage over others.

Globally, crude oil production can be dichotomously categorized into OPEC (Organization of Petroleum Exporting Countries) and non-OPEC output. OPEC is an oil cartel of eleven member nations (Algeria, Indonesia, Iran, Iraq, Kuwait, Libya, Nigeria, Saudi Arabia, the United Arab Emirates, and Venezuela) that started in 1960.[2] Total oil production from these eleven member nations constitutes OPEC production. Moreover, non-OPEC production is the total output from the world's remaining oil producing countries. Upon casual observation we see that there are fundamental differences in the production for these two production groups.

The eleven member cartel enjoys a comparative advantage in oil production. The cartel claims eighty percent of the World's proven oil reserves[3], which is roughly 814 billion barrels of oil. Most of these oil reserves are characterized as large land based concentrated pools of oil, where the marginal cost of producing the oil ranges between $1.50 a barrel to $5.00 a barrel of oil. Furthermore, as testimony to their production ability, experts estimate that OPEC could produce oil continuously for the next eighty years.[4] This is significant for experts also claim that all non-OPEC producers could only produce oil continuously for the next fifteen years. Simply non-OPEC producers do not possess the production ability that OPEC has.

Realizing their comparative advantage in the world oil market, the cartel works collusively in an effort to safeguard revenues. That is, the cartel tries to influence oil prices through organized production quotas. With its production quotas in place, OPEC's contribution to the world's oil supply has been roughly forty percent over the last nine years. As OPEC limits its oil production, the cartel enjoys relatively higher oil prices. Introductory economic theory suggests that in the face of higher prices, non-OPEC production should increase because it is now economical to do so. The reality, however, is that most non-OPEC production cannot significantly augment its oil production because they are either running at near capacity or it is simply too costly to produce. This inability of the non-OPEC producers to meet increased demand gives OPEC, in part, the ability to control world crude oil prices.

While OPEC controls a vast portion of the world's oil, they are not the only source of oil. Over the last nine years sixty percent of the world's annual production comes from non-OPEC sources and in total, non-OPEC producers own twenty percent of the world's proven oil reserves. The majority of non-OPEC production originates from three distinct regions: North America, Asia, and Europe. Logistically, non-OPEC production is unlike OPEC production. That is, the oil reserves are not all land-based, and the oil is typically diffused throughout the ground. Exploring and drilling for oil is therefore on average more costly than OPEC production. The marginal costs average between $10 a barrel to $20 a barrel of oil, and depending on the price of oil, some wells are currently unprofitable to produce.

Regionally, the world's largest oil producer outside OPEC is North America. The region holds sixty-seven billion barrels in proven reserves and produces roughly twelve million barrels a day. The United States, the second largest oil producing country in the world, accounts for almost 60 percent of the North American production followed by Mexico and Canada respectively. Following North American production is Asia where the primary source for Asian oil is the former Soviet Union. The former Soviet Union owns sixty billion barrels of oil reserves and produces nine million barrels a day. Russia ranks third in world oil production and is the largest oil exporter outside of OPEC. Other notable Asian producers are China and India, who contribute three million and less than a million barrels a day, respectively. North Sea production, off the coast of the United Kingdom and Norway, is Europe's main artery for oil production. Oil production between Norway and the United Kingdom is evenly split, with production totaling roughly six million barrels a day.

2.3 *Inventories*

In addition to the production of the world crude oil, portions of the global oil supply are not consumed, but held in stock. There are two categories of oil reserves, *discretionary* stocks and *strategic* stocks. The difference between the two categories is ownership. Discretionary stocks are inventories held by industry participants, and strategic stocks are reserves held by governments. The Energy Intelligence Group[5] estimates there are seven to eight billion barrels of oil held in total reserve with the majority held in strategic reserves.

Governments hold strategic inventories for precautionary reasons. That is, governments, mainly of the industrialized countries, hold strategic reserves to protect themselves from adverse price shocks due abrupt supply shortages. The United States owns the largest strategic oil reserve, and started building reserves in 1975 with the Energy Conservation Policy Act (EPCA). The EPCA commits up to four billion barrels of crude oil to be held in reserve. Other industrialized countries, who also wish to dampen the impact of price spikes, hold similar reserves. In total, roughly ninety percent of the world's oil storage is held in strategic reserves.

The remaining ten percent of the world's oil inventories are held in discretionary stocks. Although minor in volume, discretionary stocks provide an important function for the world's oil market. First, discretionary stocks help smooth market disturbances between market supply and demand imbalances. Second, discretionary stocks also reveal valuable information about current and future market conditions/prices. Current market conditions are monitored by the volume of crude oil through these storage facilities. For instance, if there is greater inflow than outflow, then participants witness greater production than demand and inventories will rise. Market participants notice the excess supply and prices adjust accordingly. In addition, market participants make informed forecasts about future market conditions from existing inventory levels. For example, if there is a low level of inventories going into a peak consumption period then expected future spot prices will ratchet upward in anticipation of the scarce supply.

On the whole, global crude oil stocks historically follow seasonal variations. These inventories are typically drawn down in the winter and rebuilt in the spring. In the winter, the need for heating fuels (heating oil, propane, and kerosene) in the Northern Hemisphere increases. At current production rates, the increased need for these middle distillates causes pressure on the current oil supply. To ease the burden on current production, crude oil stocks are drawn down causing storage levels to fall during the winter months. When the heating season passes and the demand for oil subsides these inventories are usually replenished. In short, this pattern creates a tendency for world oil prices to be high in the fall and low in the spring.

2.4 *Demand*

In similar form to supply the world's demand for oil also underwent dramatic changes. The 1800's were marked with new discoveries and technological advances that lead to increase the market for crude oil. In 1880, scientists found that the molecular composition of crude oil was simply a mixture of hydrocarbons. The discovery enabled scientists to separate oil into finer parts. That is, scientist used a sophisticated distillation process, which progressively heated crude oil, to produce different grades of fuel at different temperatures. These grades of fuel (natural gas, gasoline, diesel, heating oil, and asphalts) are called distillates. These distillates are distinguished by their viscosities (viscosity is the property of resistance to flow in a fluid or semifluid). Gasoline is less viscous than diesel, which is less viscous than heating oil. The viscosity of the various distillates dictates how fast these fuels burn and how they may by used. Thus, the scientists' discovery created an even greater uses for oil and thus, greater demand.

Furthermore, this discovery was timely because in 1889 two Germans, Gottlieb Daimler and Wilhelm Maybach, invented the first combustible engine (automobile). This new invention ran on straight run gasoline or kerosene, thereby creating a new use for oil. In addition, further refinements to the combustible engine and the mass production of the automobile enhanced the need for oil. Inevitably, as transportation became more widespread the need for oil became more and more pronounced.

The discussion above illustrates that the global need for energy directly influences the demand for crude oil. Industrialized countries produce the greatest amount of economic activity, and thereby, have the greatest need for energy. As expected, the countries of the Organization for Economic Cooperation and Development (OECD) consume two thirds of the world's daily oil consumption. The primary use of oil in these

OECD countries is for transportation, heating and power. The United States is responsible for twenty five percent of the world's daily oil consumption. The United States depends on private vehicles to travel relatively long distances, and uses more oil for transportation than for heat and power. Per capita, the United States stands alone in its consumption of oil, whereby the average individual consumes almost 3 gallons per day. Japan, the world's second largest consumer, utilizes nine percent of the world's oil. European consumption is evenly spread among the nations, and amounts to eight percent of world's daily oil consumption. Transportation in Europe and Asia is far less than the United States and oil is used more for power generation and heating. The per capita consumption rates reflect the lower oil utilization and equal 1.4 gallons per day. Although these consumption rates are lower than the United States, they significantly exceed the consumption of developing countries, which typically run 0.2 gallons per day per capita.

Discussed earlier, the demand for oil exhibits seasonal trends. The reason stems from the regional concentration of the OECD countries. Globally most OECD countries are located in North America, Asia and Europe, all of which are in the Northern Hemisphere. Thus during the Northern Hemisphere's cold winter months all industrialized countries have a collective need for heat, thereby causing significant increases for the demand of oil. In general, there is a swing of three to four million barrels per day (some 5 percent) between the 4th quarter of the year, when demand is the highest, to the 3rd quarter, when it is the lowest. Again, as a result, crude prices tend to peak during the winter months and fall off in the summer.

2.5 *Prices*

Throughout the latter part of the nineties, world crude oil prices[6] were subject to turbulent price swings. For instance, the thirty year historical average price of a barrel of West Texas Intermediate (WTI) is $19, and over a four year period prices for WTI crude oil were as low as $11 per barrel and as high as $38 a barrel (see Figure 1, page 17). These price swings are attributed to changes in market demand caused by abnormal weather, and market supply shocks.

Starting in 1996, crude oil prices were $22.50 a barrel. The price of oil was above its historical average of $19 due in part to a hard winter that increased the need for oil. But, beginning in 1997 the world economy witnessed the start of a global recession brought on by what economists labeled "the Asian financial crisis". The Asian financial crisis started with Thailand commercial banking system. Thai banks issued loans to many risky Asian corporations who were eventually defaulted on their loans. Default on these loans caused the Thailand banks to fold, which in turn, adversely impacted Thailand's economy. This financial effect propagated to other Asian countries affecting their economies and pushing them into a recession as well. Not everyone felt the financial distress from Asia. Some countries like the United States and Germany continued to experience high economic output, but on the whole, total world output declined. In response to falling global demand, world crude oil prices started to collapse.

Independent of the global recession, the weather during 1997 and 1998 contributed further to reduce the demand for oil. Temperatures during the 1997 and 1998 winter periods were abnormally high. These aberrant conditions led to unusually low oil consumption during the peak consumption period. The industrialized countries simply had little need for heat and so little need for oil.

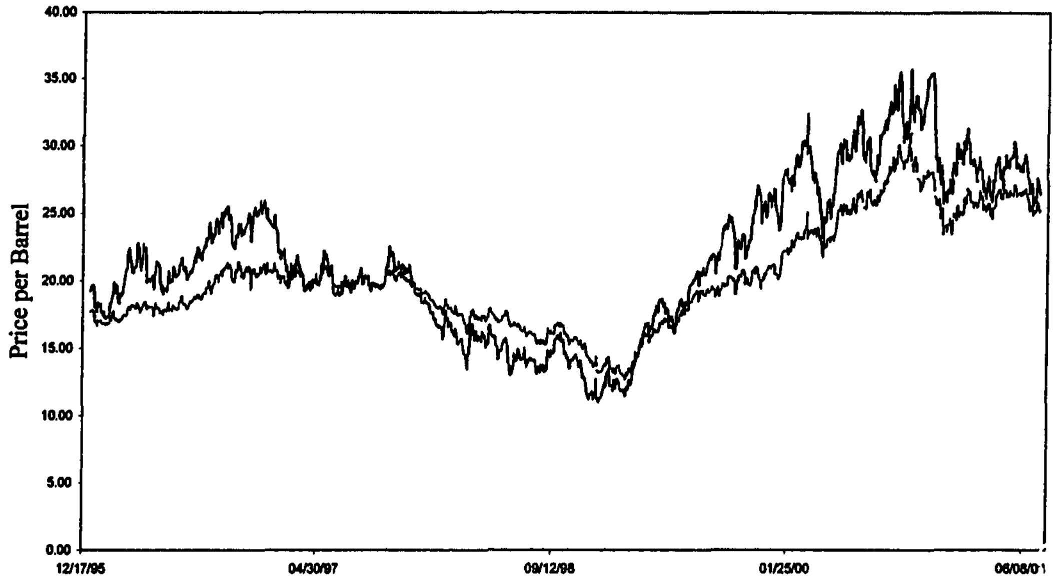
In addition to the demand shocks discussed above, there were shocks to production that contributed to the price collapse as well. These production shocks were the result of OPEC nations exceeding their production quotas. During the price decline, all the OPEC nations were losing revenues, and in order to recoup their losses, the cartel members covertly exceeded their quotas. The result was a world oil glut. Prices continued to fall and it was not until March 1998 before OPEC met and decided to collectively cut production. At this time, however, these initial production cuts were offset when Iraqi oil was allowed to trade under United Nations Security Council Resolution 986. By January 1999 global crude oil prices fell to a record level of $11 a barrel.

In April 1999, the global economy started showing signs of recovery from the Asian financial crisis. In reaction to increased world demand, the price of crude oil started increasing as well. In fact, crude oil prices actually tripled over the next year and a half. Over this time period, OPEC production cuts remained intact and the world's oil supply became inadequate to meet the growing need for world oil. OPEC production cuts and increased world demand, however, were only partially responsible for the dramatic price increase.

Weather conditions prior to 1999 and low inventory levels also contributed to the upswing in crude oil prices. The winters prior to 1999 created acute shortages in middle distillates. In fact, petroleum stocks for industrialized countries were at their lowest levels since the middle 1980's. Anticipating the upcoming peak consumption months for heating oil and diesel, and with petroleum stocks at all time lows, oil companies increased runs in their refineries to fill the petroleum reserves. With the current world oil production fixed, crude oil storage withdrawals were necessary to mitigate any price spikes due to the sudden increases in the demand for oil. Consequently, increased distillate production caused shortages in oil stocks for the OECD countries leading up to the winter of 2000. Starting the 2000 winter season crude stocks were below average and remain below average throughout the entire winter. In addition to the low storage levels, the winter temperatures for 2000 were abnormally low. The cold temperatures caused atypically high demand forcing the price of crude upward. The world price of crude oil spiked to an all time high of $38 a barrel.

Figure 1

This graph displays the time series for the WTI futures contracts listed on NYMEX. The bold line is for the nearby futures contract and the lighter line is the 18-month contract.



Endnotes

[1] All information and statistics were found from the Department of Energy's Energy Information Administration (EIA) website. This site is located at www.eia.doe.gov.

[2] The Organization of Petroleum Exporting Countries (OPEC) was formed at a conference held in Baghdad on September 10-14, 1960. There were five original members: Iran, Iraq, Kuwait, Saudi Arabia, and Venezuela. Between 1960 an 1975, the organization expanded to 13 members with the additions of Qatar, Indonesia, Libya, United Arab Emirates, Algeria, Nigeria, Ecuador, and Gabon. Currently OPEC consists of eleven members with Ecuador and Gabon withdrawing their membership in 1992 and 1995 respectively.

[3] Proven reserves are the known reserves stocks. For example, reserves in Russia and China exist but the degree in which they exist in uncertain. These levels are not proven and thus not reported.

[4] These figures are the production-to-reserve ratios reported by the EIA.

[5] Energy Intelligence Group is an independent information company that specializes in providing the highest quality business intelligence on the global oil and gas industries. There internet site is www.energyintel.com.

[6] Fundamentally, no single price for world crude oil exists. Market participants, instead, monitor spot market activity through three different benchmarks, OPEC basket, North Sea Brent crude and West Texas Intermediate (WTI). The OPEC basket made up of seven crude oils: Algeria's Saharan Blend, Indonesia's Minas, Nigeria's Bonny Light, Saudi Arabia's Arab Light, Dubai's Fateh, Venezuela's Tia Juana Light, and Mexico's

Isthmus. Brent Crude and WTI crude are both light sweet crude oils unlike the OPEC crudes, which tend to be sour (higher sulphur content), and are generally priced more than the OPEC crudes. In addition to the high quality grade, both WTI and Brent crude are appealing benchmarks because they also trade on futures markets. WTI futures trade on the New York Mercantile Exchange (NYMEX), while Brent futures trade at the International Petroleum Exchange (IPE). Organized trading for these benchmarks enhances the ability for market participants to monitor spot market trading.

Chapter 3

Commodity Forward and Futures Pricing Models

The focus of the present study is to compare models of the stochastic behavior of commodity prices in terms of their ability to value existing derivative contracts. The purpose of this chapter is to intuitively derive three futures pricing models. We start with a simple one-factor model, where we presume that all the variability in a commodity's futures price is determined entirely by a mean reverting spot price. Since the futures contract derives its value from an underlying commodity this model should work reasonably well. Modeling spot prices this way, however, we may be too restrictive in our presumption about movements in the spot price. We have no reason to believe that all commodity spot prices follow a mean reverting process. From the finance literature, (see Bessembinder et al (1995), Fama and French (1987), (1988)), there is evidence that inventories follow seasonal trends. With this in mind, we look at a second model where we consider the futures price to be a function of the spot price and a stochastic convenience yield. The convenience yield as defined by Brennan and Schwartz (1985) is the flow of services net of storage costs that accrues to an owner of the physical commodity, but not to the owner of a contract for future delivery.[1] Thus, if inventories oscillate seasonally then so will the convenience yield. For our purpose, the two-factor model should bring greater flexibility and realism to modeling commodity futures prices. To conclude, we go one step further by considering a three-factor model in which different prices for forwards and futures are found. This model includes the effects of a stochastic interest rate to the two-factor model.

To begin, we need to review the basic definitions of and distinctions between futures and forward contracts. A futures/forward contract is an agreement to buy or sell an asset at a certain time in the future for a certain price. An investor who agrees to buy has what is termed a long futures/forward position and an investor who agrees to sell has what is termed a short position. The price agreed to by the two parties is known as the futures/forward price. Now, forward contracts are negotiated directly between two parties, while futures trade on an organized exchange. As a result, forward contracts are not standardized where futures contracts are. Delivery of the underlying asset for a forward contract is specified at a unique date, and delivery or final cash settlement usually takes place. In contrast, futures have a range of delivery dates and they are usually closed out prior to delivery.

The role of the futures exchange is to organize trading so that contract defaults are minimized. In an effort to control default, futures exchanges require investors at the time they enter a contract to deposit funds into a margin account. At the end of the trading day, the account is adjusted, or *marked to market* to reflect the investor's gain or loss. If the account falls below some reserve level the exchange gives a margin call. When the call is received, the investor must deposit a sum of money making up the difference between the current balance and the initial margin. If the account appreciates above the initial margin the investor is allowed to withdraw the gain. These transactions occur at the close of each day, and they protect an investor from others defaulting on the contract. In opposition to the futures, forward contracts require no margin account.

3.1 One Factor Model

In a seminal paper, Black (1976) developed a pricing formula for a futures contract written on an spot commodity. This pricing formula is an extension of the Black and Scholes' (1973) option pricing model and Black's cost of carry solution is,

$$F(S(t),T) = S(t)e^{r(T-t)},$$

This simple pricing model is economically appealing for two reasons. One, it is arbitrage free, and two, the equation is obtained free from an investors' preferences. It is simply a powerful result. A more general and sometimes a simpler approach, however, can be used to obtain the same closed form solutions. This alternative method entails altering the true probability distribution for a particular stochastic process, and then computing a risk-neutral conditional expectation. This methodology is called equivalent martingale measures and is the method used in this study to price contingent claims.

There are useful implications in financial modeling from using equivalent martingale measures.[2] One, the risk neutral processes are purged of any risk premiums associated with their expected return. As a result, market participants may price an asset by simply discounting its forecasted future value by the *known* riskless rate of return. Second, these risk-neutral expectations are related to the arbitrage-free pricing method used by Black-Scholes. Therefore, any contingent claim that is priced under an equivalent martingale measure will itself be preference and arbitrage free.

The link between the risk-neutral expectation method of solution and the arbitrage-free method is established by the Feynman-Kac theorem.[3] Intuitively, this theorem shows that a correspondence between a certain class of conditional expectations and partial differential equations exists. A necessary condition for such a correspondence to exist is the underlying stochastic process needs to be Markovian. This condition is not too restrictive for in finance, its widely accepted by economists that asset prices follow a Markov process.

We are now ready to derive a pricing model using an equivalent martingale measure. Presume the commodity spot price follows a geometric Brownian motion[4]

$$dS(t) = \mu S(t)dt + \sigma S(t)dZ_s(t). \tag{1}$$

To obtain a solution for the futures contract price, a version of the Feynman-Kac theorem is invoked, and for tractability the market price of spot price risk is assumed to be constant.[5] The Feynman-Kac solution for the futures price is

$$F(S(T),T) = E_t^*(S(T)), \tag{2}$$

where $F(S(T),T)$ is the futures price for delivery at time T, and $S(T)$ is the commodity spot price at time T. The expectation is taken with respect to the risk-adjusted distribution for the spot price.

Stated earlier, equation (2) is the price to be paid for the commodity at time T. This price is simply the risk neutral forecast of today's spot price. Since the forecast of the spot price is preference free, we may obtain today's spot price by simply discounting the value by the riskfree rate of return. Since the risk-neutral spot price appreciates by the risk free rate, the discounted expected value will equal today's spot price. We can then say that the risk neutral process for the spot price is tingale then the futures price at time T will equal the today's spot price appreciated over the holding period by the risk free rate of return. This result is shown below by solving the expectation on the right-hand side of equation (2).

To start, let $X(t) = \ln(S(t))$. Since $X(t)$ is a twice-continuously differentiable function of $S(t)$ and continuously differentiable in t, then using Ito's lemma we can write the increment of $X(t)$ as

$$dX(t) = \frac{1}{S(t)} dS(t) - \frac{1}{[S(t)]^2} \frac{1}{2} [dS(t)]^2 + 0dt + 0dt^2. \qquad (3)$$

Now substitute in expression (3) the expressions for $dS(t)$ and $[dS(t)]^2$

$$dX(t) = \frac{1}{S(t)} (\mu S(t)dt + \sigma S(t)dZ_s(t)) - \frac{1}{[S(t)]^2} \frac{1}{2} \sigma^2 [S(t)]^2 dt,$$

$$dX(t) = \mu dt + \sigma dZ_s(t) - \frac{1}{2} \sigma^2 dt,$$

$$dX(t) = \left(\mu - \frac{1}{2}\sigma^2\right) dt + \sigma dZ_s(t). \qquad (4)$$

Equation (4) is the stochastic differential equation that depicts the salient time series characteristics of the logarithm of the spot price process. This movement is separated into two parts. There is a drift equal to $\left(\mu - \frac{1}{2}\sigma^2\right)dt$, and an innovation term given by $\sigma dZ_s(t)$.

Our objective is to find a solution for the spot price process that satisfies equation (2). By inspection of equation (1) the commodity is not risk free. Hence, the expected gross appreciation in the commodity's spot price is depicted as

$$\frac{E_t[S(t + dt)]}{S(t)} = (1 + R).$$

This gross return will on average be greater than the gross riskless rate of return. That is, on average

$$(1+R) \geq (1+r)$$

otherwise investment in the risky asset would not occur. The difference between R and r is λ. $\lambda = R - r$, represents the risk premium associated with the variability in the commodity's spot price. By definition λ is greater than zero and embedded in the mean return for $S(t)$. As a result, discounting the expected value of the terminal spot price by a risk-free discount factor will not yield a martingale.[6] We must somehow find a way to purge λ from $S(t)$ in order to price a futures contract according to expression (2). To achieve this objective we transform the probability distribution for $S(t)$ using the Girsanov theorem.[7] The Wiener process in (4) is transformed to

$$dZ_s(t) = dZ_s^*(t) - \lambda dt, \tag{5}$$

and we rewrite the stochastic differential equation for $X(t)$ above as a new risk-adjusted process:

$$dX(t) = \left(\mu - \frac{1}{2}\sigma^2\right)dt + \sigma\left(dZ_s^*(t) - \lambda dt\right). \tag{6}$$

Gibson and Schwartz (1990) show that the market price per unit of spot price risk λ can be solved. They argue the spot sale of a commodity is a contract for the immediate delivery of a commodity and in equilibrium this contract must satisfy the same partial differential equation as the contingent claim.[8] Their solution for the market price of spot price risk is

$$\lambda = \frac{\mu - r}{\sigma}. \tag{7}$$

Substituting equation (7) into equation (6) and rearranging yields

$$dX(t) = \left(\mu - \frac{1}{2}\sigma^2\right)dt + \sigma\left(dZ_s^*(t) - \left(\frac{\mu - r}{\sigma}\right)dt\right),$$

$$dX(t) = \left(r - \frac{1}{2}\sigma^2\right)dt + \sigma dZ_s^*(t). \tag{8}$$

Now integrate over equation (8) from t to T to obtain

$$\int_t^T dX(u) = \int_t^T \left(r - \frac{1}{2}\sigma^2\right)du + \sigma\int_t^T dZ_s^*(u),$$

$$X(T) - X(t) = \left(r - \frac{1}{2}\sigma^2\right)(T - t) + \sigma\int_t^T dZ_s^*(u),$$

$$X(T) = X(t) + \left(r - \frac{1}{2}\sigma^2\right)(T - t) + \sigma\int_t^T dZ_s^*(u). \tag{9}$$

Raising e to equivalent exponential powers, equation (9) yields

$$S(T) = S(t)e^{\left(r - \frac{1}{2}\sigma^2\right)(T-t) + \sigma\int_t^T dZ_s^*(u)}. \tag{10}$$

We now have a risk-adjusted expression for the terminal spot price. Recall from equation (2), the futures price is equal to the risk-adjusted expected spot price, $E_t^*(S(T))$. The question is how do we evaluate the expected value for the spot price? Consider the following. If

$$Y = \ln S,$$

$$e^Y = S,$$

$$e^{Y(t)} = S(t),$$

$$E_0\left[e^{Y(t)}\right] = E_0[S(t)] .$$

Furthermore, if $Y(t) \sim N(\mu t, \sigma^2 t)$ then

$$E_0\left[e^{Y(t)}\right]=\int_{-\infty}^{\infty} e^{y(t)}\frac{1}{\sqrt{2\pi\sigma^2 t}}e^{-\frac{1}{2}\frac{(y(t)-\mu t)^2}{\sigma^2 t}}dy(t),$$

$$E_0\left[e^{Y(t)}\right]=\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi\sigma^2 t}}e^{-\frac{1}{2}\frac{(y(t)-\mu t)^2}{\sigma^2 t}+y_t}dy(t) \quad . \tag{11}$$

The exponent is not a perfect square, but can be completed into one by multiplying the right-hand side of (11) by

$$e^{-\left(\mu t+\frac{1}{2}\sigma^2 t\right)}e^{\left(\mu t+\frac{1}{2}\sigma^2 t\right)}.$$

Substituting the above expression into (11) yields

$$E_0\left[e^{Y(t)}\right]=\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi\sigma^2 t}}e^{\left(\mu t+\frac{1}{2}\sigma^2 t\right)}e^{-\frac{1}{2}\frac{(y(t)-\mu t)^2}{\sigma^2 t}+y(t)-\left(\mu t+\frac{1}{2}\sigma^2 t\right)}dy(t) \quad .$$

Since $e^{\left(\mu t+\frac{1}{2}\sigma^2 t\right)}$ is independent of the variable of integration, it is a constant with respect to $Y(t)$, and can pass through the integral. Rewriting the equation above we obtain

$$E_0\left[e^{Y(t)}\right]=e^{\left(\mu t+\frac{1}{2}\sigma^2 t\right)}\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi\sigma^2 t}}e^{-\frac{1}{2}\frac{(y(t)-\mu t)^2}{\sigma^2 t}+y(t)-\left(\mu t+\frac{1}{2}\sigma^2 t\right)}dy(t) \quad . \tag{12}$$

Working on the exponent for the term inside the integral in equation (12), we can rewrite the exponent as

$$-\frac{1}{2}\frac{(y(t)-\mu t)^2}{\sigma^2 t}+y(t)-\left(\mu t+\frac{1}{2}\sigma^2 t\right),$$

$$-\frac{1}{2}\frac{(y(t)-\mu t)^2}{\sigma^2 t}+\frac{y(t)\sigma^2 t-\left(\mu t+\frac{1}{2}\sigma^2 t\right)\sigma^2 t}{\sigma^2 t},$$

$$\frac{-\frac{1}{2}(y(t)-\mu t)^2}{\sigma^2 t}+\frac{-\frac{1}{2}\left(-2y(t)\sigma^2 t+2\mu t\sigma^2 t-\sigma^4 t\right)}{\sigma^2 t},$$

$$\frac{-\frac{1}{2}\left((y(t)-\mu t)^2-2y(t)\sigma^2 t+2\mu t\sigma^2 t-\sigma^4 t\right)}{\sigma^2 t},$$

$$\frac{-\frac{1}{2}\left((y(t))^2-2y(t)\mu t+(\mu t)^2-2y(t)\sigma^2 t+2\mu t\sigma^2 t-\sigma^4 t\right)}{\sigma^2 t},$$

$$\frac{-\frac{1}{2}\left((y(t))^2-2y(t)\left(\mu t+\sigma^2 t\right)+(\mu t)^2+2\mu t\sigma^2 t-\sigma^4 t\right)}{\sigma^2 t},$$

$$\frac{-\frac{1}{2}\left((y(t))^2-2y(t)\left(\mu t+\sigma^2 t\right)+\left(\mu t+\sigma^2 t\right)^2\right)}{\sigma^2 t},$$

$$\frac{\left(-\frac{1}{2}\right)\left[y(t)-(\mu t+\sigma^2 t)\right]^2}{\sigma^2 t}.$$

Thus, we may rewrite equation (12) as

$$E_0\left[e^{Y(t)}\right]=e^{\left(\mu t+\frac{1}{2}\sigma^2 t\right)}\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi\sigma^2 t}}e^{-\frac{1}{2}\frac{\left(y(t)-(\mu t+\sigma^2 t)\right)^2}{\sigma^2 t}}dy(t) \quad . \tag{13}$$

Since

$$\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi\sigma^2 t}}e^{-\frac{1}{2}\frac{\left(y(t)-(\mu t+\sigma^2 t)\right)^2}{\sigma^2 t}}dy(t)$$

integrates to one for a normal random variable with mean $\mu t+\sigma t^2$ and variance $\sigma^2 t$. The expected value of $Y(t)$ may be expressed as

$$E_0\left[e^{Y(t)}\right]=e^{\left(\mu t+\frac{1}{2}\sigma^2 t\right)} \quad . \tag{14}$$

For my model, we are interested in $E_t^*\left[e^{X(T)}\right]$ where $X(t)=\ln S(t)$. Therefore, we need

$$X(t)=\ln S(t)$$

$$e^{X(t)} = S(t)$$

$$e^{X(T)} = S(T)$$

$$E_t^*\left[e^{X(T)}\right] = E_t^*[S(T)].$$

Since $X(T)$ follows a Gaussian distribution we may write the expected value as

$$E_t^*\left[e^{X(T)}\right] = E_t^*[S(T)] = e^{E_t^*(X(T)) + \frac{1}{2}V_t^*(X(T))} \quad . \tag{15}$$

The first moment for $X(T)$ is

$$E_t^*[X(T)] = E_t^*\left[X(t) + \left(r - \frac{1}{2}\sigma^2\right)(T-t) + \sigma\int_t^T dZ_s^*(u)\right],$$

$$E_t^*[X(T)] = X(t) + \left(r - \frac{1}{2}\sigma^2\right)(T-t) \; . \tag{16}$$

The second moment of $X(T)$ is

$$V_t^*[X(T)] = E_t^*\left[X(t) - E_t^*(X(t))\right]^2,$$

$$V_t^*[X(T)] = E_t^*\left[X(t) + \left(r - \frac{1}{2}\sigma^2\right)(T-t) + \sigma\int_t^T dZ_s^*(u) - \left(X(t) + \left(r - \frac{1}{2}\sigma^2\right)(T-t)\right)\right]^2,$$

$$V_t^*[X(T)] = E_t^*\left[\sigma\int_t^T dZ_s^*(u)\right]^2 \; . \tag{17}$$

Before we go further, we must be capable of evaluating the integral in expression (17). This is not a straight forward matter and can be quite difficult. The reason is the integral in equation (17) is stochastic, meaning it is not a smooth function. Since the function we are considering is not smooth, no limit exists. Consequently, we may not use the standard Riemann integral to evaluate expression (17). We need to find another method for evaluating equation (17).

To evaluate expression (17) we invoke the use of Ito's isometry.[9] Ito's isometry states

$$E\left[\int_a^b f(t)dW(t)\int_a^b g(t)dW(t)\right]=\sigma_w^2\int_a^b g(t)f(t)dt \tag{18}$$

where $dW(t) \sim N(0,\sigma\sqrt{t})$. The right hand side of expression (18) is a smooth function of time. Ito's isometry shows that the square of a stochastic integral reduces to a well know Riemann sum. Therefore, with the result in (18), the variance for a stochastic process is stated as

$$E\left[\int_a^b f(t)dW(t)\int_a^b f(t)dW(t)\right]=\sigma_w^2\int_a^b (f(t))^2 dt \tag{19}$$

$dW(t) \sim N(0,\sigma_w\sqrt{t})$. This result is for $dW(t) \sim N(0,\sigma_w\sqrt{t})$, however, we have $dZ_s^*(t)$; where $dZ_s^*(t) \sim N(0,\sqrt{t})$. Substituting the results from equation (19) into equation (17) yields

$$V_t^*[X(T)]=\sigma^2(1)\int_t^T dt,$$

$$V_t^*[X(T)]=\sigma^2(T-t). \tag{20}$$

Substitute equations (14) and (20) into equation (15) yields

$$E_t^*\left[e^{X(T)}\right]=e^{\left(X(t)+\left(r-\frac{1}{2}\sigma^2\right)(T-t)+\frac{1}{2}\sigma^2(T-t)\right)},$$

$$E_t^*\left[e^{X(T)}\right]=S(t)e^{r(T-t)}. \tag{21}$$

Equation (21) states that the risk neutral expectation of the spot price at time T is simply the future value of today's spot price appreciated by the known riskless rate of return. Just as we discussed above, this makes sense. The expectation in expression (21) is taken with respect to risk-adjusted probability distribution. The spot price is purged of any risk premium associated with it, and thereby, the drift will equal the risk free rate. We see that if we discount the risk-adjusted expected value for the spot price we get a martingale. The best forecast we have for tomorrows price will be today's spot price. The futures price is then

$$F(S(T),T) = S(t)e^{r(T-t)}. \tag{22}$$

Equation (22) is Black's well known cost of carry solution.

Given Black's cost of carry solution, we see that the futures price for a commodity contract is simply the expected value of a risk-adjusted spot price. While solution is arbitrage free and a function of the expected spot price, it may not be a reasonable depiction of a commodity's futures price. The underlying spot price is assumed to follow a geometric Brownian motion. Intuitively, this time series behavior does not seem consistent with commodity spot prices. Consider figure 1. This figure shows the time series behavior of both the nearby and eighteen month futures price contract for crude oil over the last six years. Allowing the nearby contract to proxy for the spot price, and the eighteen-month contract to proxy the expected spot price, we can see some systematic behavior in these time series. One, the spot price is more volatile than expected spot prices, and two, the spot price seems to oscillate around it's long run mean. Does this make economic sense?

Schwartz (1997) argues, in an equilibrium setting, we would expect when prices are relatively high, supply will increase since higher cost producers of the commodity will enter the market putting downward pressure on prices.[10] Conversely, when prices are relatively low, supply will decrease since some of the higher cost producers will exit the market, putting upward pressures on prices. The impact of relative prices on the supply of the commodity will induce mean reversion in commodity prices. The mean reverting nature of commodity prices has been considered in a series of recent articles (Gibson and Schwartz (1990) Cortazar and Schwartz (1994), Schwartz (1997), and Miltersen and Schwartz (1998), Brennan (1991) and Bessembinder et al (1995)). The evidence suggests that an alternative spot price process should be considered when pricing commodity futures. Consider the following model.

Presume the commodity spot price follows a mean reverting stochastic process

$$\frac{dS}{S} = k(\mu - \ln S)dt + \sigma dZ_s \tag{23}$$

where S is the spot price, μ is the long run expected return for the spot price, k is the speed of adjustment factor measuring the degree of mean reversion to the long run mean return, σ is the diffusion coefficient, and dZ_s is the increment of a standard Brownian motion.[11] Equation (23) illustrates the tendencies of commodity prices to revert to their long run mean.[12] That is, if the logarithm of the spot price is higher than the long run return the bracketed term in equation (23) is negative putting downward pressure on the incremental movement in price. Hence, the spot price reverts back toward its average and the rate of this adjustment is determined by k. The solution to a commodity's futures price given the stochastic differential equation in equation (23) is again given by Feynman-Kac in expression (2). Therefore, we need to solve for the futures price under an equivalent martingale measure.

To find a solution to the above stochastic differential equation, we start by rewriting the spot price in terms of its logarithm. We let $X = \ln S$, where X is a twice differentiable function of S and continuously differentiable with respect to time. Using Ito's Lemma, and the transformation the differential for the logarithm of the spot price is

$$dX = X_s dS + \frac{1}{2} X_{ss} dS^2 + X_t dt + X_{tt} dt^2 \ , \qquad (24)$$

where X_S and X_{SS} are the partial derivatives with respect to spot price and X_t, X_{tt} are the partial derivatives with respect to time. Substituting for X_S, X_{SS}, X_t, dS and $(dS)^2$ into equation (40) we obtain

$$dX = \left(\frac{1}{S}\right)\left[k(\mu - \ln S)Sdt + \sigma S dZ_s\right] - \frac{1}{2}\left(\frac{1}{S^2}\right)\left[\sigma^2 S^2 dt\right] \ . \qquad (25)$$

Rearranging the terms in equation (25) yields

$$dX = k(\mu - X)dt + \sigma dZ_s - \frac{1}{2}\sigma^2 dt,$$

$$dX = \left[k(\mu - X) - \frac{1}{2}\sigma^2\right]dt + \sigma dZ_s,$$

$$dX = \left[k(\alpha - X)\right]dt + \sigma dZ_s, \qquad (26)$$

where $\alpha = \mu - \frac{\sigma^2}{2k}$. Equation (26) follows a mean reverting process of the Ornstein-Uhlenbeck type. Note, the difference between equation (23) and equation (26). Equation (23) is a nonlinear function with respect to the state variable (the spot price), and equation (26) is a linear. The advantage of rewriting the spot price process is an immediate textbook solution exists for equation (26) (this result is illustrated below). Once we have the solution for (26) then we have the solution for (23). This solution leads us to

$$F(S(T),T) = E_t^*(S(T)). \qquad (27)$$

Again, $F(S(T),T)$ is the current price of a futures contract expiring at time T, and $E_t^*[]$ is the expectation operator under an equivalent probability measure.

Under standard arbitrage assumptions, the dynamics of the Ornstein-Uhlenbeck process under the equivalent martingale measure can be written as

$$dX = [k(\alpha^* - X)]dt + \sigma dZ_s^* \tag{28}$$

where $\alpha^* = \alpha - \lambda$, λ is the market price of risk (assumed constant) and dZ_s^* is the increment of a standard Brownian motion under the equivalent martingale measure. Indexing equation (28) in consideration to time, expression (28) is rewritten as

$$dX(t) = [k(\alpha^* - X(t))]dt + \sigma dZ_s^*(t). \tag{29}$$

Rearranging equation (29)

$$dX(t) + kX(t)dt = k\alpha^* dt + \sigma dZ_s^*(t),$$

and multiplying both sides of the expression above by e^{kt} yields

$$e^{kt}(dX(t) + kX(t)dt) = e^{kt}(k\alpha^* dt + \sigma dZ_s^*(t)) . \tag{30}$$

The left-hand side of equation (30) is the algebraic expression for the total differential of $e^{kt}X(t)$. For example, taking the total differential of $e^{kt}X(t)$ yields

$$d[e^{kt}X(t)] = \frac{\partial[e^{kt}X(t)]}{\partial t}dt + \frac{\partial[e^{kt}X(t)]}{\partial X(t)}dX(t),$$

$$d[e^{kt}X(t)] = ke^{kt}X(t)dt + e^{kt}dX(t) .$$

Replacing the left-hand side of equation (30) with the left-hand side of the expression above yields

$$d[e^{kt}X(t)] = e^{kt}(k\alpha^* dt + \sigma dZ_s^*(t)) . \tag{31}$$

Integrating over equation (31)

$$\int_t^T d\left[e^{ku}X(u)\right] = \int_t^T e^{ku}k\alpha^* du + \sigma\int_t^T e^{ku} dZ_s^*(u),$$

which yields

$$e^{ku}X(u)\Big|_t^T = \alpha^*\int_t^T e^{ku}kdu + \sigma\int_t^T e^{ku}dZ_s^*(u) \quad .$$

Evaluating the left-hand side and integrating the drift term of the above yields

$$e^{kT}X(T) - e^{kt}X(t) = \alpha^* e^{ku}\Big|_t^T + \sigma\int_t^T e^{ku}dZ_s^*(u) \ .$$

Normalizing on $X(T)$ yields

$$X(T) = X(t)e^{kt-kT} + \alpha^* e^{kT-kT} - \alpha^* e^{kt-kT} + \sigma e^{-kT}\int_t^T e^{ku}dZ_s^*(u),$$

or

$$X(T) = \theta X(t) + \alpha^*(1-\theta) + \sigma e^{-kT}\int_t^T e^{ku}dZ_s^*(u) \qquad (32)$$

where $\theta = e^{-k(T-t)}$. The expression for $X(T)$ is the solution for the stochastic differential equation in equation (28). A solution for the risk-adjusted spot price process is available by taking the exponential of the left-hand side and right-hand side of equation (32)

$$S(T) = \exp\left(\theta X(t) + \alpha^*(1-\theta) + \sigma e^{-kT}\int_t^T e^{ku}dZ_s^*(t)\right). \qquad (33)$$

From equation (27), we know the futures price is a function of the risk-adjusted expected spot price. With the result presented in equation (15), we need to find the first and second moments for equation (32).

Taking the risk neutral conditional expected value of equation (32) at time $t = 0$ yields

$$E_0^*[X(T)] = E_0^*\left[e^{-kT}X(t) + \alpha^*(1-e^{-kT}) + \sigma e^{-kT}\int_0^T e^{ku}dZ_s^*(u)\right] \tag{34}$$

The expected value of $dZ_s^*(u)$ is zero, thereby making $E_t^*\left[\sigma e^{-kT}\int_t^T e^{ku}dZ_s^*(s)\right] = 0$.

$$E_0^*[X(T)] = \alpha^* + (X(0)-\alpha^*)e^{-kT}. \tag{35}$$

The conditional variance for logarithm of spot price evaluated at $t = 0$ under the adjusted probability measure is

$$V_0^*[X(T)] \equiv E_0^*[X(T) - E_0^*(X(T))]^2. \tag{36}$$

Substituting in for $X(T)$, $E_0^*(X(T))$ and simplifying terms, the risk neutral variance is

$$V_0^*[X(T)] = E_0^*\left[\sigma e^{-kT}\int_0^T e^{ku}dZ_s^*(u)\right]^2$$

$$V_0^*[X(T)] = \sigma^2 e^{-2kT}E_0^*\left[\int_0^T e^{ku}dZ_s^*(u)\right]^2. \tag{37}$$

In order to evaluate equation (37), we must integrate the right-hand side. The integral in equation (37) however, is stochastic and itself a random variable. That is, *dZ* is nowhere differentiable and the integral techniques in deterministic calculus cannot be applied. The integral in equation (37), however, is a square integrable second order stochastic process and is defined in the sense of Ito. Thus, we may use Ito's isometry, to evaluate equation (37).

$$V_0^*[X(T)] = \sigma^2 e^{-2kT}(1)\int_0^T (e^{ku})^2 du,$$

$$V_0^*[X(T)] = \sigma^2 e^{-2kT}\left[\frac{1}{2k}e^{2ku}\Big|_0^T\right],$$

$$V_0^*[X(T)] = \sigma^2 e^{-2kT}\frac{1}{2k}\left[e^{2kT}-1\right],$$

$$V_0^*[X(T)] = \frac{\sigma^2 e^{-2kT}}{2k}\left[e^{2kT}-1\right],$$

$$V_0^*[X(T)] = \frac{\sigma^2}{2k}\left[1-e^{-2kT}\right] \quad . \tag{38}$$

Now substitute equations (35) and (38) into (15)

$$E_0^*[S(T)] = \exp\left[X(0)e^{-kT} + \alpha^*\left(1-e^{-kT}\right) + \frac{\sigma^2}{4k}\left[1-e^{-2kT}\right]\right] \quad . \tag{39}$$

Substitution of equation (39) into equation (27) gives an analytical solution for the futures price. The futures price is

$$F[S(T),T] = \exp\left[e^{-kT}\ln S(0) + \alpha^*\left(1-e^{-kT}\right) + \frac{\sigma^2}{4k}\left[1-e^{-2kT}\right]\right] \quad . \tag{40}$$

Expression in (40) is the solution to the partial differential equation shown in appendix B. This result is not a surprise. As discussed earlier, the Feynman-Kac solution for the risk-neutral futures price is an implicit solution to a corresponding partial differential equation. This result is explicitly shown through using the differential form of Ito's Lemma and applying the Girsanov transformation to the driving Wiener process. Illustrating,

$$dF[S(T),T] = \frac{\partial F[S(T),T]}{\partial t}dt + \frac{\partial F[S(T),t,T]}{\partial S(t)}dS(t) + \frac{1}{2}\cdot\frac{\partial F[S(T),T]}{\partial S(t)\partial S(t)}[dS(t)]^2$$

$$= F_t dt + F_s dS(t) + \frac{1}{2}F_{ss}[dS(t)]^2 \quad . \tag{41}$$

Substituting expressions for $dS(t)$ and $[dS(t)]^2$ into equation (41) yields

$$= F_t dt + F_s\left(k(\mu - \ln S(t))S(t)dt + \sigma S(t)dZ_s(t)\right) + \frac{1}{2}F_{ss}\sigma^2[S(t)]^2 dt. \tag{42}$$

Rearranging terms in (42) yields

$$= \left(\frac{1}{2}F_{ss}\sigma^2[S(t)]^2 + F_s k(\mu - \ln S(t))S(t) + F_t\right)dt + \sigma S(t)F_s dZ_s(t)\ . \tag{43}$$

Expression (43) is the stochastic differential for the futures price under the true probability distribution for the futures price. We need a risk-adjusted process. Applying the Girsanov theorem, the Brownian motion term, $dZ_s(t)$, becomes $dZ_s^*(t) - \lambda dt$. The transformed futures process is

$$= \left(\frac{1}{2}F_{ss}\sigma^2[S(t)]^2 + F_s k(\mu - \ln S(t))S(t) + F_t\right)dt + \sigma S(t)F_s\left(dZ_s^*(t) - \lambda dt\right). \tag{44}$$

Rearranging,

$$dF[S(T),T] = \left(\frac{1}{2}F_{ss}\sigma^2[S(t)]^2 + F_s k\left(\mu - \frac{\lambda\sigma}{k} - \ln S(t)\right)S(t) + F_t\right)dt + \sigma S(t)F_s dZ_s^*(t)$$

$$dF[S(T),T] = \left(\frac{1}{2}F_{ss}\sigma^2[S(t)]^2 + F_s k(\mu - \lambda'\sigma - \ln S(t))S(t) + F_t\right)dt + \sigma S(t)F_s dZ_s^*(t).$$

(45)

Note in the expression above that $\lambda' = \frac{\lambda}{k}$. Under the risk neutral probability measure, equation (45) is a martingale. That is, the expected value of (45) is zero. This means

$$E_t^*[dF] = E_t^*\left[\left(\frac{1}{2}F_{ss}\sigma^2[S(t)]^2 + F_s k(\mu - \lambda'\sigma - \ln S(t))S(t) - F_T\right)dt + \sigma S(t)F_s dZ_s^*(t)\right] = 0$$

$$= E_t^*\left[\left(\frac{1}{2}F_{ss}\sigma^2[S(t)]^2 + F_s k(\mu - \lambda'\sigma - \ln S(t))S(t) - F_T\right)dt\right] + E_t^*\left[\sigma S(t)F_s dZ_s^*(t)\right] = 0,$$

$$=\left(\frac{1}{2}F_{SS}\sigma^2[S(t)]^2+F_Sk(\mu-\lambda'\sigma-\ln S(t))S(t)-F_T\right)dt+\sigma S(t)F_S E_t^*[dZ_s^*(t)]=0,$$

$$=\left(\frac{1}{2}F_{ss}\sigma^2[S(t)]^2+F_sk(\mu-\lambda'\sigma-\ln S(t))S(t)-F_T\right)dt=0$$

From the expression above, we see the only way for the risk adjusted futures contract to be a martingale is for the drift term to equal zero. That is

$$\frac{1}{2}F_{ss}\sigma^2[S(t)]^2+F_sk(\mu-\lambda'\sigma-\ln S(t))S(t)-F_T=0. \tag{46}$$

This expression is identical to the partial differential equation above and illustrates that the Feynman-Kac solution is indeed the probabilistic solution to the partial differential equation obtain from a standard arbitrage free pricing model. The futures price obeys the stochastic differential equation

$$dF[S(T),t,T]=\sigma S(t)F_s dZ_s^*(t)\ . \tag{47}$$

3.2 Two Factor Model

In section 3.1 we derived the solution for commodity futures prices where the only source of uncertainty in the futures price was the spot price. Theoretical solutions were derived for two cases. The first case presumed the spot price process followed a geometric Brownian motion, the second presumes an Ornstein-Uhlenbeck mean reverting process. If the spot price is the only determinant for futures prices these models should work reasonably well.

The question under consideration now is whether the one factor model is a reasonable model for pricing commodity futures contracts. For a storable commodity, the one factor model may not capture all the information compounded in the expected spot price. That is, holding inventories of the physical commodity has benefits, and when the benefit is expressed in percentage terms it is called a convenience yield. The convenience yield as defined by Brennan and Schwartz (1985) is the flow of services net of storage costs that accrues to an owner of the physical commodity, but not to the owner of a contract for future delivery. When a market participant takes ownership of the physical commodity, the owner chooses where it is stored and when it is liquidated. Naturally, the owners of the physical commodity must feel there are benefits from ownership that are not obtained by holding a futures contract. In lieu of the storage costs, the benefits may include the ability to profit from temporary local shortages or the ability to keep production process running in the event of a supply disruption. Thus, if there is a net benefit to storing a commodity, any model that prices a futures contract must consider this impact.

Earlier studies attempting to price contingent claims written on storable commodities (Schwartz (1982) and Brennan and Schwartz (1985)) adjust Black's cost of carry model to include a convenience yield. For tractability, these models presume the convenience yield, δ, is a constant proportion of the commodity's price. Following Black's derivation, these studies determine Black's augmented cost of carry model as

$$F(S(t),T) = S(t)e^{(r-\delta)(T-t)} \quad . \tag{48}$$

The futures price for Black's augmented model is simply the original futures price discounted by the convenience yield. The owner of the futures contract does not receive the benefit from owing the physical before time T. Therefore, the terminal spot price is lowered by the percentage lost from holding the contract as opposed to holding the physical commodity.

The method of solution used to determine equation (48) is the traditional Black-Scholes method. Alternatively, we may derive this expression using an equivalent martingale measure. To apply this method, we need to alter the stochastic differential equation for the spot price process underlying the futures price. For our problem, the spot price process is the same as expression (1), and is

$$dS(t) = \mu S(t)dt + \sigma S(t)dZ_s(t).$$

To alter the distribution for the spot price, we must alter the distribution for the innovation term. Thus, transforming the spot price via the Girsanov theorem, the Brownian motion term becomes

$$dZ_s(t) = dZ_s^*(t) - \lambda dt.$$

Furthermore, Gibson and Schwartz (1990) (see Appendix B) show that the market price of spot price risk, λ, for this model is

$$\lambda = \frac{\mu + \delta - r}{\sigma}. \tag{49}$$

Substituting expression (49) into the transformation of the Brownian motion above yields

$$dZ_s(t) = dZ_s^*(t) - \left(\frac{\mu + \delta - r}{\sigma}\right)dt.$$

The risk-adjusted spot price process becomes

$$\frac{dS(t)}{S(t)} = (r - \delta)dt + \sigma dZ_s^*(t) \quad . \tag{50}$$

Recall the Feynman-Kac solution for the futures price. This is

$$F(S(T),T) = E_t^*(S(T)).$$

Following the argument for the derivation in section 3.1, the solution for the current system is

$$F(S(t),T) = S(t)e^{(r-\delta)(T-t)} \quad . \tag{51}$$

Expression (51) is the same as the futures price stated in equation (48). Again, the solution discounts the value of the futures contract that matures at time T by δ. Intuitively, the owner of a long futures contract is compensated for the losing the benefits of physical ownership. Note, if the convenience yield was negative then the investor with the short position will be compensated for the cost of storing the physical commodity.

The result above naturally begs the question, is it reasonable to assume that the convenience yield for a commodity will remain constant? There is evidence in the literature (Fama and French (1987), Brennan (1991), Gibson and Schwartz (1990) and Schwartz (1997)) that supports the need for a stochastic convenience yield when pricing futures contracts. In light of the empirical research, we now consider a two-factor model that includes a mean reverting stochastic convenience yield.

Consider the two-factor model presented by Gibson and Schwartz (1990). They presume the spot price and convenience yield follow the joint stochastic process below[13]

$$dS(t) = \mu S(t)dt + \sigma_s S(t)dZ_s(t) \tag{52}$$

$$d\delta(t) = k(\alpha - \delta(t))dt + \sigma_c dZ_c(t) \tag{53}$$

where μ is the instantaneous expected return in the spot price, and σ_s is the diffusion coefficient. k is the speed of adjustment parameter for the convenience yield, $\delta(t)$ around its instantaneous long run mean, α. $dZ_s(t)$ and $dZ_c(t)$ are correlated increments to standard Brownian motions. $Cov_t[dZ_s(t)dZ_c(t)] = \rho_{sc}dt$, where ρ_{sc} is the correlation coefficient between the two Brownian motions.

The spot price process follows a geometric Brownian motion, indicating the drift and diffusion coefficient change proportional to the level of the spot price. The convenience yield is presumed to follow a mean reverting process of the Ornstein-Uhlenbeck type. This specification allows for both positive and negative yields. This is an appealing characteristic, in that the owner of the physical is compensated when net storage costs are positive, and penalized when they are negative.

The advantage of the two-factor model is it allows greater flexibility in modeling futures prices. This increased flexibility, however, comes at a cost. For instance, recall the derivation for the one-factor model. There we calculate the risk adjusted expected value of the expected spot rate to find an expression for the commodity's futures price. Presently, we want again want to solve the risk neutral expected spot price, but this must now consider the distribution for both the spot price and the convenience yield. This adds greater complexity to the model. Let us now work on the two-factor model.

Invoking the Feynman-Kac theorem, we know the futures price for the two-factor model is given as

$$F(S(T),\delta(T),T) = E_t^*(S_T), \tag{54}$$

where T is maturity. $S(T)$ is the commodity spot price at maturity. $\delta(T)$ is the convenience yield, $E_t^*[\ ]$ is the expectation operator under a transformed probability measure. In order to calculate the expectation in expression (54), we need solutions for the commodity spot price, $S(T)$, the convenience yield, $\delta(T)$, and the cumulative yield, $X(T)$ as well.

Working with the spot price, let $Y(t) = \ln S(t)$. Since $Y(t)$ is a twice-continuously differentiable function of $S(t)$ and time, we may write the stochastic differential for $Y(t)$ as

$$dY(t) = Y_s dS(t) - \frac{1}{2} Y_{ss} [dS(t)]^2 .$$

Substituting in for Y_s, Y_{ss}, $dS(t)$ and $[dS(t)]^2$ I have

$$dY(t) = \frac{1}{S(t)} dS(t) - \frac{1}{2[S(t)]^2} [dS(t)]^2 ,$$

$$dY(t) = \frac{1}{S(t)} [\mu S(t) dt + \sigma_s S(t) dZ_s(t)] - \frac{1}{2[S(t)]^2} \left[\sigma_s^2 [S(t)]^2 dt\right],$$

$$= \mu dt + \sigma_s dZ_s(t) - \frac{1}{2}\sigma_s^2 dt ,$$

$$= \left[\mu - \frac{1}{2}\sigma_s^2\right] dt + \sigma_s dZ_s(t) \quad . \tag{55}$$

Now, integrating both sides of equation (55)

$$\int_t^T dY(u) = \int_t^T \left[\mu - \frac{1}{2}\sigma_s^2\right] du + \int_t^T \sigma_s dZ_s(u) ,$$

$$Y(T) - Y(t) = \int_t^T \left[\mu - \frac{1}{2}\sigma_s^2\right] du + \int_t^T \sigma_s dZ_s(u) ,$$

$$Y(T) = Y(t) + \int_t^T \left[\mu - \frac{1}{2}\sigma_s^2\right] du + \int_t^T \sigma_s dZ_s(u) \quad . \tag{56}$$

Raising e to equivalent powers, expression (56) becomes

$$S(T) = S(t) e^{\int_t^T \left[\mu - \frac{1}{2}\sigma_s^2\right] du + \int_t^T \sigma_s dZ_s(u)} . \tag{57}$$

Similar to the one factor model, it seems permissible to evaluate equation (54). This is incorrect. The reason is the convenience yield affects spot price levels due to changes in inventory levels of the commodity. That is, as inventories drop so too does the availability of the commodity. In this case, the commodity is becoming increasingly scarce. Consequently, the reduction in inventory causes an increase in the spot price. The percentage increase in the spot price due to limited inventories is called the convenience yield. Since this yield in subsumed in the spot price, it is then linked to the commodity's expected return. Therefore, when we evaluate the expected value of $S(T)$, we must consider the impact that the convenience yield has on the expected spot price. This leads us to finding solutions for the convenience yield $\delta(T)$ and the cumulative yield, $X(T)$.

Starting with equation (53) and rearranging yields

$$d\delta(t) = k\alpha dt - k\delta(t)dt + \sigma_c dZ_c(t),$$

$$d\delta(t) + k\delta(t)dt = k\alpha dt + \sigma_c dZ_c(t) \quad . \qquad (58)$$

Multiply both sides of equation (58) by e^{kt}

$$e^{kt}[d\delta(t) + k\delta(t)dt] = e^{kt}[k\alpha dt + \sigma_c dZ_c(t)]. \qquad (59)$$

We can rewrite the left-hand side of equation (59). Consider the function $e^{kt}\delta(t)$. The total differential is

$$d\left(e^{kt}\delta(t)\right) = \frac{\partial e^{kt}\delta(t)}{\partial t}dt + \frac{\partial e^{kt}\delta(t)}{\partial \delta(t)}d\delta(t)$$

$$= ke^{kt}\delta(t)dt + e^{kt}d\delta(t) \quad . \qquad (60)$$

The reader should note that the LHS of expression (59) is equal to the RHS of equation (60). Therefore, we may replace the LHS of equation (59) with the LHS of equation (60). This yields

$$d\left(e^{kt}\delta(t)\right) = e^{kt}k\alpha dt + e^{kt}\sigma_c dZ_c(t) \quad . \tag{61}$$

Integrating over equation (61)

$$\int_t^T d\left(e^{ku}\delta(u)\right) = k\alpha\int_t^T e^{ku}du + \sigma_c\int_t^T e^{ku}dZ_c(u),$$

$$e^{ku}\delta(u)\Big|_t^T = \alpha e^{ku}\Big|_t^T + \sigma_c\int_t^T e^{ku}dZ_c(u),$$

$$e^{kT}\delta(T) - e^{kt}\delta(t) = \alpha e^{kT} - \alpha e^{kt} + \sigma_c\int_t^T e^{ku}dZ_c(u),$$

$$e^{kT}\delta(T) = e^{kt}\delta(t) + \alpha e^{kT} - \alpha e^{kt} + \sigma_c\int_t^T e^{ku}dZ_c(u),$$

$$\delta(T) = e^{kt-kT}\delta(t) + \alpha e^{kT-kT} - \alpha e^{kt-kT} + \sigma_c e^{-kT}\int_t^T e^{ku}dZ_c(u),$$

$$\delta(T) = e^{-k(T-t)}\delta(t) + \alpha - \alpha e^{-k(T-t)} + \sigma_c e^{-kT}\int_t^T e^{ku}dZ_c(u),$$

$$\delta(T) = \theta\delta(t) + (1-\theta)\alpha + \sigma_c e^{-kT}\int_t^T e^{ku}dZ_c(u) \ , \tag{62}$$

where $\theta = e^{-k(T-t)}$.

We now have solutions for $S(T)$ and $\delta(T)$. Anticipating the evaluation of the expectation in expression (54), we need to find the first and second moments for the terminal spot price and convenience yield. From equation (20), I know the incremental spot price in equation (57) is lognormal with mean $\left(\mu - \frac{1}{2}\sigma_s\right)dt$ and variance $\sigma_s^2 dt$. In addition, from Levy's theorem we know the convenience yield is normally distributed with mean $\hat{\mu}$ and variance $\hat{\sigma}^2$. We obtain values for $\hat{\mu}$ and $\hat{\sigma}^2$ by taking the conditional expectation and variance of equation (62). The conditional expectation is

$$\hat{\mu} \equiv E_t(\delta(T)),$$

$$E_t[\delta(T)] = \theta\delta(t) + (1-\theta)\alpha + \sigma_c e^{-kT}\int_t^T e^{ku} E_t[dZ_c(u)],$$

$$E_t[\delta(T)] = \theta\delta(t) + (1-\theta)\alpha \quad . \tag{63}$$

The variance is defined as

$$V_t[\delta(T)] \equiv E_t[\delta(T) - E_t(\delta(T))]^2,$$

$$V_t[\delta(T)] = E_t\left[\sigma_c e^{-kT}\int_t^T e^{ku} dZ_c(u)\right]^2,$$

$$= \sigma_c^2 e^{-2kT} E_t\left[\int_t^T e^{ku} dZ_c(u)\right]^2.$$

Using Ito's isometry

$$= \sigma_c^2 e^{-2kT}\left(\sigma_z^2 \int_t^T e^{2ku} du\right),$$

$$= \sigma_c^2 e^{-2kT}\left((1)\int_t^T e^{2ku} du\right),$$

$$= \sigma_c^2 e^{-2kT}\left(\frac{1}{2k} e^{2ku}\Big|_t^T\right),$$

$$= \sigma_c^2 e^{-2kT}\left(\frac{1}{2k}(e^{2kT} - e^{2kt})\right),$$

$$= \frac{\sigma_c^2}{2k}\left(e^{2kT-2kT} - e^{2kt-2kT}\right),$$

$$= \frac{\sigma_c^2}{2k}\left(1 - e^{-2k(T-t)}\right),$$

$$= \frac{\sigma_c^2}{2k}\left(1 - \theta^2\right) \quad . \tag{64}$$

To review, we started with the stochastic differential equations for the spot price and convenience yield, and then developed solutions for $S(T)$ and $\delta(T)$. Furthermore, we have characterized the distributions for $S(T)$ and $\delta(T)$. Continuing, our next step is to find an expression for the cumulative convenience yield $X(T)$. The reason for finding a cumulative yields is that the solution for $S(T)$ is found over a holding period $T-t$. We are not just interested in the convenience yield at a particular time. Instead, we are interested in knowing the accumulated convenience yield for the entire holding period.

We define the cumulative convenience yield as

$$X(t) \equiv \int_0^t \delta(u)du, \tag{65}$$

where $X(0) = 0$. Recall equation (53)

$$d\delta(t) = k(\alpha - \delta(t))dt + \sigma_c dZ_c(t).$$

Equation (53) implies

$$\int_t^T d\delta(u) = \int_t^T k(\alpha - \delta(u))du + \int_t^T \sigma_c dZ_c(u),$$

$$\int_t^T d\delta(u) = k\alpha(T-t) - \int_t^T \delta(u)du + \int_t^T \sigma_c dZ_c(u). \tag{66}$$

It follows that

$$\delta(T)-\delta(t)=\int_t^T d\delta(u)\ , \tag{67}$$

$$X(T)-X(t)=\int_t^T \delta(u)du\,. \tag{68}$$

Inserting equations (67) and (68) into (66) we obtain

$$\delta(T)-\delta(t)=k\alpha(T-t)-k\big(X(T)-X(t)\big)+\int_t^T \sigma_c dZ_c(u)\,. \tag{69}$$

We may further substitute the expression for $\delta(T)$, equation (62), into equation (69) to obtain

$$\theta\delta(t)+(1-\theta)\alpha+\sigma_c e^{-kT}\int_t^T e^{ku}dZ_c(u)-\delta(t)=k\alpha(T-t)-k\big(X(T)-X(t)\big)+\int_t^T \sigma_c dZ_c(u)\,.$$

Our objective is to find a solution for $X(T)$. Thus, we normalize the expression above on $X(T)$.

$$k\big(X(T)-X(t)\big)=k\alpha(T-t)+\int_t^T \sigma_c dZ_c(u)+\delta(t)-\theta\delta(t)-(1-\theta)\alpha-\sigma_c e^{-kT}\int_t^T e^{ku}dZ_c(u)\,,$$

$$X(T)=X(t)+\alpha(T-t)-\frac{1}{k}(1-\theta)\alpha+\frac{1}{k}\delta(t)(1-\theta)+\frac{1}{k}\int_t^T \sigma_c dZ_c(u)-\frac{1}{k}\sigma_c e^{-kT}\int_t^T e^{ku}dZ_c(u)\,,$$

$$X(T)=X(t)+\alpha(T-t)+\frac{1}{k}(1-\theta)(\delta(t)-\alpha)+\frac{1}{k}\int_t^T \sigma_c dZ_c(u)-\frac{1}{k}\sigma_c e^{-kT}\int_t^T e^{ku}dZ_c(u)\,. \tag{70}$$

In the absence of arbitrage opportunities, and if the futures price is a martingale, then the Feynman-Kac theorem states that the futures price today is

$$F\big(S(T),\delta(T),t,T\big)=F\big(S(T),\delta(T),T\big)P(t,T)=E_t^*\big(S(T)\big)P(t,T)\,, \tag{71}$$

where $P(t,T)$ is the price of a risk-free pure discount bond at time t maturing at time T. Recall in the one factor model, we always found what the futures price is at time T. Here we are now discounting the futures price by the risk free rate to obtain today's price of a futures contract maturing at time T. To find the value today, we can discount the expected spot price by the known riskless rate of return because the futures price in expression (71) has been risk-adjusted.

In order to apply equation (71), we need to convert the stochastic processes for the spot price and convenience yield into equivalent martingales using the Girsanov theorem. The relation between the true probability measure and the martingale probability measure is

$$dZ_s(t) = dZ_s^*(t) - \left(\frac{\mu + \delta(t) - r}{\sigma_s} \right) dt, \tag{72}$$

$$dZ_c(t) = dZ_c^*(t) - \lambda dt, \tag{73}$$

where $dZ_s^*(t)$ and $dZ_c^*(t)$ are the transformed Wiener processes under the equivalent martingale probability measures. The coefficients on the dt terms are the market risk premiums associated with spot price and convenience yield respectively. Notice the risk premium for the spot price is expressed analytically. Gibson and Schwartz (1990) derive this value presuming the spot contract for the commodity complies with the models partial differential equation. The convenience yield is non-traded asset and its market price of risk cannot be solved analytically. As a consequence, we simply denote the market risk premium for the convenience yield with a λ.[14]

Working on an expression for the right-hand side of equation (71), we take equation (72) and substitute it into equation (53) obtaining

$$S(T) = S(t)\exp\left(\left[\mu - \frac{1}{2}\sigma_s^2 \right](T-t) + \sigma_s \int_t^T \left[dZ_s^*(t) - \left(\frac{\mu + \delta(t) - r}{\sigma_s} \right) du \right] \right).$$

Given the risk transformation of the spot price process, we see the convenience yield is now introduced into the spot price dynamics, as it should. The convenience impacts the expected return for the spot price, hence, when we use the risk premium to adjust the Brownian motion for the spot price, we bring the convenience yield into the analysis. That is, the convenience yield is one component of the total instantaneous return for the spot price and therefore a part of the market price of risk for the spot price. Rearranging the above yields

$$S(T) = S(t)\exp\left(\left[\mu - \frac{1}{2}\sigma_s^2\right](T-t) + \sigma_s \int_t^T dZ_s^*(u) - \sigma_s \int_t^T \left(\frac{\mu - r}{\sigma_s}\right) du - \int_t^T \delta(u) du\right). \quad (74)$$

Rearranging terms in (73) and inserting equation (68) into (73) yields

$$S(T) = S(t)\exp\left(\left[r - \frac{1}{2}\sigma_s^2\right](T-t) + \sigma_s \int_t^T dZ_s^*(u) - (X(T) - X(t))\right) \quad . \quad (75)$$

Note, as we sum over the spot price to obtain the terminal value of $S(T)$, the cumulative convenience impacts this value. This is as we expected.

Now, the transformed stochastic process for the spot price, $S(T)$ in equation (75), is still not a martingale. We have only risk-adjusted the spot price process and not the convenience yield process. Since the terminal spot price is a function of this state variable we need to transform the convenience yield as well. Taking equation (73) and substituting it into equation (70) we obtain

$$X(T) = X(t) + \alpha(T-t) + \frac{1}{k}(1-\theta)(\delta(t) - \alpha)$$

$$+ \frac{1}{k}\int_t^T \sigma_c\left(dZ_c^*(u) - \lambda du\right) - \frac{1}{k}\sigma_c e^{-kT}\int_t^T e^{ku}\left(dZ_c^*(u) - \lambda du\right). \quad (76)$$

Substitute the RHS of (76) into equation (75)

$$S(T)=S(t)\exp\left\{\left[r-\frac{1}{2}\sigma_s^2\right](T-t)+\sigma_s\int_t^T dZ_s^*(u)-\left[X(t)+\alpha(T-t)+\frac{1}{k}(1-\theta)(\delta(t)-\alpha)\right.\right.$$
$$\left.\left.+\frac{1}{k}\sigma_c\int_t^T dZ_c^*(u)-\frac{1}{k}\sigma_c\lambda\int_t^T du-\frac{1}{k}\sigma_c e^{-kT}\int_t^T e^{ku}dZ_c^*(u)+\frac{1}{k}\sigma_c\lambda e^{-kT}\int_t^T e^{ku}du-X(t)\right]\right\}. \tag{77}$$

Now, focusing our attention to the last bracket term in equation (77) we may write the expression as

$$\alpha(T-t)+\frac{1}{k}(1-\theta)(\delta(t)-\alpha)+\frac{1}{k}\sigma_c\int_t^T dZ_c^*(u)-\frac{1}{k}\sigma_c\lambda u\Big|_t^T$$
$$-\frac{1}{k}\sigma_c e^{-kT}\int_t^T e^{ku}dZ_c^*(u)+\frac{1}{k}\sigma_c\lambda e^{-kT}\frac{1}{k}e^{ku}\Big|_t^T.$$

Rearranging terms and evaluating the integrals yields

$$\alpha(T-t)+\frac{1}{k}(1-\theta)(\delta(t)-\alpha)-\frac{1}{k}\sigma_c\lambda(T-t)+\frac{1}{k^2}\sigma_c\lambda e^{-kT}\left(e^{kT}-e^{kt}\right)$$
$$+\frac{1}{k}\sigma_c\int_t^T dZ_c^*(u)-\frac{1}{k}\sigma_c e^{-kT}\int_t^T e^{ku}dZ_c^*(u),$$

$$\frac{1}{k}(1-\theta)(\delta(t)-\alpha)+\left(\alpha-\frac{1}{k}\sigma_c\lambda\right)(T-t)+\frac{1}{k^2}\sigma_c\lambda(1-\theta)$$
$$+\frac{1}{k}\sigma_c\int_t^T dZ_c^*(u)-\frac{1}{k}\sigma_c e^{-kT}\int_t^T e^{ku}dZ_c^*(u),$$

$$\frac{1}{k}\left(\delta(t)-\alpha+\frac{1}{k}\sigma_c\lambda\right)(1-\theta)+\left(\alpha-\frac{1}{k}\sigma_c\lambda\right)(T-t)$$
$$+\frac{1}{k}\sigma_c\int_t^T dZ_c^*(u)-\frac{1}{k}\sigma_c e^{-kT}\int_t^T e^{ku}dZ_c^*(u).$$

Substituting the expression above back into equation (77), the terminal spot price may be written as

$$S(T)=S(t)\exp\left\{\left[r-\frac{1}{2}\sigma_s^2\right](T-t)+\sigma_s\int_t^T dZ_s^*(u)-\left[\frac{1}{k}\left(\delta(t)-\alpha+\frac{1}{k}\sigma_c\lambda\right)(1-\theta)\right.\right.$$

$$\left.\left.+\left(\alpha-\frac{1}{k}\sigma_c\lambda\right)(T-t)+\frac{1}{k}\sigma_c\int_t^T dz_c^*(u)-\frac{1}{k}\sigma_c e^{-kT}\int_t^T e^{ku}dz_c^*(u)\right]\right\}.$$

Rearranging and grouping terms yields

$$S(T)=S(t)\exp\left\{-\left[\frac{1}{2}\sigma_s^2-r+\alpha-\frac{1}{k}\sigma_c\lambda\right](T-t)-\frac{1}{k}\left(\delta(t)-\alpha+\frac{1}{k}\sigma_c\lambda\right)(1-\theta)\right.$$

$$\left.+\sigma_s\int_t^T dZ_s^*(u)-\frac{1}{k}\sigma_c\int_t^T dZ_c^*(u)+\frac{1}{k}\sigma_c e^{-kT}\int_t^T e^{ku}dZ_c^*(u)\right\}. \quad (78)$$

From expression (71), we are interested in the discounted futures price. Therefore we multiply equation (78) by $P(t,T)$ to obtain

$$P(t,T)S(T)=S(t)\exp\left\{-\left[\frac{1}{2}\sigma_s^2+\alpha-\frac{1}{k}\sigma_c\lambda\right](T-t)-\frac{1}{k}\left(\delta(t)-\alpha+\frac{1}{k}\sigma_c\lambda\right)(1-\theta)\right.$$

$$\left.+\sigma_s\int_t^T dZ_s^*(u)-\frac{1}{k}\sigma_c\int_t^T dZ_c^*(u)+\frac{1}{k}\sigma_c e^{-kT}\int_t^T e^{ku}dZ_c^*(u)\right\}. \quad (79)$$

Simplifying equation (79), we have

$$P(t,T)S(T)=S(t)\exp[\hat{z}], \quad (80)$$

where

$$\hat{z}=-\left[\frac{1}{2}\sigma_s^2+\alpha-\frac{1}{k}\sigma_c\lambda\right](T-t)-\frac{1}{k}\left(\delta(t)-\alpha+\frac{1}{k}\sigma_c\lambda\right)(1-\theta)$$

$$+\sigma_s\int_t^T dZ_s^*(u)-\frac{1}{k}\sigma_c\int_t^T dZ_c^*(u)+\frac{1}{k}\sigma_c e^{-kT}\int_t^T e^{ku}dZ_c^*(u). \quad (81)$$

Given the result in equation (80) and (81), we may write equation (71) as

$$F(S(T),\delta(T),t,T)=E_t^*(P(t,T)S(T)),$$

$$= E_t^*\left(S(t)\exp(\hat{z})\right),$$

$$= S(t)E_t^*\left(\exp(\hat{z})\right) \quad . \tag{82}$$

From expression (81), we know $\hat{z}$ is a function of two standard Brownian motions. Given this, we know $\hat{z}$ is normally distributed and from equation (14), we may write (82) as

$$F\left(S(T), X(T), t, T\right) = S_t \exp\left(\hat{\mu} + \frac{1}{2}\hat{\sigma}^2\right) .^{15} \tag{83}$$

Where $\hat{\mu}$ and $\hat{\sigma}^2$ are the mean and variance respectively for $\hat{z}$. The mean is defined

$$\hat{\mu} \equiv E_t^*(\hat{z}),$$

$$= -\left[\frac{1}{2}\sigma_s^2 + \alpha - \frac{1}{k}\sigma_c^2\right](T-t) - \frac{1}{k}\left(\delta(t) - \alpha + \frac{1}{k}\sigma_c\lambda\right)(1-\theta)$$

$$+ \sigma_s\int_t^T E_t^*\left(dZ_s^*(t)\right) - \frac{1}{k}\sigma_c\int_t^T E_t^*\left(dZ_c^*(u)\right) + \frac{1}{k}\sigma_c e^{-kT}\int_t^T e^{ku}E_t^*\left(dZ_c^*(u)\right),$$

$$= -\left[\frac{1}{2}\sigma_s^2 + \alpha - \frac{1}{k}\sigma_c\lambda\right](T-t) - \frac{1}{k}\left(\delta(t) - \alpha + \frac{1}{k}\sigma_c\lambda\right)(1-\theta). \tag{84}$$

The reader should note that since we discounted the terminal spot price before taking the expectation, expression (84) is without the appreciation term found in the earlier pricing models. The variance is defined

$$\hat{\sigma}^2 \equiv E_t^*\left(\hat{z} - E_t^*(\hat{z})\right)^2,$$

$$= E_t^*\left(\sigma_s\int_t^T dZ_s^*(u) - \frac{1}{k}\sigma_c\int_t^T dZ_c^*(u) + \frac{1}{k}\sigma_c e^{-kT}\int_t^T e^{ku}dZ_c^*(u)\right)^2,$$

$$= E_t^*\left\{\left(\sigma_s\int_t^T dZ_s^*(u)\right)^2 + \left(\frac{1}{k}\sigma_c\int_t^T dZ_c^*(u)\right)^2 + \left(\frac{1}{k}\sigma_c e^{-kT}\int_t^T e^{ku}dZ_c^*(u)\right)^2\right.$$

$$-2\left(\sigma_s\int_t^T dZ_s^*(u)\right)\left(\frac{1}{k}\sigma_c\int_t^T dZ_c^*(u)\right)+2\left(\sigma_s\int_t^T dZ_s^*(u)\right)\left(\frac{1}{k}\sigma_c e^{-kT}\int_t^T e^{ku}dZ_c^*(u)\right)$$

$$-2\left(\frac{1}{k}\sigma_c\int_t^T dZ_c^*(u)\right)\left(\frac{1}{k}\sigma_c e^{-kT}\int_t^T e^{ku}dZ_c^*(u)\right)\Bigg\} \quad . \tag{85}$$

Using Ito's isometry to evaluate the six terms in equation (85), we have

$$E_t^*\left(\sigma_s\int_t^T dZ_s^*(u)\right)^2 = \sigma_s^2\left(\sigma_z\int_t^T du\right) = \sigma_s^2(1)(T-t) = \sigma_s^2(T-t), \tag{86}$$

$$E_t^*\left(\frac{1}{k}\sigma_c\int_t^T dZ_c^*(u)\right)^2 = \left(\frac{1}{k}\right)^2\sigma_c^2\left((1)\int_t^T du\right) = \left(\frac{1}{k}\right)^2\sigma_c^2(T-t), \tag{87}$$

$$E_t^*\left(\frac{1}{k}\sigma_c e^{-kT}\int_t^T e^{ku}dZ_c^*(u)\right)^2 = \left(\frac{1}{k}\right)^2\sigma_c^2 e^{-2kT}\left((1)\int_t^T e^{2ku}du\right) = \left(\frac{1}{k}\right)^2\frac{\sigma_c^2}{2k}e^{-2kT}\left(e^{2kT}-e^{2kt}\right),$$

$$=\left(\frac{1}{k}\right)^2\frac{\sigma_c^2}{2k}\left(1-\theta^2\right), \tag{88}$$

$$E_t^*\left(\sigma_s\int_t^T dZ_s^*(u)\right)\left(\frac{1}{k}\sigma_c\int_t^T dZ_c^*(u)\right) = \frac{1}{k}\sigma_c\sigma_s\rho_{cs}\int_t^T du = \frac{1}{k}\sigma_c\sigma_s\rho_{cs}(T-t), \tag{89}$$

$$E_t^*\left(\sigma_s\int_t^T dZ_s^*(u)\right)\left(\frac{1}{k}\sigma_c e^{-kT}\int_t^T e^{ku}dZ_c^*(u)\right) = \frac{1}{k}\sigma_c\sigma_s\rho_{cs}e^{-kT}\int_t^T e^{ku}du,$$

$$=\left(\frac{1}{k}\right)^2\sigma_c\sigma_s\rho_{cs}(1-\theta), \tag{90}$$

$$E_t^*\left(\frac{1}{k}\sigma_c\int_t^T dZ_c^*(u)\right)\left(\frac{1}{k}\sigma_c e^{-kT}\int_t^T e^{ku}dZ_c^*(u)\right) = \left(\frac{1}{k}\right)^2\sigma_c^2 e^{-kT}\int_t^T e^{ku}du,$$

$$=\left(\frac{1}{k}\right)^3\sigma_c^2(1-\theta). \tag{91}$$

Substitute equations (86) - (91) into equation (85) to obtain

$$\hat{\sigma}^2 = \sigma_s^2(T-t)+\left(\frac{1}{k}\right)^2\sigma_c^2(T-t)+\left(\frac{1}{k}\right)^2\frac{\sigma_c^2}{2k}\left(1-\theta^2\right)-2\left(\frac{1}{k}\right)\sigma_s\sigma_c\rho_{cs}(T-t)$$

$$+2\left(\frac{1}{k}\right)^2 \sigma_s\sigma_c\rho_{cs}(1-\theta)-2\left(\frac{1}{k}\right)^3 \sigma_c^2(1-\theta).$$

Collecting and rearranging terms in the expression above yields

$$\hat{\sigma}^2 = \left(\sigma_s^2 - 2\frac{1}{k}\sigma_s\sigma_c\rho_{cs} + \left(\frac{1}{k}\right)^2 \sigma_c^2\right)(T-t) + 2\left(\left(\frac{1}{k}\right)^2 \sigma_s\sigma_c\rho_{cs} - \left(\frac{1}{k}\right)^3 \sigma_c^2\right)(1-\theta)$$

$$+\left(\frac{1}{k}\right)^2 \frac{\sigma_c^2}{2k}\left(1-\theta^2\right). \tag{92}$$

Now, substituting equations (84) and (92) into (83) yields

$$F(S(T),\delta(T),t,T) = S(t)\exp\left\{-\left[\frac{1}{2}\sigma_s^2 + \alpha - \frac{1}{k}\sigma_c\lambda\right](T-t) - \frac{1}{k}\left(\delta(t) - \alpha + \frac{1}{k}\sigma_c\lambda\right)(1-\theta)\right.$$

$$+\frac{1}{2}\left[\left(\sigma_s^2 - 2\frac{1}{k}\sigma_c\sigma_s\rho_{cs} + \left(\frac{1}{k}\right)^2 \sigma_c^2\right)(T-t)\right.$$

$$\left.\left.+2\left(\left(\frac{1}{k}\right)^2 \sigma_c\sigma_s\rho_{cs} - \left(\frac{1}{k}\right)^3 \sigma_c^2\right)(1-\theta) + \left(\frac{1}{k}\right)^2 \frac{\sigma_c^2}{2k}\left(1-\theta^2\right)\right]\right\}. \tag{93}$$

Rearranging terms in the expression above yields

$$F(S(T),\delta(T),t,T) = S(t)\exp\left\{\left[-\alpha + \frac{1}{k}(\sigma_c\lambda - \sigma_c\sigma_s\rho_{cs}) + \frac{1}{2}\left(\frac{1}{k}\right)^2 \sigma_c^2\right](T-t)\right.$$

$$-\frac{1}{k}\left(\delta(t) - \alpha + \frac{1}{k}(\sigma_c\lambda - \sigma_c\sigma_s\rho_{cs}) + \left(\frac{1}{k}\right)^2 \sigma_c^2\right)(1-\theta)$$

$$\left.+\left(\frac{1}{k}\right)^2 \frac{\sigma_c^2}{4k}\left(1-\theta^2\right)\right\}. \tag{94}$$

Equation (94) is the solution to the partial differential equation below

$$\frac{1}{2}F_{ss}\sigma_s^2[S(t)]^2 + F_{s\delta}S(t)\sigma_c\sigma_s\rho_{cs} + \frac{1}{2}F_{\delta\delta}\sigma_c^2 + F_s S(t)(\mathrm{r} - \delta(t))$$

$$+F_\delta\big(k(\alpha-\delta(t))-\lambda\sigma_c\big)+F_t-rF=0\ . \qquad (95)$$

This is a nice result. As we saw in chapter 2, oil inventories fluctuated and influenced the spot price of oil. In particular, these oil reserves seem to follow a mean reverting pattern. It is this phenomenon we wish to capture, so we may price commodity futures contracts more accurately. The expression developed in equation (94) is a powerful result. That is, we now have an analytical solution that captures movements in the futures prices not just from one source but two: the spot price and the convenience yield. We should expect, the two-factor model to price commodity futures contracts more accurately than a one factor model.

It is easy to verify equation (94) as a solution to (95). Finding the partial derivatives of $F\big(S(T),\delta(T),t,T\big)$ we have

$$F_s=\exp\{\hat{z}\}, \qquad (96)$$

$$F_{ss}=0, \qquad (97)$$

$$F_\delta=-\left(\frac{1}{k}\right)(1-\theta)F, \qquad (98)$$

$$F_{\delta\delta}=\left(\frac{1}{k}\right)^2(1-\theta)^2F\ , \qquad (99)$$

$$F_{s\delta}=-(1-\theta)\frac{1}{k}\frac{1}{S(t)}F, \qquad (100)$$

$$F_t=\Bigg[\alpha-\frac{1}{k}\sigma_c\lambda+\frac{1}{k}\sigma_c\sigma_s\rho_{cs}-\frac{1}{2}\left(\frac{1}{k}\right)^2\sigma_c^2+\delta(t)\theta-\alpha\theta+\frac{1}{k}\sigma_c\lambda\theta-\frac{1}{k}\sigma_c\sigma_s\rho_{cs}\theta$$

$$+\left(\frac{1}{k}\right)^2\sigma_c^2\theta-\frac{1}{2}\left(\frac{1}{k}\right)^2\sigma_c^2\theta^2\Bigg]F. \qquad (101)$$

Insert equations (96)-(101) into (95) yields

$$\frac{1}{2}\sigma_s^2[S(t)]^2(0)+S(t)\sigma_c\sigma_s\rho_{cs}\left(-(1-\theta)\frac{1}{k}\frac{1}{S(t)}F\right)+\frac{1}{2}\sigma_c^2\left(\frac{1}{k}\right)^2(1-\theta)^2F$$

$$+S(t)\exp\{\bar{z}\}(\mathrm{r}-\delta(t))+(k(\alpha-\delta(t))-\lambda\sigma_c)\left(-\left(\frac{1}{k}\right)(1-\theta)F\right)$$

$$+\left[\alpha-\frac{1}{k}\sigma_c\lambda+\frac{1}{k}\sigma_c\sigma_s\rho_{cs}-\frac{1}{2}\left(\frac{1}{k}\right)^2\sigma_c^2+\delta(t)\theta-\alpha\theta+\frac{1}{k}\sigma_c\lambda\theta\right.$$

$$\left.-\frac{1}{k}\sigma_c\sigma_s\rho_{cs}\theta+\left(\frac{1}{k}\right)^2\sigma_c^2\theta-\frac{1}{2}\left(\frac{1}{k}\right)^2\sigma_c^2\theta^2\right]F-rF=0. \qquad (102)$$

Performing the arithmetic above, the equality in equation (100) holds. Therefore, equation (94) is a solution to the partial differential equation in (95).

Once again, the result above is expected. The partial differential equation in (95) is implied in the transformed futures price given by the Feynman-Kac solution. Consider the following illustration. Take the total differential of the discounted value for the futures

$$d\left(e^{-rt}F(\ )\right)=\frac{\partial e^{-rt}F(\ )}{\partial e^{-rt}}de^{-rt}+\frac{\partial e^{-rt}F(\ )}{\partial F(\ )}dF(\ ),$$

$$=Fde^{-rt}+e^{-rt}dF(\ ),$$

$$=F(-r)e^{-rt}dt+e^{-rt}dF(\ ). \qquad (103)$$

If $F(S(T),\delta(T),t,T)$ is a twice differentiable function with respect to $S(t)$, $\delta(t)$ and time, we may use Ito's lemma to write the increment of $F(S(T),\delta(T),t,T)$ as

$$=-re^{-rt}Fdt+e^{-rt}\left(F_tdt+F_sdS(t)+\frac{1}{2}F_{ss}[dS(t)]^2+F_{s\delta}dS(t)d\delta(t)\right.$$

$$\left.+\frac{1}{2}F_{\delta\delta}[d\delta(t)]^2+F_\delta d\delta(t)\right)\quad . \qquad (104)$$

Substituting the dynamics for the state variables in equation (104) yields

$$= -re^{-rt}Fdt + e^{-rt}\left(F_t dt + F_s\left(\mu S(t)dt + \sigma_s S(t)dZ_s(t)\right) + \frac{1}{2}F_{ss}\sigma_s^2[S(t)]^2 dt\right.$$

$$\left. + F_{s\delta}\sigma_c\sigma_s\rho_{cs}S(t)dt + \frac{1}{2}F_{\delta\delta}\sigma_c^2 dt + F_\delta\left(k(\alpha - \delta(t))dt + \sigma_c dZ_c(t)\right)\right). \quad (105)$$

The expression in (105) is not a risk adjusted process. The Girsanov theorem provides the risk neutral transformations for the driving Wiener processes in (105). These are

$$dZ_s(t) = dZ_s^*(t) - \frac{\mu + \delta(t) - r}{\sigma_s}dt, \quad (106)$$

$$dZ_c(t) = dZ_c^*(t) - \lambda dt. \quad (107)$$

Substituting (106) and (107) into (105) yields

$$= -re^{-rt}Fdt + e^{-rt}\left(F_t dt + F_s\left(\mu S(t)dt + \sigma_s S(t)\left(dZ_s^*(t) - \frac{\mu + \delta(t) - r}{\sigma_s}dt\right)\right)\right.$$

$$+ \frac{1}{2}F_{ss}\sigma_s^2[S(t)]^2 dt + F_{S\delta}\sigma_c\sigma_s\rho_{cs}S(t)dt + \frac{1}{2}F_{\delta\delta}\sigma_c^2 dt$$

$$\left. F_\delta\left(k(\alpha - \delta(t))dt + \sigma_c\left(dZ_c^*(t) - \lambda dt\right)\right)\right). \quad (108)$$

Rearranging terms in (108) yields

$$d\left(e^{-rt}F(S(T),\delta(T),t,T)\right) = e^{-rt}\left\{\left(\frac{1}{2}F_{ss}\sigma_s^2[S(t)]^2 + F_{s\delta}S(t)\sigma_c\sigma_s\rho_{cs} + \frac{1}{2}F_{\delta\delta}\sigma_c^2\right.\right.$$

$$\left. + F_s S(t)(\mathrm{r} - \delta(t)) + F_\delta\left(k(\alpha - \delta(t)) - \lambda\sigma_c\right) + F_t - rF\right)dt$$

$$\left. + \sigma_s S(t)F_s dZ_s^*(t) + \sigma_c F_\delta dZ_c^*(t)\right\} \quad . \quad (109)$$

If (109) is a martingale, then the expected value must be zero. This implies

$$E_t^*\left[d\left(e^{-rt}F(S(T),\delta(T),t,T)\right)\right] = E_t^*\left[e^{-rt}\left\{\left(\frac{1}{2}F_{ss}\sigma_s^2[S(t)]^2 + F_{s\delta}S(t)\sigma_c\sigma_s\rho_{cs} + \frac{1}{2}F_{\delta\delta}\sigma_c^2\right.\right.\right.$$

$$+ F_s S(t)(\mathrm{r} - \delta(t)) + F_\delta\left(k(\alpha - \delta(t)) - \lambda\sigma_c\right) + F_t - rF\Big)dt$$

$$+\sigma_s S(t)F_s dZ_s^*(t)+\sigma_c F_\delta dZ_c^*(t)\}\big]=0,$$

$$=E_t^*\Big[e^{-rt}\Big\{\Big(\frac{1}{2}F_{ss}\sigma_s^2[S(t)]^2+F_{s\delta}S(t)\sigma_c\sigma_s\rho_{cs}+\frac{1}{2}F_{\delta\delta}\sigma_c^2$$

$$+F_sS(t)(\mathrm{r}-\delta(t))+F_\delta\big(k(\alpha-\delta(t))-\lambda\sigma_c\big)+F_t-rF\Big)dt\Big\}\Big]$$

$$+e^{-rt}\sigma_sS(t)F_sE_t^*\big[dZ_s^*(t)\big]+e^{-rt}\sigma_cF_\delta E_t^*\big[dZ_c^*(t)\big]=0,$$

$$=e^{-rt}\Big\{\Big(\frac{1}{2}F_{ss}\sigma_s^2[S(t)]^2+F_{s\delta}S(t)\sigma_c\sigma_s\rho_{cs}+\frac{1}{2}F_{\delta\delta}\sigma_c^2$$
$$+F_sS(t)(\mathrm{r}-\delta(t))+F_\delta\big(k(\alpha-\delta(t))-\lambda\sigma_c\big)+F_t-rF\Big)dt\Big\}=0.$$

Therefore, the drift term's coefficient must equal zero. That is,

$$\frac{1}{2}F_{ss}\sigma_s^2[S(t)]^2+F_{s\delta}S(t)\sigma_c\sigma_s\rho_{cs}+\frac{1}{2}F_{\delta\delta}\sigma_c^2+F_sS(t)(\mathrm{r}-\delta(t))$$

$$+F_\delta\big(k(\alpha-\delta(t))-\lambda\sigma_c\big)+F_t-rF=0\ .$$

This is identical to expression (95). This result is expected. Recall, the Feynman-Kac theorem shows a correspondence between a conditional expectation and a particular class of partial differential equations. We took the solution of a risk neutral conditional expectation and were able to obtain the associate partial differential equation. Thus, we have illustrated that both method yield the same result.

3.3 Three-factor model

So far we have two cases that model the stochastic behavior of commodity prices that take into account mean reversion. We now consider a third model by extending the two-factor model to include a stochastic interest rate. Up to now, we have held interest rates constant, and as such, we have treated forward and futures contracts synonymously.[16] This will no longer be the case. Since futures prices are *marked to* *market* each day a variable interest rate will necessarily change the value of the futures contract over time. Therefore, new pricing solutions for both the forwards and futures contracts need to be derived.

3.31 Forward Prices

Consider the following joint stochastic process for the spot price, convenience yield and interest rate

$$dS(t) = (r(t) - \delta(t))S(t)dt + \sigma_s S(t)dZ_s^*(t), \tag{110}$$

$$d\delta(t) = (k_c(\alpha - \delta(t)) - \lambda\sigma_c)dt + \sigma_c dZ_c^*(t), \tag{111}$$

$$dr(t) = \left(f_t(s,t) + k_r f(s,t) + \frac{\sigma_r^2}{2k_r}\left(1 - e^{-2k_r(t-s)}\right) - k_r r(t) \right)dt + \sigma_r dZ_r^*(t), \tag{112}$$

where σ_s, σ_c and σ_r are the diffusion coefficients for the commodity spot price, the convenience yield and the interest rate processes, respectively. k_c and k_r are the speed of adjustment factors for the convenience yield, $\delta(t)$, and interest rate $r(t)$, α is the long run mean for the convenience yield. $dZ_s^*(t)$, $dZ_c^*(t)$ and $dZ_r^*(t)$ are the risk-adjusted increments of standard Brownian motions for the spot price, convenience yield and interest rate respectively. $Cov_t\lfloor dZ_s^*(t)dZ_c^*(t)\rfloor = \rho_{sc}dt$, $Cov_t\lfloor dZ_s^*(t)dZ_r^*(t)\rfloor = \rho_{rs}dt$, and $Cov_t\lfloor dZ_r^*(t)dZ_c^*(t)\rfloor = \rho_{rc}dt$. ρ_{sc}, ρ_{sr}, and ρ_{rc} are the correlation coefficients between the risk-adjusted Wiener processes above. $f(s,t)$ is the instantaneous forward rate, and $f_t(s,t)$ is the derivative of the instantaneous forward rate with respect to maturity.[17]

Notice in the specification above, the model starts with the risk-adjusted processes for the state variables. The current presentation is different than the derivations for the one-factor and two-factor model. Recall in those derivations the state variables are not adjusted until we find solutions to the respective stochastic differential equations. It turns out that the timing of the risk-neutral transformation does not alter the solution for the pricing equation. That is, we could have easily started with the risk-adjusted state variables and still arrived to the same solution. With this said, we start with the risk-neutral state variables and derive solutions to both forward and futures prices.

The pricing equation for the three-factor forward contract is given by the Feynman-Kac solution

$$J(S(T),\delta(T),r(T),T) = E_t^*(S(T))\frac{1}{P(t,T)}. \tag{113}$$

Expression (113) shows the forward price to be the future value of the risk-adjusted expected spot price. In the earlier models, we expressed the futures/forward price to be the risk-adjusted expected value. For each of those models we presume the interest is deterministic, here it is not. Our model needs to account for random movements in the interest rate. This is discussed below.

To model the forward price, we start with the transformation $X(t) = \ln S(t) - \int_t^T r(v)dv$. This transformed process for the forward price is different from the process used in the derivation for the two-factor model, where we use $X(t) = \ln S(t)$. The reason for the asymmetric treatment between the two-factor model and the three-factor model is due to the trading practices of a forward contract. Forward contracts are instruments that are directly negotiated between two parties. Since these instruments do not trade on an exchange, they are not ***marked-to-market***, which means random changes in the interest rate will not impact the forward price. Therefore, if we allow interest rates to move randomly, then we must account for this when modeling the forward price.

A natural question then is why model forward contracts with a variable interest rate? In a financial market interest rates fluctuate randomly. Any model that treats interest rates as a constant, fails to incorporate a facet of known market behavior to their model. For the present case, interest rates are random but as we will see they do not impact the price of a forward contract.

Expression (113) shows the forward price is simply the risk-neutral expected value of the spot price. To evaluate this expression, we start with the transformation $X(t) = \ln S(t) - \int_t^T r(v)dv$. Applying Ito's lemma to the transformation, the stochastic differential for $X(t)$ is shown to be

$$dX(t) = X_s dS(t) + \frac{1}{2} X_{ss} [dS(t)]^2 - d\left[\int_t^T r(v)dv\right],$$

$$dX(t) = \frac{1}{S(t)}\left[(r(t) - \delta(t))S(t)dt + \sigma_s S(t)dZ_s^*(t)\right] - \frac{1}{2}\left(\frac{1}{[S(t)]^2}\right)[S(t)]^2 \sigma_s^2 dt - d\left[\int_t^T r(v)dv\right],$$

$$dX(t) = \left[r(t) - \delta(t) - \frac{1}{2}\sigma_s^2\right]dt + \sigma_s dZ_s^*(t) - d\left[\int_t^T r(v)dv\right]. \qquad (114)$$

We integrate over equation (114) to obtain

$$\int_t^T dX(v) = \int_t^T r(v)dv - \int_t^T \delta(v)dv - \frac{1}{2}\sigma_s^2 \int_t^T dv + \sigma_s \int_t^T dZ_s^*(v) - \int_t^T r(v)dv,$$

$$X(T) - X(t) = \int_t^T r(v)dv - \int_t^T \delta(v)dv - \frac{1}{2}\sigma_s^2 \int_t^T dv + \sigma_s \int_t^T dZ_s^*(v) - \int_t^T r(v)dv,$$

$$X(T) = \ln S(t) - \frac{1}{2}\sigma_s^2 (T - t) - \int_t^T \delta(v)dv + \sigma_s \int_t^T dZ_s^*(v). \qquad (115)$$

The distribution of $X(T)$ is normal, and in this case, the solution for the forward price is

$$J(S(T),\delta(T),r(T),T)=e^{E_t^*[X(T)]+\frac{1}{2}V_t^*[X(T)]}\frac{1}{P(t,T)}. \quad (116)$$

Expression (116) shows the solution for the forward contract as a function of the first and second moment of $X(T)$. Therefore, we need the mean and variance of $X(T)$. The expected value of $X(T)$ is

$$E_t^*[X(T)]=\ln S(t)-\frac{1}{2}\sigma_s^2(T-t)-\int_t^T E_t^*[\delta(v)]dv. \quad (117)$$

To solve the integral in equation (117) we need the solution for the risk-neutralized convenience yield. The solution is found as follows.

$$d\delta(t)=(k_c(\alpha-\delta(t))-\lambda\sigma_c)dt+\sigma_c dZ_c^*(t),$$

$$d\delta(t)+k_c\delta(t)dt=(k_c\alpha-\lambda\sigma_c)dt+\sigma_c dZ_c^*(t),$$

$$e^{k_c t}[d\delta(t)+k_c\delta(t)dt]=e^{k_c t}[(k_c\alpha-\lambda\sigma_c)dt+\sigma_c dZ_c^*(t)],$$

$$d[\delta(t)e^{k_c t}]=e^{k_c t}[(k_c\alpha-\lambda\sigma_c)dt+\sigma_c dZ_c^*(t)],$$

$$\int_t^v d[\delta(s)e^{k_c s}]=\int_t^v e^{k_c s}(k_c\alpha-\lambda\sigma_c)ds+\sigma_c\int_t^v e^{k_c s}dZ_c^*(s),$$

$$e^{k_c v}\delta(v)-e^{k_c t}\delta(t)=e^{k_c v}\left(\alpha-\frac{\lambda\sigma_c}{k_c}\right)-e^{k_c t}\left(\alpha-\frac{\lambda\sigma_c}{k_c}\right)+\sigma_c\int_t^v e^{k_c s}dZ_c^*(s),$$

$$\delta(v)=e^{-k_c(v-t)}\delta(t)+\left(\alpha-\frac{\lambda\sigma_c}{k_c}\right)-\left(\alpha-\frac{\lambda\sigma_c}{k_c}\right)e^{-k_c(v-t)}+\sigma_c e^{-k_c v}\int_t^v e^{k_c s}dZ_c^*(s). \quad (118)$$

Taking the expected value of (118) and plugging it into (117) yields

$$E_t^*[X(T)]=\ln S(t)-\frac{1}{2}\sigma_s^2(T-t)-\int_t^T\left(e^{-k_c(v-t)}\delta(t)+\left(\alpha-\frac{\lambda\sigma_c}{k_c}\right)-\left(\alpha-\frac{\lambda\sigma_c}{k_c}\right)e^{-k_c(v-t)}\right)dv,$$

$$= \ln S(t) - \frac{1}{2}\sigma_s^2(T-t) - \left[-\frac{1}{k_c}\delta(t)e^{-k_c(v-t)}\Big|_t^T + \left(\alpha - \frac{\lambda\sigma_c}{k_c}\right)(T-t) + \frac{1}{k_c}\left(\alpha - \frac{\lambda\sigma_c}{k_c}\right)e^{-k_c(v-t)}\Big|_t^T \right]$$

$$= \ln S(t) - \frac{1}{2}\sigma_s^2(T-t) - \left[-\frac{1}{k_c}\delta(t)\left(e^{-k_c(T-t)} - e^{-k_c(t-t)}\right) + \left(\alpha - \frac{\lambda\sigma_c}{k_c}\right)(T-t) \right.$$

$$\left. + \frac{1}{k_c}\left(\alpha - \frac{\lambda\sigma_c}{k_c}\right)\left(e^{-k_c(T-t)} - e^{-k_c(t-t)}\right) \right],$$

$$= \ln S(t) - \frac{1}{2}\sigma_s^2(T-t) - \left[-\frac{1}{k_c}\delta(t)\left(e^{-k_c(T-t)} - 1\right) + \left(\alpha - \frac{\lambda\sigma_c}{k_c}\right)(T-t) \right.$$

$$\left. + \frac{1}{k_c}\left(\alpha - \frac{\lambda\sigma_c}{k_c}\right)\left(e^{-k_c(T-t)} - 1\right) \right],$$

$$= \ln S(t) - \frac{1}{2}\sigma_s^2(T-t) - \left[\frac{1}{k_c}\delta(t)\left(1 - e^{-k_c(T-t)}\right) + \left(\alpha - \frac{\lambda\sigma_c}{k_c}\right)(T-t) \right.$$

$$\left. - \frac{1}{k_c}\left(\alpha - \frac{\lambda\sigma_c}{k_c}\right)\left(1 - e^{-k_c(T-t)}\right) \right],$$

$$= \ln S(t) - \frac{1}{2}\sigma_s^2(T-t) - \left(\alpha - \frac{\lambda\sigma_c}{k_c}\right)(T-t) - \left(\delta(t) - \alpha + \frac{\lambda\sigma_c}{k_c}\right)\left(\frac{1 - e^{-k_c(T-t)}}{k_c}\right)$$

$$= \ln S(t) - \frac{1}{2}\sigma_s^2\tau - \left(\alpha - \frac{\lambda\sigma_c}{k_c}\right)\tau - \left(\delta(t) - \alpha + \frac{\lambda\sigma_c}{k_c}\right)(H_c(\tau)). \qquad (119)$$

where $\tau = T - t$ and $H_c(\tau) = \dfrac{1 - e^{-k_c\tau}}{k_c}$. The variance for equation (116) is

$$V_t^*[X(T)] = V_t^*\left[\ln S(t) - \frac{1}{2}\sigma_s^2(T-t) - \int_t^T \delta(v)dv + \sigma_s \int_t^T dZ_s^*(v) \right],$$

$$= V_t^*\left[\int_t^T \delta(v)dv\right] + V_t^*\left[\sigma_s \int_t^T dZ_s^*(v)\right] - 2Cov_t^*\left[\int_t^T \delta(v)dv, \sigma_s \int_t^T dZ_s^*(v)\right]. \qquad (120)$$

The variance for $X(T)$ follows from Ito's isometry and the variances of the risk neutralized convenience yield and spot price. The variance for the spot price is

$$V_t^*\left[\sigma_s\int_t^T dZ_s^*(v)\right]=\sigma_s^2(1)\int_t^T dv=\sigma_s^2(T-t)=\sigma_s^2\tau\ . \tag{121}$$

Working on the variance for the convenience yield

$$V_t^*\left[\int_t^T \delta(v)dv\right]=V_t^*\left[\int_t^T \sigma_c e^{-k_c v}\int_t^v e^{k_c s}dZ_c^*(s)dv\right].$$

Switching the order of integration[18]

$$=V_t^*\left[\int_t^T e^{k_c s}\int_s^T \sigma_c e^{-k_c v}dv dZ_c^*(s)\right],$$

$$=V_t^*\left[\int_t^T e^{k_c s}\left[-\frac{\sigma_c}{k_c}\left(e^{-k_c T}-e^{-k_c s}\right)\right]dZ_c^*(s)\right],$$

$$=V_t^*\left[\int_t^T e^{k_c s}\left[\frac{\sigma_c}{k_c}\left(e^{-k_c s}-e^{-k_c T}\right)\right]dZ_c^*(s)\right],$$

$$=V_t^*\left[\int_t^T \frac{\sigma_c}{k_c}\left(1-e^{-k_c(T-s)}\right)dZ_c^*(s)\right].$$

Now using Ito's isometry we obtain

$$=\frac{\sigma_c^2}{k_c^2}\int_t^T\left(1-e^{-k_c(T-s)}\right)^2 ds,$$

$$=\frac{\sigma_c^2}{k_c^2}\int_t^T\left(1-2e^{-k_c(T-s)}+e^{-2k_c(T-s)}\right)ds,$$

$$=\frac{\sigma_c^2}{k_c^2}(T-t)-\frac{2\sigma_c^2}{k_c^2}\left(\frac{1}{k_c}\right)\left(e^{-k_c(T-T)}-e^{-k_c(T-t)}\right)+\frac{\sigma_c^2}{k_c^2}\left(\frac{1}{2k_c}\right)\left(e^{-2k_c(T-T)}-e^{-2k_c(T-t)}\right),$$

$$=\frac{\sigma_c^2}{k_c^2}(T-t)-\frac{2\sigma_c^2}{k_c^2}\left(\frac{1}{k_c}\right)\left(1-e^{-k_c(T-t)}\right)+\frac{\sigma_c^2}{k_c^2}\left(\frac{1}{2k_c}\right)\left(1-e^{-2k_c(T-t)}\right),$$

$$= \frac{\sigma_c^2}{k_c^2}\tau - \frac{2\sigma_c^2}{k_c^2}\left(\frac{1}{k_c}\right)\left(1 - e^{-k_c\tau}\right) + \frac{\sigma_c^2}{k_c^2}\left(\frac{1}{2k_c}\right)\left(1 - e^{-2k_c\tau}\right),$$

$$= \frac{\sigma_c^2}{k_c^2}\tau - \frac{\sigma_c^2}{k_c^2}\left(\frac{1}{k_c}\right)\left(2 - 2e^{-k_c\tau}\right) + \frac{\sigma_c^2}{k_c^2}\left(\frac{1}{2k_c}\right)\left(1 - e^{-2k_c\tau}\right),$$

$$= \frac{\sigma_c^2}{k_c^2}\tau - \frac{\sigma_c^2}{k_c^3}\left(1 - e^{-k_c\tau}\right) - \frac{\sigma_c^2}{k_c^3}\left(1 - e^{-k_c\tau}\right) + \frac{\sigma_c^2}{2k_c^3}\left(1 - e^{-2k_c\tau}\right),$$

$$= \frac{\sigma_c^2}{k_c^2}\tau - \frac{\sigma_c^2}{k_c^3}\left(1 - e^{-k_c\tau}\right) + \frac{\sigma_c^2}{2k_c^3}\left(1 - e^{-2k_c\tau} + 2e^{-k_c\tau} - 2\right),$$

$$= \frac{\sigma_c^2}{k_c^2}\tau - \frac{\sigma_c^2}{k_c^3}\left(1 - e^{-k_c\tau}\right) + \frac{\sigma_c^2}{2k_c^3}\left(-1 - e^{-2k_c\tau} + 2e^{-k_c\tau}\right),$$

$$= \frac{\sigma_c^2}{k_c^2}\tau - \frac{\sigma_c^2}{k_c^3}\left(1 - e^{-k_c\tau}\right) - \frac{\sigma_c^2}{2k_c^3}\left(1 - 2e^{-k_c\tau} + e^{-2k_c\tau}\right),$$

$$= \frac{\sigma_c^2}{k_c^2}\tau - \frac{\sigma_c^2}{k_c^3}\left(1 - e^{-k_c\tau}\right) - \frac{\sigma_c^2}{2k_c^3}\left(1 - e^{-k_c\tau}\right)^2,$$

$$= \frac{\sigma_c^2}{k_c^2}\tau - \frac{\sigma_c^2 H_c(\tau)}{k_c^2} - \frac{\sigma_c^2 H_c^2(\tau)}{2k_c},$$

$$= -\left(H_c(\tau) - \tau\right)\frac{\sigma_c^2}{k_c^2} - \frac{\sigma_c^2 H_c^2(\tau)}{2k_c} \quad . \tag{122}$$

The covariance between the spot price and convenience yield is

$$Cov_t^*\left[\int_t^T \delta(v)dv, \sigma_s \int_t^T dZ_s^*(v)\right] = Cov_t^*\left[\int_t^T \sigma_c e^{-k_c v}\int_t^v e^{k_c s} dZ_c^*(s)dv, \sigma_s \int_t^T dZ_s^*(v)\right],$$

$$= Cov_t^*\left[\int_t^T e^{k_c s}\int_s^T \sigma_c e^{-k_c v} dv dZ_c^*(s), \sigma_s \int_t^T dZ_s^*(v)\right],$$

$$= Cov_t^*\left[\int_t^T e^{k_c s}\left[-\frac{\sigma_c}{k_c}\left(e^{-k_c T} - e^{-k_c s}\right)\right] dZ_c^*(s), \sigma_s \int_t^T dZ_s^*(v)\right],$$

$$= Cov_t^*\left[\frac{\sigma_c}{k_c}\int_t^T\left(1-e^{-k_c(T-s)}\right)dZ_c^*(s), \sigma_s\int_t^T dZ_s^*(v)\right].$$

Using Ito's isometry we get

$$= \frac{\sigma_s\sigma_c\rho_{cs}}{k_c}\int_t^T\left(1-e^{-k_c(T-s)}\right)ds,$$

$$= \frac{\sigma_s\sigma_c\rho_{cs}}{k_c}\left((T-t)-\left(\frac{1}{k_c}\right)\left(e^{-k_c(T-T)}-e^{-k_c(T-t)}\right)\right),$$

$$= \frac{\sigma_s\sigma_c\rho_{cs}}{k_c}\left((T-t)-\left(\frac{1}{k_c}\right)\left(1-e^{-k_c(T-t)}\right)\right),$$

$$= \frac{\sigma_s\sigma_c\rho_{cs}}{k_c}(\tau - H_c(\tau)) \quad . \tag{123}$$

Substituting (121), (122) and (123) back into (120) yields

$$V_t^*[X(T)] = \left[-(H_c(\tau)-\tau)\frac{\sigma_c^2}{k_c^2} - \frac{\sigma_c^2 H_c^2(\tau)}{2k_c}\right] + \sigma_s^2\tau - \frac{2\sigma_s\sigma_c\rho_{cs}}{k_c}(\tau - H_c(\tau)). \tag{124}$$

Taking the expressions in (119) and (124), and substituting into expression (113) yields

$$J(S(T),\delta(T),r(T),T) = \exp\left\{\ln S(t) - \frac{1}{2}\sigma_s^2\tau - \left(\alpha - \frac{\lambda\sigma_c}{k_c}\right)\tau - \left(\delta(t) - \alpha + \frac{\lambda\sigma_c}{k_c}\right)(H_c(\tau))\right.$$
$$\left. + \frac{1}{2}\left(\left[-(H_c(\tau)-\tau)\frac{\sigma_c^2}{k_c^2} - \frac{\sigma_c^2 H_c^2(\tau)}{2k_c}\right] + \sigma_s^2\tau - \frac{2\sigma_s\sigma_c\rho_{cs}}{k_c}(\tau - H_c(\tau))\right)\right\}\frac{1}{P(t,T)}.$$

Rearranging the above yields

$$J(S(T),\delta(T),r(T),T) = S(t)\exp\left\{-\left(\alpha - \frac{\lambda\sigma_c}{k_c}\right)\tau - \left(\delta(t) - \alpha + \frac{\lambda\sigma_c}{k_c}\right)(H_c(\tau))\right.$$
$$\left. + \frac{1}{2}\left(\left[-(H_c(\tau)-\tau)\frac{\sigma_c^2}{k_c^2} - \frac{\sigma_c^2 H_c^2(\tau)}{2k_c}\right] - \frac{2\sigma_s\sigma_c\rho_{cs}}{k_c}(\tau - H_c(\tau))\right)\right\}\frac{1}{P(t,T)}$$

$$= S(t)\exp\left\{-\left(\frac{k_c\alpha}{k_c}-\frac{\lambda\sigma_c}{k_c}\right)\tau-\left(\delta(t)-\frac{k_c\alpha}{k_c}+\frac{\lambda\sigma_c}{k_c}\right)(H_c(\tau))\right.$$

$$\left.+\frac{1}{2}\left(\left[-(H_c(\tau)-\tau)\frac{\sigma_c^2}{k_c^2}-\frac{\sigma_c^2H_c^2(\tau)}{2k_c}\right]-\frac{2\sigma_s\sigma_c\rho_{cs}}{k_c}(\tau-H_c(\tau))\right)\right\}\frac{1}{P(t,T)},$$

$$= S(t)\exp\left\{-\left(\frac{k_c\alpha-\lambda\sigma_c}{k_c}\right)\tau-\left(\delta(t)-\frac{k_c\alpha-\lambda\sigma_c}{k_c}\right)(H_c(\tau))\right.$$

$$\left.+\frac{1}{2}\left(-(H_c(\tau)-\tau)\left(\frac{\sigma_c^2}{k_c^2}-\frac{2\sigma_s\sigma_c\rho_{cs}}{k_c}\right)-\frac{\sigma_c^2H_c^2(\tau)}{2k_c}\right)\right\}\frac{1}{P(t,T)},$$

$$= S(t)\exp\left\{(H_c(\tau)-\tau)\left(\frac{k_c\alpha-\lambda\sigma_c}{k_c}\right)-\delta(t)H_c(\tau)\right.$$

$$\left.-(H_c(\tau)-\tau)\left(\frac{\frac{\sigma_c^2}{2}}{k_c^2}-\frac{k_c\sigma_s\sigma_c\rho_{cs}}{k_c^2}\right)-\frac{\sigma_c^2H_c^2(\tau)}{4k_c}\right\}\frac{1}{P(t,T)},$$

$$= S(t)\exp\left\{(H_c(\tau)-\tau)\left(\frac{k_c^2\alpha-k_c\lambda\sigma_c}{k_c^2}\right)-\delta(t)H_c(\tau)\right.$$

$$\left.-(H_c(\tau)-\tau)\left(\frac{\frac{\sigma_c^2}{2}-k_c\sigma_s\sigma_c\rho_{cs}}{k_c^2}\right)-\frac{\sigma_c^2H_c^2(\tau)}{4k_c}\right\}\frac{1}{P(t,T)},$$

$$= S(t)\exp\left\{(H_c(\tau)-\tau)\left(\frac{k_c^2\alpha-k_c\lambda\sigma_c-\frac{\sigma_c^2}{2}+k_c\sigma_s\sigma_c\rho_{cs}}{k_c^2}\right)-\delta(t)H_c(\tau)\right.$$

$$\left.-\frac{\sigma_c^2H_c^2(\tau)}{4k_c}\right\}\frac{1}{P(t,T)}.$$

Finally we obtain

$$J(S(t),\delta(t),r(T),T) = S(t)\exp\left\{\frac{(H_c(\tau)-\tau)\left(k_c^2\alpha - k_c\lambda\sigma_c - \frac{\sigma_c^2}{2} + \rho_{cs}\sigma_s\sigma_c k_c\right)}{k_c^2} - \frac{\sigma_c^2 H_c^2(\tau)}{4k_c}\right.$$

$$\left. - H_c(\tau)\delta(t)\right\}\frac{1}{P(t,T)}.$$

Simplifying

$$J(S(t),\delta(t),r(T),T) = S(t)A(\tau)e^{-H_c(\tau)\delta(t)}\frac{1}{P(t,T)}, \tag{125}$$

where $A(\tau) = \exp\left(\frac{(H_c(\tau)-\tau)\left(k_c^2\alpha - k_c\lambda\sigma_c - \frac{\sigma_c^2}{2} + \rho_{cs}\sigma_s\sigma_c k_c\right)}{k_c^2} - \frac{\sigma_c^2 H_c^2(\tau)}{4k_c}\right)$.

Expression (125) is identical to the solution for the two-factor pricing model (expression (94)). This should make sense. We stated earlier that the trading practice of a forward contract makes its price invariant to the interest rate process. Therefore in the face of a variable interest rate, we model the forward price by taking the interest rate out of the model up front. Since the forward contract is modeled independent of the interest rate, the solution we derive is identical to the two-factor model, as it should be. We now, turn our focus to modeling the futures price given a stochastic interest rate..

3.32 Futures Prices

For the three-factor futures price we use the joint stochastic process for the forward contract above. This is

$$dS(t) = (r(t) - \delta(t))S(t)dt + \sigma_s S(t)dZ_s^*(t), \tag{126}$$

$$d\delta(t) = \left(k_c(\alpha - \delta(t)) - \lambda\sigma_c\right)dt + \sigma_c dZ_c^*(t), \tag{127}$$

$$dr(t) = \left(f_t(s,t) + k_r f(s,t) + \frac{\sigma_r^2}{2k_r}\left(1 - e^{-2k_r(t-s)}\right) - k_r r(t) \right) dt + \sigma_r dZ_r^*(t). \quad (128)$$

The Feynman-Kac solution for the futures price is

$$F(S(t),\delta(t),T) = E_t^*(S(T)) = E_t^*(e^{G(T)}). \quad (129)$$

Expression (129) is different from (113). A futures contract and unlike the forward contract it is *marked-to-market.* Therefore, random movements in the interest rate will impact the futures price so we do not amend the spot price process.

In order to find a solution for the terminal spot price let $G(t) = \ln S(t)$. The diffusion for $G(t)$ follows from the transformation, Ito's lemma, and the risk-adjusted dynamics given for the spot price in equation (126) and is

$$dG(t) = G_s dS(t) + \frac{1}{2} G_{ss} [dS(t)]^2,$$

$$dG(t) = \frac{1}{S(t)}\left[(r(t) - \delta(t))S(t)dt + \sigma_s S(t) dZ_s^*(t)\right] - \frac{1}{2}\left(\frac{1}{[S(t)]^2}\right)[S(t)]^2 \sigma_s^2 dt,$$

$$dG(t) = \left[r(t) - \delta(t) - \frac{1}{2}\sigma_s^2 \right] dt + \sigma_s dZ_s^*(t). \quad (130)$$

Now integrate over equation (130)

$$\int_t^T dG(v) = \int_t^T r(v)dv - \int_t^T \delta(v)dv - \frac{1}{2}\sigma_s^2 \int_t^T dv + \sigma_s \int_t^T dZ_s^*(v),$$

$$G(T) - G(t) = \int_t^T r(v)dv - \int_t^T \delta(v)dv - \frac{1}{2}\sigma_s^2 \int_t^T dv + \sigma_s \int_t^T dZ_s^*(v),$$

$$G(T) = \ln S(t) - \frac{1}{2}\sigma_s^2 (T - t) - \int_t^T \delta(v)dv + \int_t^T r(v)dv + \sigma_s \int_t^T dZ_s^*(v). \quad (131)$$

Similar to the forward price model, the distribution of $G(T)$ is normal, thereby giving the solution for the futures price as

$$F(S(T),\delta(T),r(T),T) = e^{E_t^*[G(T)]+\frac{1}{2}V_t^*[G(T)]} \tag{132}$$

The solution for $F(S(T),\delta(T),r(T),T)$ involves the expected value and variance of $G(T)$. The expected value of $G(T)$ is

$$E_t^*[G(T)] = \ln S(t) - \frac{1}{2}\sigma_s^2(T-t) - \int_t^T E_t^*[\delta(v)]dv + \int_t^T E_t^*[r(v)dv]. \tag{133}$$

To compute equation (133), solutions for the risk-neutralized convenience yield and spot interest rate processes are needed. From expression (118), we have the expression for the convenience yield. This is

$$\delta(v) = e^{-k_c(v-t)}\delta(t) + \left(\alpha - \frac{\lambda\sigma_c}{k_c}\right) - \left(\alpha - \frac{\lambda\sigma_c}{k_c}\right)e^{-k_c(v-t)} + \sigma_c e^{-k_c v}\int_t^v e^{k_c s} dZ_c^*(s). \tag{134}$$

Thus, we only need to find a solution for the risk-neutral interest rate.

The arbitrage free dynamics of the spot interest rate in expression (128) is a special case of the Heath, Jarrow and Morton (1992) model.[19] The Heath, Jarrow and Morton (HJM) model is an arbitrage free model of the term structure of interest rates. To develop their model, HJM work directly with the forward rates. From this model, the arbitrage-free dynamics for the instantaneous spot rate can be determined. Therefore, we need to understand how to model the forward rate dynamics.

To obtain an arbitrage free process for the instantaneous forward rate we consider the relationship between the forward rates and a default free zero coupon bond. Once the relationship between the pure discount bond price and forward rates is found, the risk neutral dynamics for the default free bonds can be used to determine the process for the forward rates. Given the forward rate process, we may then find the spot rate since the spot rate is simply the nearest forward rate. Let us proceed with the model of the instantaneous forward rate.

The relationship between the default free discount pure bond prices $P(t,T_i)$, $T_i < T^{\max}$, with maturity T_i and forward rates $f(t,T)$ is

$$P(t,T) = e^{-\int_t^T f(t,u)du} . \tag{135}$$

Note that there is no expectation operator involved in this expression, because the $f(t,u)$ are all forward rates observed at time t. They are rates on forward loans that will begin at future dates $u > t$ and last an infinitesimal period du. Assume that for a typical bond with maturity T we are given the following stochastic differential equation

$$dP(t,T) = \mu(t,T,P)P(t,T)dt + \sigma(t,T,P)P(t,T)dV(t), \tag{136}$$

where $dV(t)$ is a Wiener increment under the true probability measure. Now, bonds are traded assets. Thus they have an expected return and volatility measure that may be estimated. In addition, the prices must adhere to a no arbitrage rule and we may define the risk premium associated with these stocks. But most importantly, since bond prices are financial instruments that are arbitrage free there exists an equivalent probability measure for these fixed income securities. Therefore, in a risk neutral world with application of the Girsanov theorem, the drift coefficient can be modified as in the case of the Black-Scholes framework

$$dP(t,T) = r(t)P(t,T)dt + \sigma(t,T,P)P(t,T)dZ_r^*(t), \tag{137}$$

where $r(t)$ is the risk-free instantaneous spot rate, and $dZ_r^*(t)$ is the new incremental Weiner process under the risk-neutral probabilities. Note, the unknown drift in the bond dynamics is eliminated.

Given the stochastic differential equations for the bonds we can get the arbitrage free dynamics for the forward rates. Begin with

$$f(t,T,T+\Delta) = \frac{\log P(t,T) - \log P(t,T+\Delta)}{(T+\Delta)-T}, \tag{138}$$

where a noninfinitesimal interval $0 < \Delta$ is used to define the non-instantaneous forward rates, $f(t,T,T+\Delta)$, for a loan that begins at T and ends at $T+\Delta$. This is done by considering two bonds that are identical in all aspects, except for their maturity, which are Δ apart.

Now, to get the arbitrage free dynamics of forward rates, apply Ito's lemma to the right hand side of (138), and use the risk-adjusted drifts whenever needed. Working on the first expression in the numerator of (138) yields

$$d[\log P(t,T)] = \frac{1}{P(t,T)} dP(t,T) - \frac{1}{2P^2(t,T)} dP^2(t,T) \tag{139}$$

Substituting in for the risk-adjusted dynamics of $P(t,T)$ into (139)

$$d[\log P(t,T)] = \frac{1}{P(t,T)} \left(r(t)P(t,T)dt + \sigma(t,T,P)P(t,T)dZ_r^*(t)\right)$$

$$-\frac{1}{2P^2(t,T)}\left(\sigma^2(t,T,P)P^2(t,T)\right)dt.$$

Simplifying the expression above yields

$$d[\log P(t,T)] = \left(r(t) - \frac{1}{2}\sigma^2(t,T,P)\right)dt + \sigma(t,T,P)dZ_r^*(t). \tag{140}$$

Now working on the second term in the numerator of (138)

$$d[\log P(t,T+\Delta)] = \frac{1}{P(t,T+\Delta)} dP(t,T+\Delta) - \frac{1}{2P^2(t,T+\Delta)} dP^2(t,T+\Delta). \quad (141)$$

Substituting in for the risk-adjusted dynamics of $P(t,T+\Delta)$ into (141) yields

$$d[\log P(t,T+\Delta)] = \frac{1}{P(t,T+\Delta)}\left(r(t)P(t,T+\Delta)dt + \sigma(t,T+\Delta,P)P(t,T+\Delta)dZ_r^*(t)\right)$$

$$-\frac{1}{2P^2(t,T+\Delta)}\left(\sigma^2(t,T+\Delta,P)P^2(t,T+\Delta)\right)dt.$$

Simplifying yields

$$d[\log P(t,T+\Delta)] = \left(r(t) - \frac{1}{2}\sigma^2(t,T+\Delta,P)\right)dt + \sigma(t,T,P)dZ_r^*(t). \quad (142)$$

It is important to realize that the first terms in the drift of the Stochastic differentials for $P(t,T)$ and $P(t,T+\Delta)$ are the same because the dynamics under consideration are arbitrage-free. Under the risk-neutral probabilities, discount bonds with different maturities will have expected rates of returns equal to the risk free rate r. This is essentially the same argument used in switching to the known (constant) risk free rate r in the drift of the stock price process utilized in the Black-Scholes derivation.

Now substitute the expressions (140) and (142) into the stochastic differential of expression (138), and cancel the drift terms, $r(t)$, to obtain

$$df(t,T,T+\Delta) = \frac{1}{2\Delta}\left(\sigma^2(t,T+\Delta,P(t,T+\Delta)) - \sigma^2(t,T,P(t,T))\right)dt$$

$$\frac{1}{\Delta}\left(\sigma(t,T+\Delta,P(t,T+\Delta)) - \sigma(t,T,P(t,T))\right)dZ_r^*(t). \quad (143)$$

This is the final result of applying Ito's lemma to (138). Expression (143) is the arbitrage free dynamics of a forward rate on a loan that begins at time T and ends Δ period later.

Now, if we can let $\Delta \to 0$ we will obtain the dynamics of the instantaneous forward rate. To do this, note that the way expression (143) is written. On the right hand side, we have two terms that are of the form

$$\frac{g(x+\Delta)-g(x)}{\Delta}.$$

In expressions like these, letting $\Delta \to 0$ means taking the standard derivative of $g(x)$ with respect to x, Writing these terms in brackets separately and then letting $\Delta \to 0$ amounts to taking the derivative of the two terms on the right hand side with respect to T. Doing this gives

$$\lim_{\Delta\to 0}\frac{1}{2\Delta}\left(\sigma^2(t,T+\Delta,P(t,T+\Delta))-\sigma^2(t,T,P(t,T))\right)$$

$$=\sigma(t,T,P(t,T))\left[\frac{\partial\sigma(t,T,P(t,T))}{\partial T}\right],$$

$$\lim_{\Delta\to 0}\frac{1}{\Delta}\left(\sigma(t,T+\Delta,P(t,T+\Delta))-\sigma(t,T,P(t,T))\right)$$

$$=\left[\frac{\partial\sigma(t,T,P(t,T))}{\partial T}\right].$$

Putting these together in expression (143) we get the corresponding stochastic differential equation for the instantaneous forward rate

$$\lim_{\Delta\to 0} df(t,T,T+\Delta)=df(t,T),$$

or

$$df(t,T)=\sigma(t,T,P(t,T))\left[\frac{\partial\sigma(t,T,P(t,T))}{\partial T}\right]dt+\left[\frac{\partial\sigma(t,T,P(t,T))}{\partial T}\right]dZ_r^*(s). \quad (144)$$

Expression (144) is the HJM risk neutral dynamics for the instantaneous forward rate that is arbitrage free. To reach this result the relationship between a bond price and forward rates is obtained using an arbitrage argument. Then the arbitrage free dynamics are written for $P(t,T)$. Given the Stochastic differentials for the bond prices, we derive the dynamics for the instantaneous forward rate. This process (expression (144)) is arbitrage free, and the risk neutral drift for $f(t,T)$ is

$$\mu_f = \sigma(t,T,P(t,T))\left[\frac{\partial\sigma(t,T,P(t,T))}{\partial T}\right].$$

The instantaneous diffusion coefficient is

$$\sigma_f = \frac{\partial\sigma(t,T,P(t,T))}{\partial T}.$$

The expression $\frac{\partial\sigma(t,T,P(t,T))}{\partial T}$ is the volatility for the forward rate $f(t,T)$, which is given above. Analytically the forward rate volatility is equal to the partial derivative of time t's pure discount bond volatility with respect to maturity. Intuitively, this means the forward rate's volatility is a portion of the total volatility for $P(t,T)$. As the notation suggest, the expression $\sigma(t,T,P(t,T))$ is the volatility of a pure discount bond at time t maturing at time T. Furthermore, we may rewrite this term. Recall from expression (135) that the price of pure discount bond is

$$P(t,T) = e^{-\int_t^T f(t,u)du}.$$

Since the price of a pure discount bond is a function of the forward rates, its variability is caused by these forward rates. Thus the volatility for the pure discount bond may be written as

$$\sigma(t,T,P(t,T)) = \int_t^T \frac{\partial\sigma(t,u,P(t,u))}{\partial u} du.$$

We have an expression for the risk-neutral forward rate process and now turn our attention to the spot rate. The question is what does the above dynamics for the forward rates imply for the spot rate? The question is relevant because the spot rate corresponds to the nearest infinitesimal forward loan, which formally stated is

$$r(t) = f(t,t)$$

for all t. The expression for the spot rate will extend from the stochastic process for the instantaneous forward rate. Working with expression (144), we take the integral for $f(t,T)$ to obtain

$$f(t,T) = f(0,T) + \int_0^t \frac{\partial\sigma(s,T,P(s,T))}{\partial T}\left[\int_s^T \frac{\partial\sigma(s,u,P(s,u))}{\partial u}du\right]ds + \int_0^t \frac{\partial\sigma(s,T,P(s,T))}{\partial T}dZ_r^*(s).$$

Next, select $T = t$ to get a representation for the spot rate $r(t)$

$$r(t) = f(0,t) + \int_0^t \frac{\partial\sigma(s,t,P(s,t))}{\partial t}\left[\int_s^t \frac{\partial\sigma(s,u,P(s,u))}{\partial u}du\right]ds + \int_0^t \frac{\partial\sigma(s,t,P(s,t))}{\partial t}dZ_r^*(s).$$

(145)

Expression (145) is the general expression for the risk neutral spot rate for the HJM model. We can consider a subset of the HJM processes for the instantaneous forward rate. Let us presume that the volatility of the forward rates is an exponentially dampened volatility structure. That is,

$$\sigma_f(t,T) = \sigma_r e^{-k_r(T-t)}.$$

This representation exploits the fact that the near term forward rates are more volatile than distant forward rates. Implicit in the model is that the same Wiener process impacts all forwards. This seems incongruent with certain theories for the terms structure, but the volatility process stated above allows for difference in the volatilities of different forward rates. Given the volatility above we may write the forward rate dynamics in equation (144) as

$$df(t,T) = \left(\sigma_r e^{-k_r(v-t)} \int_t^v \sigma_r e^{-k_r(s-t)} ds \right) dt + \sigma_r e^{-k_r(v-t)} dZ_r^*(t),$$

or

$$df(t,v) = \frac{-\sigma_r^2}{k_r}\left(e^{-2k_r(v-t)} - e^{-k_r(v-t)}\right)dt + \sigma_r e^{-k_r(v-t)} dZ_r^*(t). \tag{146}$$

To obtain the solution for the spot rate process we may directly use equation (145) or follow the derivation above.[20] Following the derivation above to obtain the spot rate process, we integrate over expression (146)

$$\int_t^v df(s,v)ds = \int_t^v \frac{-\sigma_r^2}{k_r}\left(e^{-2k_r(v-s)} - e^{-k_r(v-s)}\right)ds + \int_t^v \sigma_r e^{-k_r(v-s)} dZ_r^*(v),$$

$$f(v,v) - f(t,v) = \frac{\sigma_r^2}{k_r^2}\left(e^{-k_r(v-v)} - e^{-k_r(v-t)}\right) - \frac{\sigma_r^2}{2k_r^2}\left(e^{-2k_r(v-v)} - e^{-2k_r(v-t)}\right)$$

$$+ \int_t^v \sigma_r e^{-k_r(v-s)} dZ_r^*(s),$$

$$r(v) = f(t,v) + \frac{\sigma_r^2}{k_r^2}\left(1 - e^{-k_r(v-t)}\right) - \frac{\sigma_r^2}{2k_r^2}\left(1 - e^{-2k_r(v-t)}\right) + \int_t^v \sigma_r e^{-k_r(v-s)} dZ_r^*(s),$$

$$= f(t,v) + \frac{\sigma_r^2}{k_r^2} - \frac{\sigma_r^2}{k_r^2} e^{-k_r(v-t)} - \frac{\sigma_r^2}{2k_r^2} + \frac{\sigma_r^2}{2k_r^2} e^{-2k_r(v-t)} + \int_t^v \sigma_r e^{-k_r(v-s)} dZ_r^*(s),$$

$$= f(t,v) + \frac{\sigma_r^2}{2k_r^2}\left(2 - 2e^{-k_r(v-t)} - 1 + e^{-2k_r(v-t)}\right) + \sigma_r e^{-k_r v} \int_t^v e^{k_r s} dZ_r^*(s),$$

$$= f(t,v) + \frac{\sigma_r^2}{2k_r^2}\left(1 - 2e^{-k_r(v-t)} + e^{-2k_r(v-t)}\right) + \sigma_r e^{-k_r v} \int_t^v e^{k_r s} dZ_r^*(s). \tag{147}$$

The solution for the instantaneous spot rate, expression (147), is an artifact of the presumed dynamics of the forward rate, and the stochastic differential for the spot rate is stated in equation (128). To understand the solution in equation (147) we need to understand equation (128).[21] Equation (128) is a specific form of a linear stochastic differential equation with a time varying coefficient. That is, equation (128) describes the dynamics for the stochastic differential $dr(t)$ around its mean, where the mean itself is time varying. Movements in the mean of the instantaneous spot rate are due to shifts in the term structure. As the term structure shifts over time, it pulls the mean of the spot rate with it. We see from the HJM model that these movements in the spot rate mean are not unexpected. Expression (144) shows that once the dynamics of the term structure are known the instantaneous drift is known too. While the expected movements in the mean spot rate are known, the actual movements in the spot rates are random. The process in (128) simply describes how the spot rate oscillates around its time varying mean.

We have solutions for the convenience yield and spot interest rate in expressions (134) and (147), and we may now solve the integrals in equation (133). We start with the convenience yield. From expression (118), we know that

$$\int_t^T E_t^*[\delta(v)]dv = \left(\alpha - \frac{\lambda\sigma_c}{k_c}\right)\tau - \left(\delta(t) - \alpha + \frac{\lambda\sigma_c}{k_c}\right)(H_c(\tau)).$$

Working on the spot interest rate we have

$$\int_t^T E_t^*[r(v)]dv = \int_t^T\left(f(t,v) + \frac{\sigma_r^2}{2k_r^2}\left(1 - 2e^{-k_r(v-t)} + e^{-2k_r(v-t)}\right)\right)dv,$$

$$= \int_t^T f(t,v)dv + \int_t^T \frac{\sigma_r^2}{2k_r^2}\left(1 - 2e^{-k_r(v-t)} + e^{-2k_r(v-t)}\right)dv,$$

$$= \int_t^T f(t,v)dv + \frac{\sigma_r^2}{2k_r^2}(T-t) - \frac{2\sigma_r^2}{2k_r^2}\int_t^T e^{-k_r(v-t)}dv + \frac{\sigma_r^2}{2k_r^2}\int_t^T e^{-2k_r(v-t)}dv,$$

$$= \int_t^T f(t,v)dv + \frac{\sigma_r^2}{2k_r^2}\tau + \frac{2\sigma_r^2}{2k_r^3}\left[e^{-k_r(T-t)} - 1\right] - \frac{\sigma_r^2}{4k_r^3}\left[e^{-2k_r(T-t)} - 1\right],$$

$$= \int_t^T f(t,v)dv + \frac{\sigma_r^2}{2k_r^2}\tau - \frac{2\sigma_r^2}{2k_r^3}\left[1 - e^{-k_r\tau}\right] + \frac{\sigma_r^2}{4k_r^3}\left[1 - e^{-2k_r\tau}\right],$$

$$= \int_t^T f(t,v)dv + \frac{\sigma_r^2}{2k_r^2}\tau - \frac{\sigma_r^2}{2k_r^3}\left[2 - 2e^{-k_r\tau}\right] + \frac{\sigma_r^2}{4k_r^3}\left[1 - e^{-2k_r\tau}\right],$$

$$= \int_t^T f(t,v)dv + \frac{\sigma_r^2}{2k_r^2}\tau - \frac{\sigma_r^2}{2k_r^3}\left(1 - e^{-k_r\tau}\right) - \frac{\sigma_r^2}{2k_r^3}\left(1 - e^{-k_r\tau}\right) + \frac{\sigma_r^2}{4k_r^3}\left(1 - e^{-2k_r\tau}\right),$$

$$= \int_t^T f(t,v)dv + \frac{\sigma_r^2}{2k_r^2}\tau - \frac{\sigma_r^2}{2k_r^3}\left(1 - e^{-k_r\tau}\right) + \frac{\sigma_r^2}{4k_r^3}\left(1 - e^{-2k_r\tau} + 2e^{-k_r\tau} - 2\right),$$

$$= \int_t^T f(t,v)dv + \frac{\sigma_r^2}{2k_r^2}\tau - \frac{\sigma_r^2}{2k_r^3}\left(1 - e^{-k_r\tau}\right) - \frac{\sigma_r^2}{4k_r^3}\left(1 - e^{-k_r\tau}\right)^2,$$

$$= \int_t^T f(t,v)dv + \frac{\sigma_r^2}{2k_r^2}\tau - \frac{\sigma_r^2 H_r(\tau)}{2k_r^2} - \frac{\sigma_r^2 H_r^2(\tau)}{4k_r},$$

$$= \int_t^T f(t,v)dv - \left(H_r(\tau) - \tau\right)\frac{\sigma_r^2}{2k_r^2} - \frac{\sigma_r^2 H_r^2(\tau)}{4k_r}.$$

Now substitute the results above into (133) we obtain

$$E_t^*[G(T)] = \ln S(t) - \frac{1}{2}\sigma_s^2(T-t) - \left(\alpha - \frac{\lambda\sigma_c}{k_c}\right)\tau - \left(\delta(t) - \alpha + \frac{\lambda\sigma_c}{k_c}\right)(H_c(\tau))$$

$$+ \int_t^T f(t,v)dv - \left(H_r(\tau) - \tau\right)\frac{\sigma_r^2}{2k_r^2} - \frac{\sigma_r^2 H_r^2(\tau)}{4k_r}. \qquad (148)$$

The variance of $G(T)$ is

$$V_t^*[G(T)] = V_t^*\left[\int_t^T \delta(v)dv\right] + V_t^*\left[\sigma_s\int_t^T dZ_s^*(v)\right] + V_t^*\left[\int_t^T r(v)dv\right] - 2Cov_t^*\left[\int_t^T \delta(v)dv, \sigma_s\int_t^T dZ_s^*(v)\right]$$

$$+2Cov_t^*\left[\int_t^T r(v)dv, \sigma_s \int_t^T dZ_s^*(v)\right] - 2Cov_t^*\left[\int_t^T \delta(v)dv, \int_t^T r(v)dv\right]. \quad (149)$$

We know from expressions (121), (122) and (123) in the three-factor forward model that

$$V_t^*\left[\sigma_s \int_t^T dZ_s^*(v)\right] = \sigma_s^2 \tau, \quad (150)$$

$$\int_t^T V_t^*[\delta(v)] = -(H_c(\tau) - \tau)\frac{\sigma_c^2}{k_c^2} - \frac{\sigma_c^2 H_c^2(\tau)}{2k_c}, \quad (151)$$

and

$$Cov_t^*\left[\int_t^T \delta(v)dv, \sigma_s \int_t^T dZ_s^*(v)\right] = \frac{\sigma_s \sigma_c \rho_{cs}}{k_c}(\tau - H_c(\tau)). \quad (152)$$

Therefore, we need to find expressions for the remaining terms. Starting with the variance for the interest rate.

$$V_t^*\left[\int_t^T r(v)dv\right] = V_t^*\left[\int_t^T \sigma_r e^{-k_r v} \int_t^v e^{k_r s} dZ_r^*(s) dv\right].$$

Switching the order of integration

$$= V_t^*\left[\int_t^T e^{k_r s} \int_s^T \sigma_r e^{-k_r v} dv dZ_r^*(s)\right],$$

$$= V_t^*\left[\int_t^T e^{k_r s}\left[-\frac{\sigma_r}{k_r}\left(e^{-k_r T} - e^{-k_r s}\right)\right] dZ_r^*(s)\right],$$

$$= V_t^*\left[\int_t^T e^{k_r s}\left[\frac{\sigma_r}{k_r}\left(e^{-k_r s} - e^{-k_r T}\right)\right] dZ_r^*(s)\right],$$

$$= V_t^*\left[\int_t^T \frac{\sigma_r}{k_r}\left(1 - e^{-k_r (T-s)}\right) dZ_r^*(s)\right].$$

Using Ito's isometry we may write the above expression as

$$=\frac{\sigma_r^2}{k_r^2}\int_t^T\left(1-e^{-k_r(T-s)}\right)^2 ds,$$

$$=\frac{\sigma_r^2}{k_r^2}\int_t^T\left(1-2e^{-k_r(T-s)}+e^{-2k_r(T-s)}\right)ds,$$

$$=\frac{\sigma_r^2}{k_r^2}(T-t)-\frac{2\sigma_r^2}{k_r^2}\left(\frac{1}{k_r}\right)\left(e^{-k_r(T-T)}-e^{-k_r(T-t)}\right)+\frac{\sigma_r^2}{k_r^2}\left(\frac{1}{2k_r}\right)\left(e^{-2k_r(T-T)}-e^{-2k_r(T-t)}\right),$$

$$=\frac{\sigma_r^2}{k_r^2}(T-t)-\frac{2\sigma_r^2}{k_r^2}\left(\frac{1}{k_r}\right)\left(1-e^{-k_r(T-t)}\right)+\frac{\sigma_r^2}{k_r^2}\left(\frac{1}{2k_r}\right)\left(1-e^{-2k_r(T-t)}\right),$$

$$=\frac{\sigma_r^2}{k_r^2}\tau-\frac{2\sigma_r^2}{k_r^2}\left(\frac{1}{k_r}\right)\left(1-e^{-k_r\tau}\right)+\frac{\sigma_r^2}{k_r^2}\left(\frac{1}{2k_r}\right)\left(1-e^{-2k_r\tau}\right),$$

$$=\frac{\sigma_r^2}{k_r^2}\tau-\frac{\sigma_r^2}{k_r^2}\left(\frac{1}{k_r}\right)\left(2-2e^{-k_r\tau}\right)+\frac{\sigma_r^2}{k_r^2}\left(\frac{1}{2k_r}\right)\left(1-e^{-2k_r\tau}\right),$$

$$=\frac{\sigma_r^2}{k_r^2}\tau-\frac{\sigma_r^2}{k_r^3}\left(1-e^{-k_r\tau}\right)-\frac{\sigma_r^2}{k_r^3}\left(1-e^{-k_r\tau}\right)+\frac{\sigma_r^2}{2k_r^3}\left(1-e^{-2k_r\tau}\right),$$

$$=\frac{\sigma_r^2}{k_r^2}\tau-\frac{\sigma_r^2}{k_r^3}\left(1-e^{-k_r\tau}\right)+\frac{\sigma_r^2}{2k_r^3}\left(1-e^{-2k_r\tau}+2e^{-k_r\tau}-2\right),$$

$$=\frac{\sigma_r^2}{k_r^2}\tau-\frac{\sigma_r^2}{k_r^3}\left(1-e^{-k_r\tau}\right)+\frac{\sigma_r^2}{2k_r^3}\left(-1-e^{-2k_r\tau}+2e^{-k_r\tau}\right),$$

$$=\frac{\sigma_r^2}{k_r^2}\tau-\frac{\sigma_r^2}{k_r^3}\left(1-e^{-k_r\tau}\right)-\frac{\sigma_r^2}{2k_r^3}\left(1-2e^{-k_r\tau}+e^{-2k_r\tau}\right),$$

$$=\frac{\sigma_r^2}{k_r^2}\tau-\frac{\sigma_r^2}{k_r^3}\left(1-e^{-k_r\tau}\right)-\frac{\sigma_r^2}{2k_r^3}\left(1-e^{-k_r\tau}\right)^2,$$

$$=\frac{\sigma_r^2}{k_r^2}\tau-\frac{\sigma_r^2H_r(\tau)}{k_r^2}-\frac{\sigma_r^2H_r^2(\tau)}{2k_r},$$

$$=-\left(H_r(\tau)-\tau\right)\frac{\sigma_r^2}{k_r^2}-\frac{\sigma_r^2H_r^2(\tau)}{2k_r}. \qquad (153)$$

The covariance between the spot interest rate and the spot price is

$$Cov_t^*\left[\int_t^T r(v)dv, \sigma_s \int_t^T dZ_s^*(v)\right] = Cov_t^*\left[\int_t^T \sigma_r e^{-k_r v}\int_t^v e^{k_r s} dZ_r^*(s)dv, \sigma_s \int_t^T dZ_s^*(s)\right],$$

$$= Cov_t^*\left[\int_t^T e^{k_r s}\int_s^T \sigma_r e^{-k_r v} dv dZ_r^*(s), \sigma_s \int_t^T dZ_s^*(s)\right],$$

$$= Cov_t^*\left[\int_t^T e^{k_r s}\left[-\frac{\sigma_r}{k_r}\left(e^{-k_r T} - e^{-k_r s}\right)\right]dZ_r^*(s), \sigma_s \int_t^T dZ_s^*(s)\right],$$

$$= Cov_t^*\left[\frac{\sigma_r}{k_r}\int_t^T \left(1 - e^{-k_r(T-s)}\right)dZ_r^*(s), \sigma_s \int_t^T dZ_s^*(s)\right].$$

Using Ito's isometry yields

$$= \frac{\sigma_s \sigma_r \rho_{rs}}{k_r}\int_t^T \left(1 - e^{-k_r(T-s)}\right)ds,$$

$$= \frac{\sigma_s \sigma_r \rho_{rs}}{k_r}\left((T-t) - \left(\frac{1}{k_r}\right)\left(e^{-k_r(T-T)} - e^{-k_r(T-t)}\right)\right),$$

$$= \frac{\sigma_s \sigma_r \rho_{rs}}{k_r}\left((T-t) - \left(\frac{1}{k_r}\right)\left(1 - e^{-k_r(T-t)}\right)\right),$$

$$= \frac{\sigma_s \sigma_r \rho_{rs}}{k_r}\left(\tau - H_r(\tau)\right). \tag{154}$$

The covariance between the convenience yield and spot rate is

$$Cov_t^*\left[\int_t^T \delta(v)dv, \int_t^T r(v)dv\right] = Cov_t^*\left[\int_t^T \sigma_c e^{-k_c v}\int_t^v e^{k_c s} dZ_c^*(s)dv, \int_t^T \sigma_r e^{-k_r v}\int_t^v e^{k_r s} dZ_r^*(s)dv\right],$$

$$= Cov_t^*\left[\int_t^T e^{k_c s}\int_s^T \sigma_c e^{-k_c v} dv dZ_c^*(s), \int_t^T e^{k_r s}\int_s^T \sigma_r e^{-k_r v} dv dZ_r^*(s)\right],$$

$$= Cov_t^*\left[\int_t^T e^{k_c s}\left[-\frac{\sigma_c}{k_c}\left(e^{-k_c T} - e^{-k_c s}\right)\right]dZ_c^*(s), \int_t^T e^{k_r s}\left[-\frac{\sigma_r}{k_r}\left(e^{-k_r T} - e^{-k_r s}\right)\right]dZ_r^*(s)\right],$$

$$= \frac{\sigma_c \sigma_r}{k_c k_r} Cov_t^*\left[\int_t^T \left(1 - e^{-k_c(T-s)}\right)dZ_c^*(s), \int_t^T \left(1 - e^{-k_r(T-s)}\right)dZ_r^*(s)\right].$$

Using Ito's isometry we obtain

$$= \frac{\sigma_c \sigma_r \rho_{rc}}{k_c k_r}\int_t^T \left(1 - e^{-k_c(T-s)}\right)\left(1 - e^{-k_r(T-s)}\right)ds,$$

$$= \frac{\sigma_c \sigma_r \rho_{rc}}{k_c k_r}\int_t^T \left(1 - e^{-k_r(T-s)} - e^{-k_c(T-s)} + e^{-k_r(T-s)}e^{-k_c(T-s)}\right)ds,$$

$$= \frac{\sigma_c \sigma_r \rho_{rc}}{k_c k_r}\left[(T-t) - \frac{1}{k_r}\left(1 - e^{-k_r(T-t)}\right) - \frac{1}{k_c}\left(1 - e^{-k_c(T-t)}\right) + \frac{1}{k_r + k_c}\left(1 - e^{-k_r(T-t)}e^{-k_c(T-t)}\right)\right],$$

$$= \frac{\sigma_c \sigma_r \rho_{rc}}{k_c k_r}\left[\tau - H_r(\tau) - H_c(\tau) + \frac{1}{k_r + k_c}\left(1 - e^{-k_r \tau}e^{-k_c \tau}\right)\right],$$

$$= \frac{\sigma_c \sigma_r \rho_{rc}}{k_c k_r}\left[\frac{\tau(k_r + k_c)}{k_r + k_c} - \frac{H_r(\tau)(k_r + k_c)}{k_r + k_c} - \frac{H_c(\tau)(k_r + k_c)}{k_r + k_c} + \frac{\left(1 - e^{-k_r \tau}e^{-k_c \tau}\right)}{k_r + k_c}\right],$$

$$= \frac{\sigma_c \sigma_r \rho_{rc}}{k_r + k_c}\left[\frac{\tau(k_r + k_c)}{k_c k_r} - \frac{H_r(\tau)(k_r + k_c)}{k_c k_r} - \frac{H_c(\tau)(k_r + k_c)}{k_c k_r} + \frac{\left(1 - e^{-k_r \tau}e^{-k_c \tau}\right)}{k_c k_r}\right],$$

$$= \frac{-\sigma_c \sigma_r \rho_{rc}}{k_r + k_c}\left[\frac{(H_r(\tau) - \tau)k_c}{k_c k_r} + \frac{(H_c(\tau) - \tau))k_r}{k_c k_r} + \frac{H_r(\tau)k_r}{k_c k_r} + \frac{H_c(\tau)k_c}{k_c k_r} - \frac{\left(1 - e^{-k_r \tau}e^{-k_c \tau}\right)}{k_c k_r}\right],$$

$$= \frac{-\sigma_c \sigma_r \rho_{rc}}{k_r + k_c}\left[\frac{(H_r(\tau) - \tau)}{k_r} + \frac{(H_c(\tau) - \tau))}{k_c} + \frac{\left(1 - e^{-k_r \tau} - e^{-k_c \tau} + e^{-k_r \tau}e^{-k_c \tau}\right)}{k_c k_r}\right],$$

$$= \frac{-\sigma_c \sigma_r \rho_{rc}}{k_r + k_c}\left[\frac{(H_r(\tau) - \tau)}{k_r} + \frac{(H_c(\tau) - \tau))}{k_c} + \frac{\left(1 - e^{-k_c \tau}\right)\left(1 - e^{-k_r \tau}\right)}{k_c k_r}\right],$$

$$= \frac{-\sigma_c \sigma_r \rho_{rc}}{k_r + k_c}\left[\frac{(H_r(\tau) - \tau)}{k_r} + \frac{(H_c(\tau) - \tau))}{k_c} + H_r(\tau)H_c(\tau)\right]. \qquad (155)$$

Now substitute expression (150)-(155) into (149) to obtain the variance

$$V_t^*[G(T)]=\left(-(H_c(\tau)-\tau)\frac{\sigma_c^2}{k_c^2}-\frac{\sigma_c^2 H_c^2(\tau)}{2k_c}\right)+\sigma_s^2\tau+\left(-(H_r(\tau)-\tau)\frac{\sigma_r^2}{k_r^2}-\frac{\sigma_r^2 H_r^2(\tau)}{2k_r}\right)$$

$$-2\left(\frac{\sigma_s\sigma_c\rho_{cs}}{k_c}(\tau-H_c(\tau))\right)+2\left(\frac{\sigma_s\sigma_r\rho_{rs}}{k_r}(\tau-H_r(\tau))\right)$$

$$-2\left(\frac{-\sigma_c\sigma_r\rho_{rc}}{k_r+k_c}\left[\frac{(H_r(\tau)-\tau)}{k_r}+\frac{(H_c(\tau)-\tau))}{k_c}+H_r(\tau)H_c(\tau)\right]\right). \qquad (156)$$

Using the expected value and variance for $G(T)$ in expressions (148) and (166) we can obtain the formula for the three-factor futures price. Substituting (148) and (166) into equation (129) yields

$$F(S(T),\delta(T),r(T),T)=E_t^*(e^{G(T)})=e^{E_t^*[G(T)]+\frac{1}{2}V_t^*[G(T)]},$$

$$F(S(T),\delta(T),r(T),T)=\exp\left\{\ln S(t)-\frac{1}{2}\sigma_s^2(T-t)-\left(\left(\alpha-\frac{\lambda\sigma_c}{k_c}\right)\tau\right.\right.$$

$$-\left(\delta(t)-\alpha+\frac{\lambda\sigma_c}{k_c}\right)H_c(\tau)\Bigg)+\int_t^T f(t,v)dv-(H_r(\tau)-\tau)\frac{\sigma_r^2}{2k_r^2}-\frac{\sigma_r^2 H_r^2(\tau)}{4k_r}$$

$$+\frac{1}{2}\left(\left(-(H_c(\tau)-\tau)\frac{\sigma_c^2}{k_c^2}-\frac{\sigma_c^2 H_c^2(\tau)}{2k_c}\right)+\sigma_s^2\tau\right.$$

$$+\left(-(H_r(\tau)-\tau)\frac{\sigma_r^2}{k_r^2}-\frac{\sigma_r^2 H_r^2(\tau)}{2k_r}\right)-2\left(\frac{\sigma_s\sigma_c\rho_{cs}}{k_c}(\tau-H_c(\tau))\right)$$

$$+2\left(\frac{\sigma_s\sigma_r\rho_{rs}}{k_r}(\tau-H_r(\tau))\right)$$

$$\left.\left.-2\left(\frac{-\sigma_c\sigma_r\rho_{rc}}{k_r+k_c}\left[\frac{(H_r(\tau)-\tau)}{k_r}+\frac{(H_c(\tau)-\tau))}{k_c}+H_r(\tau)H_c(\tau)\right]\right)\right)\right\},$$

$$= \exp\left\{ \ln S(t) - \frac{1}{2}\sigma_s^2(T-t) - \left(\frac{k_c^2\alpha}{k_c^2} - \frac{k_c\lambda\sigma_c}{k_c^2}\right)\tau - \left(\delta(t) - \frac{k_c^2\alpha}{k_c^2} + \frac{k_c\lambda\sigma_c}{k_c^2}\right)(H_c(\tau)) \right.$$

$$+ \int_t^T f(t,v)dv - (H_r(\tau)-\tau)\frac{\sigma_r^2}{2k_r^2} - \frac{\sigma_r^2 H_r^2(\tau)}{4k_r}$$

$$+ \frac{1}{2}\left(\left(-(H_c(\tau)-\tau)\frac{\sigma_c^2}{k_c^2} - \frac{\sigma_c^2 H_c^2(\tau)}{2k_c}\right) + \sigma_s^2\tau\right.$$

$$+ \left(-(H_r(\tau)-\tau)\frac{\sigma_r^2}{k_r^2} - \frac{\sigma_r^2 H_r^2(\tau)}{2k_r}\right) - 2\left(\frac{\sigma_s\sigma_c\rho_{cs}}{k_c}(\tau - H_c(\tau))\right)$$

$$+ 2\left(\frac{\sigma_s\sigma_r\rho_{rs}}{k_r}(\tau - H_r(\tau))\right)$$

$$\left.\left. - 2\left(\frac{-\sigma_c\sigma_r\rho_{rc}}{k_r+k_c}\left[\frac{(H_r(\tau)-\tau)}{k_r} + \frac{(H_c(\tau)-\tau))}{k_c} + H_r(\tau)H_c(\tau)\right]\right)\right)\right\},$$

$$= \exp\left\{ \ln S(t) - \frac{1}{2}\sigma_s^2\tau + (H_c(\tau)-\tau)\left(\frac{k_c^2\alpha - k_c\lambda\sigma_c}{k_c^2}\right) - \delta(t)H_c(\tau) \right.$$

$$+ \int_t^T f(t,v)dv - (H_r(\tau)-\tau)\frac{\sigma_r^2}{2k_r^2} - \frac{\sigma_r^2 H_r^2(\tau)}{4k_r}$$

$$-(H_c(\tau)-\tau)\frac{\sigma_c^2}{2k_c^2} - \frac{\sigma_c^2 H_c^2(\tau)}{4k_c} + \frac{1}{2}\sigma_s^2\tau$$

$$-(H_r(\tau)-\tau)\frac{\sigma_r^2}{2k_r^2} - \frac{\sigma_r^2 H_r^2(\tau)}{4k_r} - \frac{\sigma_s\sigma_c\rho_{cs}}{k_c}(\tau - H_c(\tau))$$

$$+ \frac{\sigma_s\sigma_r\rho_{rs}}{k_r}(\tau - H_r(\tau))$$

$$\left. \frac{-\sigma_c\sigma_r\rho_{rc}}{k_r+k_c}\left[\frac{(H_r(\tau)-\tau)}{k_r} + \frac{(H_c(\tau)-\tau))}{k_c} + H_r(\tau)H_c(\tau)\right]\right\},$$

$$= \exp\left\{ \ln S(t) + \left(H_c(\tau) - \tau\right)\left(\frac{k_c^2\alpha - k_c\lambda\sigma - \frac{\sigma_c^2}{2} + k_c\sigma_s\sigma_c\rho_{cs}}{k_c^2} \right) - \frac{\sigma_c^2 H_c^2(\tau)}{4k_c} \right.$$

$$+ \frac{\sigma_c\sigma_r\rho_{rc}}{k_r + k_c}\left[\frac{\left(H_r(\tau) - \tau\right)}{k_r} + \frac{\left(H_c(\tau) - \tau\right)}{k_c} + H_r(\tau)H_c(\tau) \right]$$

$$-\left(H_r(\tau) - \tau\right)\frac{\sigma_r^2}{k_r^2} - \frac{\sigma_r^2 H_r^2(\tau)}{2k_r}$$

$$\left. + \left(\frac{\sigma_s\sigma_r\rho_{rs}}{k_r}\left(\tau - H_r(\tau)\right) \right) \right\} \exp\{-\delta(t)H_c(\tau)\}\exp\left\{ + \int_t^T f(t,v)dv \right\},$$

$$= S(t)\exp\left\{ \frac{\left(H_c(\tau) - \tau\right)\left(k_c^2\alpha - k_c\lambda\sigma_c - \frac{\sigma_c^2}{2} + \rho_{cs}\sigma_s\sigma_c k_c \right)}{k_c^2} - \frac{\sigma_c^2 H_c^2(\tau)}{4k_c} \right.$$

$$+ \frac{\sigma_c\sigma_r\rho_{rc}}{k_r + k_c}\left[\frac{\left(H_r(\tau) - \tau\right)}{k_r} + \frac{\left(H_c(\tau) - \tau\right)}{k_c} + H_r(\tau)H_c(\tau) \right]$$

$$\left. + \left(\frac{\sigma_s\sigma_r\rho_{rs}}{k_r}\left(\tau - H_r(\tau)\right) \right) - \left(H_r(\tau) - \tau\right)\frac{\sigma_r^2}{k_r^2} - \frac{\sigma_r^2 H_r^2(\tau)}{2k_r} \right\} \exp\{-H_c(\tau)\delta(t)\}\frac{1}{P(t,T)}.$$

Simplifying the above, we have

$$F(S(T),\delta(T),r(T),T) = S(t)A(\tau)D_1(\tau)D_2(\tau)D_3(\tau)e^{-H_c(\tau)\delta(t)}\frac{1}{P(t,T)}, \tag{157}$$

where $$A(\tau) = \exp\left(\frac{\left(H_c(\tau) - \tau\right)\left(k_c^2\alpha - k_c\lambda\sigma_c - \frac{\sigma_c^2}{2} + \rho_{cs}\sigma_s\sigma_c k_c \right)}{k_c^2} - \frac{\sigma_c^2 H_c^2(\tau)}{4k_c} \right), \tag{158}$$

$$D_1(\tau) = \exp\left[\frac{\sigma_c\sigma_r\rho_{rc}}{k_r + k_c}\left[\frac{\left(H_r(\tau) - \tau\right)}{k_r} + \frac{\left(H_c(\tau) - \tau\right)}{k_c} + H_r(\tau)H_c(\tau) \right] \right], \tag{159}$$

$$D_2(\tau) = \exp\left[\left(\frac{\sigma_s\sigma_r\rho_{rs}}{k_r}(\tau - H_r(\tau))\right)\right], \qquad (160)$$

$$D_3(\tau) = \exp\left[-(H_r(\tau) - \tau)\frac{\sigma_r^2}{k_r^2} - \frac{\sigma_r^2 H_r^2(\tau)}{2k_r}\right]. \qquad (161)$$

The closed form solution for the three-factor futures pricing model inculcates a large amount of information. First, we see it is a function of the current spot price, current convenience yield, the current price of the pure discount bond maturing at time *T*. In addition the risk neutral drift and diffusion terms for the each state variable is compounded into the futures price. The solution is more general than the one and two factor solutions and we see that the three-factor futures equation can reduce to the two-factor futures/forward equation. That is, when $D_1(\tau) = D_2(\tau) = D_3(\tau) = 1$ the three factor solution is the same as the two-factor solution.

What do the coefficients $D_1(\tau), D_2(\tau), D_3(\tau)$ capture. The term $D_1(\tau)$ is a premium or a discount depending on the correlation between the interest rate and the convenience yield. If the two processes seem to move in tandem with one another then $D_1(\tau)$ will be positive, putting a premium on the futures price. The term $D_2(\tau)$ is a premium or the discount due to the correlation between the spot price and interest rates. This will work in the same as $D_1(\tau)$.

Expression (161), $D_3(\tau)$, introduces the premium or discount applied to the futures price due to the volatility of the interest rate. The interest rate volatility is an exponentially dampened function. For very short time periods this expression is greater than one, thereby adding a premium to the futures price. As the time to maturity increases this expression is decreasing and in particular is less than one. The indication is that as time to maturity goes up greater discounts are applied to the futures price due to decreased volatility..

Endnotes

[1] Economists have extensively studied the impact of a convenience yield upon commodity markets. Some of the most noticeable work in this area is Working (1948), Brennan (1958) and Telser (1958).

[2] A more detailed discussion of equivalent martingales measures is found in Appendix C.

[3] For the conditional expectation method of solution we do not construct an arbitrage free portfolio directly. We presume that we may alter the probability distribution for a random process. This altered process becomes a martingale and Harrison and Kreps (1979) and Harrison and Pliska (1979) show that if assets follow a martingale process then no arbitrage opportunities exist with this asset.

[4] For the model to be consistent with the general efficient market hypothesis of Fama (1970) and Samuelsone (1965b), the dynamics of the unanticipated part of the asset price motions should be a martingale. That is, the innovation term should follow a random walk. These price vibrations for example, are due to a temporary imbalance between supply and demand, changes in the capitalization rates, changes in economic outlook, or other new information that causes marginal changes in the asset's value. A good candidate to model this behavior is geometric Brownian motion. The properties of this process in an economic context are discussed in Cootner (1964), Samuelson (1965a, 1973), Merton (1971, 1973a, 1973b), and Merton and Samuelson (1974).

[5] This is equivalent to assuming in a general equilibrium framework that the representative investor has a logarithmic utility function. In this special case, the marginal utility of wealth is independent of wealth. Consequently, the market price of risk which is the covariance between the change in the spot price with the rate of change

in the marginal utility is constant. Under the above assumption, the derivation for the market price of risk in a general equilibrium framework is done by Cox et al (1985a, 1985b). Gibson and Schwartz (1990) make the same assumption for the representative agent.

[6] The astute reader will notice that the futures price in equation (2) is not discounted. The futures price is said to equal the expected spot price. This is the price to be paid at the terminal date of the contract. If we wanted to know the futures price today we would have to discount the expected spot price. Thus, if the spot price has a risk premium embedded in it, then discounting the expected spot price by the known riskless rate of return would not yield a Martingale.

[7] Refer to appendix C for a more formal discussion.

[8] A discussion of the Gibson and Schwartz (1990) model is found in Appendix B

[9] A formal discussion of Ito's isometry and stochastic calculus is presented in Appendix D. A more serious reader might find Neftci (2000), Hoel et al (1972), or Kushner (1995) a better presentation.

[10] For the cautious reader the above argument is presented in the analysis as a review for past literature and only serves as a point of departure for later models. This argument assumes that in equilibrium the market always clears. That is supply and demand both adjust simultaneously to reach a market clearing price. The price swings discussed above, however, could be indicative of aberrant demand conditions that temporarily alter market prices. That is, during abnormal weather conditions we would see an increase in demand that occasions a shift along the supply curve forcing prices up. Supply in the short term will be reasonably fixed and therefore will not adjust. Once the abnormal

conditions fade demand will return to its normal level and prices will fall. The price movements are in this case are strictly due to demand not supply.

[11] This is modeled used in Schwartz's (1997) Journal of finance presidential address.

[12] The reader should note we are using the log of the spot price in expression (23) to capture the tendency of spot prices to revert around their mean. This is necessary because the process is geometric. That is, both the left and right hand sides of expression (23) are expressed in terms of returns. Thus we need to use the log of the spot price to model the deviations around the expected return.

[13] This joint process is the model introduced by Gibson and Schwartz (1990). Derivation of their pricing equation is found in appendix B. The closed form solution for this model is found in Bjerksund (1991).

[14] Since the spot commodity is a traded good we can state its risk premium analytically. We use the equilibrium condition for the futures contract to develop this quantity. The market price of risk for the convenience yield however, is not possible to express analytically. The convenience yield is not a traded good and we do not know its form. The parameter λ will remain in the analysis and must be estimated. In another analysis by Miltersen and Schwartz (1998), they use the method by Heath et al (1992) to bring the term structure of convenience yields into the analysis. This makes there results independent of the market price of convenience yield risk. Their result however, is only good for instruments that have long term maturities. Futures have maturities of just over a year, and do not lend themselves to this type of model. Thus, we use the previous method mentioned above.

[15] We can characterize the distribution of the futures price by finding the convolution between the pdfs for the spot price and convenience yields' Brownian motions. To demonstrate consider two random variables X and Y with pdf's given as $f_x(x)$ and $f_y(y)$. Let $Z = X + Y$, then $f_z(z) = \int_{-\infty}^{\infty} f_x(x) f_y(z-x)dx$, which is called the convolution of $f_x(x)$ and $f_y(y)$. Let $f_x(x) = \frac{1}{\sqrt{2\pi\sigma_x^2}} \exp\left(-\frac{(x-\mu_x)^2}{2\sigma_x^2}\right)$ and $f_x(y) = \frac{1}{\sqrt{2\pi\sigma_y^2}} \exp\left(-\frac{(y-\mu_y)^2}{2\sigma_y^2}\right)$, then what can we say about Z? If $X \sim N(0,\sigma^2)$ and $Y \sim N(0,1)$ then $Z = X + Y \sim N(0,\sigma^2+1)$. Proof

$$f_{x+y}(t) = \int_{-\infty}^{\infty} f_x(x) f_y(t-x)dx$$

$$= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{x^2}{2\sigma^2}\right) \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(t-y)^2}{2}\right) dx$$

$$= \frac{1}{2\pi\sigma} \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2}\left(\frac{x^2}{\sigma^2} - \frac{(t-x)^2}{1}\right)\right) dx. \qquad (1)$$

Rewriting the term inside the exponential we have

$$\frac{x^2}{\sigma^2} - \frac{(t-x)^2}{1} = \frac{x^2}{\sigma^2} + t^2 - 2tx + x^2$$

$$= \frac{x^2}{\sigma^2} + t^2 - 2tx + x^2 \frac{\sigma^2}{\sigma^2}$$

$$= \left(\frac{1+\sigma^2}{\sigma^2}\right) x^2 - 2tx + t^2. \qquad (2)$$

Now let $c = \frac{\sigma}{\sqrt{1+\sigma^2}}$, then the right hand side of expression (2) becomes

$$= \left(\frac{x^2}{c^2}\right) - 2tx + t^2$$

$$= \left(\frac{x^2}{c^2}\right) - 2tx + c^2t^2 + t^2 - c^2t^2$$

$$= \left(\frac{x}{c} - ct\right)^2 + \left(1-c^2\right)t^2$$

$$= \left(\frac{x-c^2t}{c}\right)^2 + \left(1-c^2\right)t^2.$$

Substituting this back into the integrand in expression (1)

$$f_{x+y}(x) = \frac{1}{2\pi\sigma}\int_{-\infty}^{\infty}\exp\left(-\frac{1}{2}\left(\frac{\left(x-c^2t\right)^2}{c^2} + \left(1-c^2\right)t^2\right)\right)dx$$

$$= \frac{1}{2\pi\sigma}\int_{-\infty}^{\infty}\exp\left(-\frac{1}{2}\left(\frac{\left(x-c^2t\right)^2}{c^2}\right)\right)dx\exp\left(-\frac{\left(1-c^2\right)t^2}{2}\right)$$

$$= \frac{1}{2\pi\sigma}\exp\left(-\frac{\left(1-c^2\right)t^2}{2}\right)\int_{-\infty}^{\infty}\exp\left(-\frac{1}{2}\left(\frac{\left(x-c^2t\right)^2}{c^2}\right)\right)dx$$

$$= \frac{1}{\sqrt{2\pi}\frac{\sigma}{c}}\exp\left(-\frac{\left(1-c^2\right)t^2}{2}\right)\frac{1}{\sqrt{2\pi}c}\int_{-\infty}^{\infty}\exp\left(-\frac{1}{2}\left(\left(\frac{x-c^2t}{c}\right)^2\right)\right)dx.$$

The integrand now sums to one and the above reduces to

$$= \frac{1}{\sqrt{2\pi}\left(\frac{\sigma}{c}\right)}\exp\left(-\frac{\left(1-c^2\right)t^2}{2}\right). \tag{3}$$

We can rewrite $(1-c^2)$ this is

$$(1-c^2)=1-\frac{\sigma}{1+\sigma^2}=\frac{1}{1+\sigma^2}.$$

In addition we may also write

$$\frac{\sigma}{c}=(1+\sigma^2)^{1/2}.$$

Substituting these expressions into equation (3) we obtain

$$=\frac{1}{\sqrt{2\pi(1+\sigma^2)}}\exp\left(-\frac{t^2}{2(1+\sigma^2)}\right). \qquad (4)$$

Expression (4) is the pdf of a random variable that is normally distributed with a mean of zero and a variance equal to $1+\sigma^2$.

[16] We have priced futures and forward contract and treated them as the same contract. With the marking to market for futures contract one may think the price of a futures contract is different from a forward contract. Cox et al (1981), and Jarrow and Oldfield (1981) show when interest rates are constants forwards and futures prices are the same. Consider the following example.

Suppose that futures contract lasts for n days and that $F(i)$ is the futures price at the end of the day i ($0 < i < n$). Define δ as the risk-free rate per day (assumed constant). Consider the following strategy.

1. Take a long futures position of e^{δ} at the end of day 0. The beginning of the contract.
2. Increase a long position to $e^{2\delta}$ at the end of day 1.
3. Increase a long position to $e^{3\delta}$ at the end of day 2.

And so on.

By the beginning of day i, the investor has a long position of $e^{\delta i}$. The profit (possibly negative) from the position on day i is

$$(F(i)-F(i-1))e^{\delta i}.$$

Assume that this is compounded at the risk-free rate until the end of day n. Its value at the end of day n is

$$(F(i)-F(i-1))e^{\delta i}e^{(n-i)\delta}=(F(i)-F(i-1))e^{n\delta}$$

The value at the end of day n of the entire investment horizon is therefore

$$\sum_{i=1}^{n}(F(i)-F(i-1))e^{n\delta}$$

That is,

$$[(F(n)-F(n-1))+(F(n-1)-F(n-2))+\cdots+(F(1)-F(0))]e^{n\delta}=(F(n)-F(0))e^{n\delta}$$

Since $F(n)$ is the same as the terminal asset price, $S(T)$, the terminal value of the investment strategy can be written

$$(S(T)-F(0))e^{n\delta}.$$

An investment of $F(0)$ in a risk-free bond combined with the strategy just given yields

$$F(0)e^{n\delta}+(S(T)-F(0))e^{n\delta}=S(T)e^{n\delta}$$

at time T. No investment is required for all the long futures positions described. It follows that an amount $F(0)$ can be invested to give an amount $S(T)e^{n\delta}$ at time T.

Suppose next that the forward price at the end of day 0 is $J(0)$. By investing $J(0)$ in a riskless bond and taking a long forward position of $e^{\delta n}$ forward contracts, an amount $S(T)e^{n\delta}$ is also guaranteed at time T. Thus, there are two investment strategies, one requiring an initial outlay of $F(0)$, the other requiring an initial outlay of $J(0)$, that yield $S(T)e^{n\delta}$ at time T. It follows that in the absence of arbitrage opportunities

$$J(0) = F(0)$$

In other words, the futures price and the forward price are identical. Note that in this proof there is nothing special about the time period of one day. The futures price based on a contract with weekly settlements is also the same as the forward price when corresponding assumptions are made.

[17] For the cautious reader, $f_T(t,T)$ is the derivative of the forward rate with respect to the maturity. As the maturity changes so does the forward rate.

[18] To understand how we may switch the order of integration we need to review some basic properties of the double integral. We start with a bounded region Ω in the xy-plane. We assume that Ω is a basic region. That is, we assume that boundary of Ω consists of a finite number of arcs $y = \varphi(x)$, or $x = \psi(y)$. Now, we want to define the double integral

$$\iint_{\Omega} f(x,y)\,dxdy.$$

To do this, we surround Ω by a rectangle R. We now extend f to all of R by setting f equal to zero outside of Ω. This extended function of f is bound on R, and it is continuous on all of R except possibly at the boundary of Ω. In spite of these possible discontinuities, f is still integrable on R; that is, there still exists a unique number I such that

$$L_f(P) \le I \le U_f(P)$$

for all partitions P of R. Where $L_f(P)$ is the lower sum and $U_f(P)$ is the upper sum. This number I is by definition the double integral

$$\iint_R f(x,y)dxdy$$

We define the double integral over Ω by setting

$$\iint_\Omega f(x,y)dxdy = \iint_R f(x,y)dxdy.$$

If f is nonnegative over Ω, the extended f is nonnegative on all of R. The double integral gives the volume of the solid trapped between the surface $z = f(x,y)$, and the rectangle R. But since the surface has height 0 outside of Ω, the volume outside of Ω is 0. It follows then that

$$\iint_\Omega f(x,y)dxdy$$

gives the volume of the solid T bounded above by $z = f(x,y)$ and below by Ω:

$$volume\ of\ T = \iint_\Omega f(x,y)dxdy.$$

The double integral

$$\iint_\Omega 1dxdy = \iint_\Omega dxdy$$

gives the volume of a solid of constant height 1 over Ω. In square units this is the area of Ω:

$$Area\ of\ \Omega = \iint_\Omega dxdy.$$

If an integral

$$\int_a^b f(x)dx$$

proves difficult to evaluate, it is not because of the interval [a,b] but because of the integrand *f*. Difficulty in evaluating a double integral

$$\iint_\Omega f(x,y)dxdy$$

can come from two sources: from the integrand *f* and from the base region Ω. Even such a simple looking integral as $\iint_\Omega 1dxdy$ is difficult to evaluate if Ω is complicated.

To evaluate the double integral we use the iterated integral approach. Ω is a basic region and we know that the double integral exists. The fundamental idea is that the double integral over sets of this structure can be reduced to a pair of ordinary integrals.

To evaluate the double integral above we do the following. The projection of Ω onto the *x*-axis is a closed interval [a,b] and Ω consists of all points of (*x,y*) with

$$a \le x \le b \text{ and } \varphi_1(x) \le y \le \varphi_2(x).$$

Then

$$\iint_\Omega f(x,y)dxdy = \int_a^b \int_{\varphi_1(x)}^{\varphi_2(x)} f(x,y)dydx$$

Here we have to first calculate

$$\int_{\varphi_1(x)}^{\varphi_2(x)} f(x,y)dy$$

by integrating *f*(*x,y*) with respect to *y* from $\varphi_1(x)$ to $\varphi_2(x)$. The resulting expression is a function of *x* alone, which we then integrate with respect to *x* from $x = a$ to $x = b$. We could have solved the double integral another way. That is, we can switch the order of integration and obtain the same result. Consider the following. The projection of Ω onto the *x*-axis is a closed interval [a,b] and Ω consists of all points of *(x,y)* with

$$a \le y \le b \text{ and } \varphi_1(y) \le x \le \varphi_2(y).$$

Then

$$\iint_{\Omega} f(x,y)dxdy = \int_{a}^{b} \int_{\varphi_1(y)}^{\varphi_2(y)} f(x,y)dxdy$$

This time we first calculate

$$\int_{\varphi_1(y)}^{\varphi_2(y)} f(x,y)dx$$

by integrating *f(x,y)* with respect to *x* from $\varphi_1(y)$ to $\varphi_2(y)$. The resulting expression is a function of *x* alone, which we then integrate with respect to *y* from $y = a$ to $y = b$.

Let us consider an example to elucidate the discussion above. Take the region Ω, which is bounded by the functions $y = x^2$ $(x = y^{1/2})$ and $y = x^{1/4}$ $(x = y^4)$. We want to evaluate the double integral

$$\iint_{\Omega} \left(x^{1/2} - y^2\right) dxdy .$$

The projection of Ω onto the *x*-axis is the closed interval [0,1] and Ω can be characterized as the set of all *(x,y)* with

$$0 \le x \le 1 \text{ and } x^2 \le y \le x^{1/4}.$$

Thus,

$$\iint_{\Omega}\left(x^{1/2}-y^{2}\right)dydx=\int_{0}^{1}\int_{x^{2}}^{x^{1/4}}\left(x^{1/2}-y^{2}\right)dydx$$

$$=\int_{0}^{1}\left[x^{1/2}y-\frac{1}{3}y^{3}\right]_{x^{2}}^{x^{1/4}}dx$$

$$=\int_{0}^{1}\left[\frac{2}{3}x^{3/4}-x^{5/2}+\frac{1}{3}x^{6}\right]dx$$

$$=\left[\frac{8}{21}x^{7/4}-\frac{2}{7}x^{7/2}+\frac{1}{21}x^{7}\right]_{0}^{1}$$

$$=\frac{8}{21}-\frac{2}{7}+\frac{1}{21}=\frac{1}{7}.$$

We can also switch the order of integration to achieve the same result. The projection of Ω onto the y-axis is the closed interval $[0,1]$ and Ω can be characterized as the set of all (x,y) with

$$0\le y\le 1 \text{ and } y^{4}\le x\le y^{1/2}.$$

Thus,

$$\iint_{\Omega}\left(x^{1/2}-y^{2}\right)dydx=\int_{0}^{1}\int_{y^{4}}^{y^{1/2}}\left(x^{1/2}-y^{2}\right)dxdy$$

$$=\int_{0}^{1}\left[\frac{2}{3}x^{3/2}-y^{2}x\right]_{y^{4}}^{y^{1/2}}dy$$

$$=\int_{0}^{1}\left[\frac{2}{3}y^{3/4}-y^{5/2}+\frac{1}{3}y^{6}\right]dy$$

$$=\left[\frac{8}{21}y^{7/4}-\frac{2}{7}y^{7/2}+\frac{1}{21}y^{7}\right]_{0}^{1}$$

$$= \frac{8}{21} - \frac{2}{7} + \frac{1}{21} = \frac{1}{7}.$$

[19] If liquid bonds that determine the term structure are all influenced by the same unpredictable Wiener process W_t, the respective prices must somehow be related to each other as suggested by the pricing relation:

$$B(t,T) = E_t^{*}\left[e^{-\int_t^T r_s ds} \right]$$

The classical approach to pricing interest rate sensitive securities is an attempt to extract these arbitrage relations from the $B(t,T)$ and then summarize them within an arbitrage-free spot rate model. This is indeed a complicated task of indirect accounting for a complex set of arbitrage relations between market prices. The Heath-Jarrow-Morton (1992), or as known as HJM approach, attacks these arbitrage restrictions directly by bringing the forward rates to the forefront. The risk-adjusted diffusion process for the HJM forward rate is

$$df(t,T) = \mu_f(t,T)dt + \sigma_f(t,T)dZ^{*}(t).$$

We can describe the movements of the above diffusion process with a binomial lattice model. To construct the lattice we use

$$f(t+\Delta t,T) = \begin{cases} f(t,T) + \mu_f(t,T)\Delta t + \sigma_f(t,T)\sqrt{\Delta t} \\ f(t,T) + \mu_f(t,T)\Delta t - \sigma_f(t,T)\sqrt{\Delta t} \end{cases},$$

Where $\mu_f(t,T) = \tanh(\omega(t,T))\sigma_f(t,T)$,

$$\tanh(\omega(t,T)) = \frac{e^{\omega} - e^{-\omega}}{e^{\omega} + e^{-\omega}}, \text{ and}$$

$$\omega(t,T) = \int_{t+\Delta t}^{T} \sigma_f(t,s)ds.$$

We start with a flat term structure

$$f(0,0) = r(0) = .1, f(0,1) = .1, f(0,2) = .1, \cdots.$$

and let $\Delta t = 1$ and $\sigma_f(t,T) = 0.02$. The correction term $\mu_f(t,T)$ is found by the following

$$\omega(t,T) = \int_{t+1}^{T} \sigma_f ds = (T - t - 1)\sigma_f$$

$$\omega(0,1) = (1 - 0 - 1)(.02) = 0,$$

$$\omega(0,2) = (2 - 0 - 1)(.02) = .02.$$

Substituting the above into

$$\tanh(\omega(t,T)) = \frac{e^{\omega} - e^{-\omega}}{e^{\omega} + e^{-\omega}}$$

yields

$$\tanh(\omega(0,1)) = \frac{e^{0} - e^{-0}}{e^{0} + e^{-0}} = 0,$$

$$\tanh(\omega(0,1)) = \frac{e^{.02} - e^{-.02}}{e^{.02} + e^{-.02}} = .019997334.$$

Therefore

$$\mu_f(t,T)\Delta t = \tanh(\omega(t,T))\sigma_f(t,T)\Delta t$$

is

$$\mu_f(0,1)(1) = \tanh(\omega(0,1))(.02)(1) = (0)(.02)(1) = 0$$

$$\mu_f(0,2)(1) = \tanh(\omega(0,2))(.02)(1) = (0.019997334)(.02)(1) = 0.0004.$$

The forward rates at time t = 1 are

$$r(1)_u = f(1,1)_u = .1 + 0 + .02 = .12,$$

$$f(1,2)_u = .1 + 0.0004 + .02 = .1204,$$

$$r(1)_d = f(1,1)_d = .1 + 0.02 = .08,$$

$$f(1,2)_d = .1 + 0.0004 - .02 = .0804,$$

The forward rates at time t = 2 are

$$r(2)_{uu} = f(2,2)_{uu} = .1204 + 0 + .02 = .1404,$$

$$r(2)_{dd} = f(1,1)_{dd} = .0804 + 0 + .02 = .0804,$$

$$r(2)_{du} = f(1,1)_{du} = .0804 + 0 + .02 = .1004,$$

$$r(2)_{ud} = f(2,2)_{ud} = .1204 + 0 - .02 = .1004.$$

$r(2)_{uu} = .1404$

$r(1)_u = .12$
$f(1,2)_u = .1204$

$r(0) = .1$
$f(0,1) = .1$
$f(0,2) = .1$

$r(2)_{ud} = .1004$

$r(1)_d = .08$
$f(1,2)_d = .0804$

$r(2)_{dd} = .0604$

Given the lattice above we can calculate the price of the pure discount bonds. The bonds are given by

$$P(t,T) = \$1e^{-\sum f(t,T)}.$$

The corresponding pure discount bond prices at time t = 0 are

$$P(0,1) = 1e^{-.1} = .904837418,$$

$$P(0,2) = 1e^{-.1(2)} = .818730753,$$

$$P(0,3) = 1e^{-.1(3)} = .740818221.$$

Bond prices at time t =1 are

$$P(1,2)_u = 1e^{-.12} = .886920437,$$

$$P(1,3)_u = 1e^{-(.12+.1204)} = .786313315,$$

$$P(1,2)_d = 1e^{-.08} = .923116346,$$

$$P(1,3)_d = 1e^{-(.08+.0804)} = .851803045.$$

Bond prices at time t = 3 are

$$P(2,3)_{uu} = 1e^{-.1404} = .869010608,$$

$$P(2,3)_{ud} = 1e^{-.1004} = .904475603,$$

$$P(2,3)_{dd} = 1e^{-.0604} = .941387903.$$

The lattice for the bond prices is

.869010608

.886920437
.786313315

.904837418
.818730753
.740818221

.904475603

.923116346
.851803045

.941387903

The bond prices above are arbitrage free. Hence, we may calculate the price of a bond today under an equivalent martingale measure. That is

$$P(t,T) = E_t^*[P(t+\Delta t,T)]P(t,t+\Delta t).$$

The price of a three year pure discount bond is

$$P(0,3) = E_0^*[P(2,3)]P(0,2),$$

$$P(0,3) = \left[(p^*)^2 P(2,3)_{uu} + (1-p^*)(p^*)P(2,3)_{ud} + (1-p^*)^2 P(2,3)_{dd}\right]P(0,2)$$

$$P(0,3) = \left[\frac{1}{4}(.869010608) + \frac{1}{2}(.9044750603) + \frac{1}{4}(.941387903)\right](.818730753)$$

$P(0,3) = .740818221.$

Note, the correction term $\mu_f(t,T)$ is needed to insure no arbitrage opportunities exist. If we did not use this correction term then we cannot live up to the no arbitrage condition. Consider the following. Let

$$f(t+\Delta t,T) = \begin{cases} f(t,T) + \sigma_f(t,T)\sqrt{\Delta t} \\ f(t,T) - \sigma_f(t,T)\sqrt{\Delta t} \end{cases}.$$

The lattice for the forward rates would be

$r(2)_{uu} = .14$

$r(1)_u = .12$
$f(1,2)_u = .12$

$r(0) = .1$
$f(0,1) = .1$
$f(0,2) = .1$

$r(2)_{ud} = .10$

$r(1)_d = .08$
$f(1,2)_d = .08$

$r(2)_{dd} = .06$

The possible prices of a discount bond at time t = 2 maturing at time t = 3 are

$$P(2,3)_{uu} = 1e^{-.14} = .86358235,$$

$$P(2,3)_{ud} = 1e^{-.10} = .904837418,$$

$$P(2,3)_{dd} = 1e^{-.06} = .941764534.$$

The price of a three year discount bond under an equivalent martingale is

$$P(0,3) = \left[\frac{1}{4}(.86358235) + \frac{1}{2}(.904837418) + \frac{1}{4}(.941764534)\right](.818730753)$$

$$P(0,3) = .739932364.$$

This price is not equal to the three year discount bond price using today's term structure. Thus the above process permits arbitrage opportunities.

[20] Using equation (145),

$$r(t) = f(0,t) + \int_0^t \frac{\partial\sigma(s,t,P(s,t))}{\partial t}\left[\int_s^t \frac{\partial\sigma(s,u,P(s,u))}{\partial u}du\right]ds + \int_0^t \frac{\partial\sigma(s,t,P(s,t))}{\partial t}dZ_r^*(s),$$

to derive the spot rate we have,

$$r(v) = f(t,v) + \int_t^v \sigma_f(s,v)\left[\int_s^v \sigma_p(s,y)dy\right]ds + \int_t^v \sigma_f(s,v)dZ_r^*(s).$$

If $\sigma_f(s,v) = \sigma_r e^{-k_r(v-s)}$ then the above may be written as

$$r(v) = f(t,v) + \int_t^v \sigma_r e^{-k_r(v-s)}\left[\int_s^v \sigma_r e^{-k_r(y-s)}dy\right]ds + \int_t^v \sigma_r e^{-k_r(v-s)}dZ_r^*(s).$$

Evaluating the expression above yields

$$r(v) = f(t,v) + \int_t^v \sigma_r e^{-k_r(v-s)}\left(\frac{-\sigma_r}{k_r}\right)\left[e^{-k_r(v-s)} - e^{-k_r(s-s)}\right]ds + \int_t^v \sigma_r e^{-k_r(v-s)}dZ_r^*(s),$$

$$= f(t,v) + \left(\frac{-\sigma_r^2}{k_r}\right)\int_t^v e^{-k_r(v-s)}\left[e^{-k_r(v-s)} - e^{-k_r(s-s)}\right]ds + \int_t^v \sigma_r e^{-k_r(v-s)}dZ_r^*(s),$$

$$= f(t,v) + \left(\frac{-\sigma_r^2}{k_r}\right)\int_t^v e^{-k_r(v-s)}\left[e^{-k_r(v-s)} - 1\right]ds + \int_t^v \sigma_r e^{-k_r(v-s)}dZ_r^*(s),$$

$$= f(t,v) + \left(\frac{-\sigma_r^2}{k_r}\right)\int_t^v \left[e^{-2k_r(v-s)} - e^{-k_r(v-s)}\right]ds + \int_t^v \sigma_r e^{-k_r(v-s)} dZ_r^*(s),$$

$$= f(t,v) - \left(\frac{\sigma_r^2}{k_r}\right)\int_t^v e^{-2k_r(v-s)} ds + \left(\frac{\sigma_r^2}{k_r}\right)\int_t^v e^{-k_r(v-s)} ds + \int_t^v \sigma_r e^{-k_r(v-s)} dZ_r^*(s),$$

$$= f(t,v) + \frac{\sigma_r^2}{k_r^2}\left(e^{-k_r(v-v)} - e^{-k_r(v-t)}\right) - \frac{\sigma_r^2}{2k_r^2}\left(e^{-2k_r(v-v)} - e^{-2k_r(v-t)}\right) + \int_t^v \sigma_r e^{-k_r(v-s)} dZ_r^*(s),$$

$$= f(t,v) + \frac{\sigma_r^2}{k_r^2}\left(1 - e^{-k_r(v-t)}\right) - \frac{\sigma_r^2}{2k_r^2}\left(1 - e^{-2k_r(v-t)}\right) + \int_t^v \sigma_r e^{-k_r(v-s)} dZ_r^*(s),$$

$$= f(t,v) + \frac{\sigma_r^2}{k_r^2} - \frac{\sigma_r^2}{k_r^2} e^{-k_r(v-t)} - \frac{\sigma_r^2}{2k_r^2} + \frac{\sigma_r^2}{2k_r^2} e^{-2k_r(v-t)} + \int_t^v \sigma_r e^{-k_r(v-s)} dZ_r^*(s),$$

$$= f(t,v) + \frac{\sigma_r^2}{2k_r^2}\left(2 - 2e^{-k_r(v-t)} - 1 + e^{-2k_r(v-t)}\right) + \sigma_r e^{-k_r v}\int_t^v e^{k_r s} dZ_r^*(s),$$

$$= f(t,v) + \frac{\sigma_r^2}{2k_r^2}\left(1 - 2e^{-k_r(v-t)} + e^{-2k_r(v-t)}\right) + \sigma_r e^{-k_r v}\int_t^v e^{k_r s} dZ_r^*(s).$$

The expression immediately above is the same as equation (147), where we derived this expression from (145) not (146).

[21] During the development of the instantaneous spot rate process, we relied extensively on the HJM methodology. This method uses the arbitrage free bond price dynamics to determine the forward rate dynamics. It turns out that the only input to determine the forward rate dynamics is the volatility of the term structure. Once we have the volatility then we have the forward rate dynamics which includes the instantaneous spot rate. The derivation for this process is economically appealing, insightful and promotes the use of this specification in developing asset pricing models. The forward rate dynamics used in

the HJM model, however, is not unique to them. These stochastic differential equations are a particular class of time varying coefficient stochastic processes used extensively in the engineering literature.

As we have already seen an economic derivation of the instantaneous spot rate from the HJM perspective, we now concentrate on the intuitiveness of this derivation from a mathematical perspective. The purpose for using a time varying coefficient stochastic process to model forward rate movements is it allows the forward rate to revert around its mean while its average moves over time. HJM realizes that as new information arrives each period the term structure changes. The actual magnitude of the change may depend of the forward rate itself, on its maturity date, and on other factors. The change in the forward rate need not be constant for all maturities. Given these asymmetric changes in the forward rate, we could expect forward rates to oscillate around their means and for their averages to move across time as well. Thus, the use of a time varying coefficient stochastic process would be a natural selection to model this behavior.

What follows is the derivation of the instantaneous spot rate (expression 147) using the general specification of the time varying coefficient process. In addition, we will be capable of deriving the stochastic differential of the instantaneous spot rate (expression 128).

Let dynamics of the instantaneous spot rate be described as

$$dr(t) = \left(k_r m(t) + m'(t) - k_r r(t)\right)dt + \sigma dW(t) \qquad (1)$$

where $m(t) = f(t,v) + \frac{\sigma^2}{2k_r^2}\left(1 - e^{-k_r(t-v)}\right)^2$,

$$= f(t,v) + \frac{\sigma^2}{2k_r^2}\left(1 - 2e^{-k_r(t-v)} + e^{-2k_r(t-v)}\right). \qquad (2)$$

Given the above specification, we may restate the stochastic differential for the instantaneous spot rate. To accomplish this, we must first find $m'(t)$. This is

$$m'(t) = \frac{\partial\left(f(t,v) + \frac{\sigma^2}{2k_r^2}\left(1 - 2e^{-k_r(t-v)} + e^{-2k_r(t-v)}\right)\right)}{\partial v}$$

$$= f_v(t,v) + \frac{\sigma^2}{k_r^2}e^{-k_r(t-v)} - \frac{\sigma^2}{k_r^2}e^{-2k_r(t-v)}. \qquad (3)$$

Using equations (2) and (3) in (1), we can obtain equation (128) in chapter 3. This is done as follows

$$dr(t) = \left(k_r\left(f(t,v) + \frac{\sigma^2}{2k_r^2}\left(1 - 2e^{-k_r(t-v)} + e^{-2k_r(t-v)}\right)\right) + f_v(t,v) + \frac{\sigma^2}{k_r^2}e^{-k_r(t-v)} - \frac{\sigma^2}{k_r^2}e^{-2k_r(t-v)}\right.$$

$$\left. - k_r r(t)\right)dt + \sigma dW(t),$$

$$= \left(f_v(t,v) + k_r f(t,v) + \frac{\sigma^2}{2k_r} - \frac{\sigma^2}{k_r}e^{-k_r(t-v)} + \frac{\sigma^2}{2k_r}e^{-2k_r(t-v)} + \frac{\sigma^2}{k_r}e^{-k_r(t-v)}\right.$$

$$\left. - \frac{\sigma^2}{k_r}e^{-2k_r(t-v)} - k_r r(t)\right)dt + \sigma dW(t),$$

$$= \left(f_v(t,v) + k_r f(t,v) + \frac{\sigma^2}{2k_r} + \frac{\sigma^2}{2k_r}e^{-2k_r(t-v)} - \frac{\sigma^2}{k_r}e^{-2k_r(t-v)} - k_r r(t)\right)dt + \sigma dW(t),$$

$$= \left(f_v(t,v) + k_r f(t,v) + \frac{\sigma^2}{2k_r} - \frac{\sigma^2}{2k_r}e^{-2k_r(t-v)} - k_r r(t)\right)dt + \sigma dW(t),$$

$$= \left(f_v(t,v) + k_r f(t,v) + \frac{\sigma^2}{2k_r}\left(1 - e^{-2k_r(t-v)}\right) - k_r r(t)\right)dt + \sigma dW(t). \qquad (4)$$

Equation (4) is the stochastic differential we wanted to find. We may now derive expression (147). This is

$$dr(t) = \left(k_r m(t) + m'(t) - k_r r(t)\right)dt + \sigma dW(t),$$

$$dr(t) + k_r r(t)dt = \left(k_r m(t) + m'(t)\right)dt + \sigma dW(t),$$

$$e^{k_r t}\left(dr(t) + k_r r(t)dt\right) = e^{k_r t}\left(\left(k_r m(t) + m'(t)\right)dt + \sigma dW(t)\right),$$

$$d\left(r(t)e^{k_r t}\right) = e^{k_r t}\left(\left(k_r m(t) + m'(t)\right)dt + \sigma dW(t)\right),$$

$$d\left(r(t)e^{k_r t}\right) = \left(e^{k_r t}k_r m(t) + e^{k_r t}m'(t)\right)dt + \sigma e^{k_r t}dW(t),$$

$$d\left(r(t)e^{k_r t}\right) = d\left(e^{k_r t}m(t)\right) + \sigma e^{k_r t}dW(t),$$

$$\int_0^t d\left(r(s)e^{k_r s}\right) = \int_0^t d\left(e^{k_r s}m(s)\right) + \sigma\int_0^t e^{k_r s}dW(s),$$

$$d\left(r(s)e^{k_r s}\right)\Big|_0^t = d\left(e^{k_r s}m(s)\right)\Big|_0^t + \sigma\int_0^t e^{k_r s}dW(s),$$

$$r(t)e^{k_r t} - r(0)e^{k_r 0} = e^{k_r t}m(t) - e^{k_r 0}m(0) + \sigma\int_0^t e^{k_r s}dW(s).$$

Recall that

$$m(t) = f(0,t) + \frac{\sigma^2}{2k_r^2}\left(1 - e^{-k_r t}\right)^2,$$

thus $m(0) = f(0,0) + 0$.

Using the expressions above we find

$$r(t)e^{k_r t} - r(0) = e^{k_r t}\left(f(0,t) + \frac{\sigma^2}{2k_r^2}\left(1 - e^{-k_r t}\right)^2\right) - f(0,0) + \sigma\int_0^t e^{k_r s}dW(s),$$

$$r(t)e^{k_r t} = e^{k_r t}\left(f(0,t) + \frac{\sigma^2}{2k_r^2}\left(1 - e^{-k_r t}\right)^2\right) - f(0,0) + r(0) + \sigma\int_0^t e^{k_r s}dW(s),$$

$$r(t)e^{k_r t} = e^{k_r t}\left(f(0,t) + \frac{\sigma^2}{2k_r^2}\left(1 - e^{-k_r t}\right)^2 \right) + \sigma \int_0^t e^{k_r s} dW(s),$$

$$r(t) = e^{-k_r t} e^{k_r t}\left(f(0,t) + \frac{\sigma^2}{2k_r^2}\left(1 - e^{-k_r t}\right)^2 \right) + e^{-k_r t} \sigma \int_0^t e^{k_r s} dW(s),$$

$$r(t) = \left(f(0,t) + \frac{\sigma^2}{2k_r^2}\left(1 - e^{-k_r t}\right)^2 \right) + e^{-k_r t} \sigma \int_0^t e^{k_r s} dW(s),$$

$$r(t) = \left(f(0,t) + \frac{\sigma^2}{2k_r^2}\left(1 - 2e^{-k_r t} + e^{-2k_r t}\right) \right) + e^{-k_r t} \sigma \int_0^t e^{k_r s} dW(s). \qquad (5)$$

This is equivalent to equation (147) in chapter 3.

Chapter 4

Commodity Futures Option Pricing Models

The focus of our analysis to this point is centered on pricing commodity futures contracts. We begin with a simple mean reverting spot price and then add greater complexity by introducing a stochastic convenience yield as well as a random interest rate. In each of these models, we note that the solutions to the underlying systems of stochastic differential equations are themselves random processes. Consequently, there is an implicit price dynamic inherent in the futures price, which we may state formally (This was detailed in chapter 3 for the one- and two-factor models). Once the futures price dynamics are in place, we may begin to price option contracts written on the commodity futures contract. The objective of this chapter is to develop a pricing model for these options.

The theoretical underpinnings of the option models in this chapter are predicated on the original work of Black and Scholes (1973). Anticipating the use of the Black-Scholes methodology we begin with a review of this analysis. As we reflect on Black-Scholes work, we explore two different approaches used in the literature to derive the Black-Scholes formula. The first approach is the traditional arbitrage methodology used in the Black-Scholes analysis. The second method is a partial expectation model working with an equivalent martingale measure. Once we conclude our review, the remaining analysis for this chapter proceeds with the partial expectation method of solution.

In the final section of this chapter, we extend the above option models by introducing the effects of a discrete random jump to the futures price process. Occasionally, commodities are influenced by aberrant market conditions, whereby, prices tend to exhibit discrete jumps. For the case of crude oil, these aberrant conditions may arise because of the weather, war, or other political unrest in a particular country. This evidence suggests commodity prices may not follow a geometric Brownian motion all the time, which leads us to investigate other means of pricing futures options. Therefore, the remaining portion of this chapter investigates jump-diffusion models in the finance literature and their ability to price options written on commodity futures.

4.1 Black-Scholes Option Pricing Model

The Black-Scholes option pricing model prices options that are written on a non-dividend paying stock. To derive this formula, we need to understand how the underlying asset behaves. That is, we need to know the salient time series characteristics of the stock price. Black-Scholes posit that the stock prices follow a geometric Brownian motion, which we formally describe as

$$\frac{dS(t)}{S(t)} = \mu dt + \sigma dZ(t), \tag{1}$$

where μ is the mean return and σ is the diffusion coefficient. $dZ(t)$ is the increment of a standard Brownian motion. Given the return dynamics of the stock price we can now say something about the derivative security written on it.

Suppose an option contract can be written as a twice-continuously differentiable function of the stock price and time, namely $C(S,t)$. If the stock price follows the dynamics described in expression (1) then the option return dynamics can be written in a similar form as

$$\frac{dC(S,t)}{C(S,t)} = \mu_c dt + \sigma_c dZ(t), \tag{2}$$

where μ_c is the mean return and σ_c is the diffusion coefficient.[1] The term $dZ(t)$ is the increment of a standard Brownian motion, and later we will discuss the nature of this innovation term.

Equation (2) is a general expression for the return dynamics of the option contract. We know the option is a function of the stock price so it too must follow a stochastic differential equation with drift μ_c and diffusion σ_c, but it need not follow a geometric Brownian motion. We may determine the actual drift and diffusion terms for the option by formally developing the stochastic differential equation in expression (2). We use Ito's lemma to develop this dynamic. Invoking Ito's lemma the increment for the option contract is

$$
\begin{aligned}
dC(S,t) &= C_s dS(t) + \frac{1}{2}C_{ss}[dS(t)]^2 + C_t dt, \\
&= C_s[\mu S(t)dt + \sigma S(t)dZ(t)] + \frac{1}{2}C_{ss}\sigma^2[S(t)]^2 dt + C_t dt, \\
&= \left[\frac{1}{2}C_{ss}\sigma^2[S(t)]^2 + \mu S(t)C_s + C_t\right]dt + \sigma S(t)C_s dZ(t). \qquad (3)
\end{aligned}
$$

Expression (3) is the stochastic differential of the call option's price. To express these movements in terms of returns we divide the left-hand side and the right-hand side by $C(S,t)$. That is,

$$
\frac{dC(S,t)}{C(S,t)} = \left[\left(\frac{1}{2}C_{ss}\sigma^2[S(t)]^2 + \mu S(t)C_s + C_t\right)\Big/ C(S,t)\right]dt + [\sigma S(t)C_s / C(S,t)]dZ(t). \qquad (4)
$$

Expressions (4) and (2) are expressed in the same units (returns) and they are describing the behavior of the same process. Hence, they are equal and we can equate the drift and diffusion terms in both expressions. We obtain

$$\mu_c = \left[\frac{1}{2}C_{ss}\sigma^2[S(t)]^2 + \mu S(t)C_s + C_t\right] \Big/ C(S,t), \qquad (5)$$

$$\sigma_c = \sigma S(t)C_s / C(S,t). \qquad (6)$$

Earlier in equation (2) we did not formally define the innovation term. We now see that the Brownian motion in equation (4) is the same as the Brownian motion in the spot price shown equation (1). Thus, the call option is driven by the same innovation term as the spot price.

With the return characteristics for the stock and option contract in place, consider the following investment strategy. We hold portions of the stock, the option, and a riskless asset in a portfolio, where the portfolio weights are denoted as $w_1, w_2,$ and w_3 respectively. The portfolio weights sum to one, $\sum_{j=1}^{3} w_j = 1$, and the value of the portfolio is denoted as $A(t)$. Since the portfolio is a function of both the stock and the option, we may express the return dynamics of the portfolio as a stochastic process. This dynamic is expressed in similar fashion as the stock return dynamics. Formally,

$$\frac{dA(t)}{A(t)} = \mu_A dt + \sigma_A dZ(t), \qquad (7)$$

where μ_A is the mean return and σ_A is the diffusion coefficient. For now we define $dZ(t)$ as the increment of a standard Brownian motion, with an expected value of zero and a variance of $\sqrt{t}$. Below, we will see that this term is the linear combination of the individual asset diffusion coefficients.

Since the portfolio is a linear combination of the three assets, its drift and diffusion terms are linear combinations of the drift and diffusion coefficients of the individual assets. The drift term is expressed as

$$\mu_A = w_1\mu + w_2\mu_c + w_3\mu_r .$$

Using the constraint $w_1 + w_2 + w_3 = 1$ and the fact that the expected return on the riskless asset is equal to r the above expression is rewritten as

$$\mu_A = w_1\mu + w_2\mu_c + (1 - w_1 + w_2)r ,$$

$$\mu_A = w_1(\mu - r) + w_2(\mu_c - r) + r \qquad (8)$$

The diffusion coefficient for the portfolio is

$$\sigma_A = w_1\sigma + w_2\sigma_c + w_3\sigma_r .$$

By definition $\sigma_r = 0$ so we are left with

$$\sigma_A = w_1\sigma + w_2\sigma_c \qquad (9)$$

The variance of the portfolio is a function of the volatilities for the underlying asset and the derivative security.

We now have the dynamics of the portfolio. In obtaining the drift and diffusion terms we simply took the linear combination of the underlying assets coefficients. This property is of interest to an investor. The reason lies in the diffusion term. The diffusion term for the portfolio is the sum of the two diffusion coefficients for the option and stock processes. These terms are not unrelated. Both the stock and option processes are driven by the same innovation term. Such dependence makes it possible for an investor to take offsetting positions in these assets to eliminate the randomness from the portfolio. Therefore, an investor may strategically choose a weighting scheme so that the diffusion coefficient equals zero or $\sigma_A = 0$. Consequently, if the portfolio in expression (7) is risk-free, then the return to this portfolio over the investment horizon should equal the riskless rate of return. That is, $\mu_A = r$. Hence, we can write expressions (8) and (9) as

$$\mu_A = w_1(\mu - r) + w_2(\mu_c - r) + r = r$$

$$\sigma_A = w_1\sigma + w_2\sigma_c = 0,$$

or

$$\mu_A - r = w_1(\mu - r) + w_2(\mu_c - r) = 0 \tag{10}$$

$$\sigma_A = w_1\sigma + w_2\sigma_c = 0. \tag{11}$$

We may express the above system in the following form

$$\begin{bmatrix} \mu - r & \mu_c - r \\ \sigma & \sigma_c \end{bmatrix} \begin{bmatrix} w_1^* \\ w_2^* \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}. \tag{12}$$

The system of equations in (12) is known as a homogeneous equation system. There are two possible solutions to this system. The first is called the trivial solution where the weights are all equal to zero, $w_i's = 0$. This solution, however, is not a viable solution since by definition there must be some investment. The second solution and more appropriate is the non-trivial solution where the portfolio weights are all non zero, $w_i's \neq 0$. The only way to obtain a nontrivial solution from a homogeneous system of equations is if the coefficient matrix is singular. That is,

$$\begin{vmatrix} \mu - r & \mu_c - r \\ \sigma & \sigma_c \end{vmatrix} = 0.$$

The above condition implies that the row vector $[\mu - r, \mu_c - r]$ is a multiple of the row vector $[\sigma, \sigma_c]$; consequently one of these two equations is redundant. Dealing with the first equation we see

$$w_1^*(\mu - r) + w_2^*(\mu_c - r) = 0,$$

$$w_1^*(\mu - r) = -w_2^*(\mu_c - r),$$

$$w_1^* = -w_2^* \frac{(\mu_c - r)}{(\mu - r)}.$$

Now substituting this expression into the second equation we have

$$\left(-w_2^* \frac{(\mu_c - r)}{(\mu - r)}\right)\sigma + w_2^*\sigma_c = 0,$$

$$w_2^*\left[\left(-\frac{(\mu_c - r)}{(\mu - r)}\right)\sigma + \sigma_c\right] = 0.$$

We are looking for the non-trivial solution $w_i^* \neq 0$. The only way the equation above equals zero is if the expression inside the brackets equals zero. This implies

$$\left(-\frac{(\mu_c - r)}{(\mu - r)}\right)\sigma + \sigma_c = 0,$$

$$\left(\frac{(\mu_c - r)}{(\mu - r)}\right)\sigma = \sigma_c,$$

$$\sigma(\mu_c - r) = \sigma_c(\mu - r),$$

$$\frac{(\mu_c - r)}{\sigma_c} = \frac{(\mu - r)}{\sigma}. \tag{13}$$

The non-trivial solution for the homogeneous system of equations in expression (12) is

$$w_i^* = \frac{\gamma_i - r}{s_i}, \tag{14}$$

where $\gamma_i = [\mu, \mu_c]$ and $s_i = [\sigma, \sigma_c]$. To illustrate the above consider the expression

$$w_1^* = -w_2^* \frac{(\mu_c - r)}{(\mu - r)}.$$

Substitute this expression into

$$w_1^* \sigma + w_2^* \sigma_c = 0,$$

$$-w_2^* \frac{(\mu_c - r)}{(\mu - r)} \sigma + w_2^* \sigma_c = 0. \tag{15}$$

If $w_2^* = \frac{(\mu_c - r)}{\sigma_c}$, then expression (15) becomes

$$-\left(\frac{(\mu_c - r)}{\sigma_c}\right) \frac{(\mu_c - r)}{(\mu - r)} \sigma + \left(\frac{(\mu_c - r)}{\sigma_c}\right) \sigma_c = 0,$$

$$-\frac{(\mu_c - r)(\mu_c - r)}{\sigma_c (\mu - r)} \sigma + \left(\frac{(\mu_c - r)}{\sigma_c}\right) \sigma_c = 0,$$

$$-\frac{(\mu_c - r)^2}{\sigma_c^2} + \left(\frac{(\mu_c - r)}{\sigma_c}\right) \frac{(\mu - r)}{\sigma} = 0. \tag{16}$$

Recall from expression (13), that

$$\frac{(\mu_c - r)}{\sigma_c} = \frac{(\mu - r)}{\sigma}.$$

Therefore, expression (16) reduces to

$$-\frac{(\mu_c - r)^2}{\sigma_c^2} - \frac{(\mu - r)^2}{\sigma^2} = 0.$$

The result above shows that the optimal weighting scheme,

$$w_i^* = \frac{\gamma_i - r}{s_i},$$

is a solution to the above system of homogeneous equations.

Now substituting expressions (5) and (6) into expression (13) we obtain

$$\frac{\mu - r}{\sigma} = \frac{\left\{\left[\frac{1}{2}C_{ss}\sigma^2[S(t)]^2 + \mu S(t)C_s + C_t\right]\Big/C(S,t)\right\} - r}{\sigma S(t)C_s/C(S,t)},$$

$$\frac{(\mu - r)\sigma S(t)C_s}{\sigma C(S,t)} = \frac{\left[\frac{1}{2}C_{ss}\sigma^2[S(t)]^2 + \mu S(t)C_s + C_t\right]}{C(S,t)} - r,$$

$$\frac{(\mu - r)\sigma S(t)C_s}{\sigma} = \frac{1}{2}C_{ss}\sigma^2[S(t)]^2 + \mu S(t)C_s + C_t - rC(S,t),$$

$$(\mu - r)S(t)C_s = \frac{1}{2}C_{ss}\sigma^2[S(t)]^2 + \mu S(t)C_s + C_t - rC(S,t),$$

$$-\mu S(t)C_s + \mu S(t)C_s + \frac{1}{2}C_{ss}\sigma^2[S(t)]^2 + rS(t)C_s + C_t - rC(S,t) = 0,$$

$$\frac{1}{2}C_{ss}\sigma^2[S(t)]^2 + C_s rS(t) - rC(S,t) + C_t = 0. \tag{17}$$

Expression (17) is a partial differential equation for an option contract written on a stock whose price follows a geometric Brownian motion. This is a unique result in that we can now express the option dynamics deterministically. That is, earlier we begin with a stochastic differential equation for option prices, but after using a no arbitrage condition, we eliminate the randomness from the option price movements. This allows us to find a pricing formula for the call option. We only need the necessary boundary conditions to solve the partial differential equation numerically or analytically. For the Black-Scholes model, the boundary conditions for the option are

$$C(0,\tau) = 0, \tag{18}$$

$$C(S,0) = \max[0, S(T) - X], \tag{19}$$

where X is the exercise price of the option and $\tau = T - t$ is the time to maturity. Intuitively, expressions (18) and (19) are contractual clauses for an option contract. Expression (18) states that anytime the spot price is zero, which implies that no market exists for the asset, the option will be worthless. Expression (19) states that when the option matures the value of the option will equal the greater of the two amounts, $S(T) - X$ or 0. The function (solution) that satisfies (17), (18), and (19) is

$$C(S,t) = S(t)N(d_1) - e^{-r\tau} XN(d_2), \tag{20}$$

where
$$N(d_1) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{d_1} e^{-\frac{z^2}{2}} dz, \tag{21}$$

$$d_1 = \left[\ln\left(\frac{S(t)}{X}\right) + \left(r + \frac{1}{2}\sigma^2\right)\tau \right] \Big/ \sigma\sqrt{\tau}, \tag{22}$$

$$d_2 = d_1 - \sigma\sqrt{\tau}. \tag{23}$$

Expression (20) is the well known Black-Scholes pricing formula for a call option. The expression is a function of the underlying stock price, the exercise price, the volatility of the stock price, the risk free rate of interest and time to maturity.

The derivation for the Black-Scholes option pricing formula presented above is the traditional partial equilibrium model presented in many financial economic textbooks. From the above analysis, we see that the Black-Scholes method of solution is capable of pricing options with a partial differential equation that is the predicate of an arbitrage portfolio. There is, however, an alternative method to obtaining the Black-Scholes solution. We may develop the Black-Scholes option pricing solution via a partial expectation model where the expectation is taken with an equivalent martingale measure.

The Black-Scholes methodology of option pricing exploits a partial differential equation implied by an arbitrage-free portfolio. Recall, from chapter three, that recent methods of derivative pricing do not necessarily exploit partial differential equations, instead they rest on converting prices of such assets into martingales. This is done through transforming the underlying probability distributions of the assets.[2] We know, for instance, that the spot price process

$$\frac{dS(t)}{S(t)} = \mu dt + \sigma dZ(t),$$

can be transformed using the Girsanov theorem. The Girsanov theorem is a method for changing the probability distribution for a continuous time stochastic process. Invoking the theorem, the Brownian motion for the spot price above maybe written as

$$dZ(t) = dZ^{*}(t) - \lambda dt.$$

In chapter three, expression (7) shows the market price of risk for the spot price to be

$$\lambda = \frac{\mu - r}{\sigma}.$$

The risk adjusted Brownian motion becomes

$$dZ(t) = dZ^{*}(t) - \left(\frac{\mu - r}{\sigma}\right)dt.$$

Substituting the risk adjusted Weiner process into the spot price process yields

$$\frac{dS(t)}{S(t)} = \mu dt + \sigma\left[dZ^{*}(t) - \left(\frac{\mu - r}{\sigma}\right)dt\right],$$

$$\frac{dS(t)}{S(t)} = \mu dt - (\mu - r)dt + \sigma dZ^{*}(t),$$

$$\frac{dS(t)}{S(t)} = rdt + \sigma dZ^{*}(t). \tag{24}$$

We have successfully altered the drift term for the spot price by altering the underlying probability distribution. As a result, we see that the spot price now trends at a known rate, *r*, as opposed to an unknown rate, μ. As we will see below, this has useful implications for pricing any contingent claim written on the spot price.

The spot price process in equation (24) is a geometric Brownian motion under an equivalent martingale measure. An alternative representation of expression (24) can be obtained. Let $H(t) = \ln S(t,T)$. The transformation, and Ito's lemma yield the following process for the increment of $H(t)$

$$dH = H_S dS + \frac{1}{2} H_{SS} dS^2,$$

$$dH = \frac{1}{S}(S)\left(rdt + \sigma dZ^*(t)\right) - \frac{1}{2}\frac{1}{S^2}\left(S^2\sigma^2 dt\right),$$

$$dH = rdt + \sigma dZ^*(t) - \frac{1}{2}\sigma^2 dt,$$

$$dH = \left(r - \frac{1}{2}\sigma^2\right)dt + \sigma dZ^*(t).$$

Integrating over the above expression yields

$$\int_t^T dH = \int_t^T\left(r - \frac{1}{2}\sigma^2\right)ds + \int_t^T \sigma dZ^*(s),$$

$$H(T) - H(t) = \int_t^T\left(r - \frac{1}{2}\sigma^2\right)ds + \int_t^T \sigma dZ^*(s),$$

$$H(T) - H(t) = \left(r - \frac{1}{2}\sigma^2\right)(T - t) + \int_t^T \sigma dZ^*(s)$$

$$H(T) = H(t) + \left(r - \frac{1}{2}\sigma^2\right)\tau + \int_t^T \sigma dZ^*(s),$$

$$S(T) = S(t)\exp[Y(T)], \qquad (25)$$

where $Y(T) = \left(r - \frac{1}{2}\sigma^2\right)\tau + \int_t^T \sigma dZ^*(s)$, and $\tau = T - t$ is the time to maturity. Expression (25) shows that the log returns, $\ln\left(\frac{S(T)}{S(t)}\right) = Y(T)$, for the spot price are normally distributed with a mean of $\left(r - \frac{1}{2}\sigma^2\right)\tau$ and a variance of $\sigma^2\tau$.[3] Furthermore, the conditional forecast for the spot price is

$$E_t^*[S(T)] = S(t)e^{r\tau}. \qquad (26)$$

The derivation for expression (26) is as follows. From expression (25), we know

$$S(T) = S(t)\exp[Y(T)].$$

Taking the expected value of the above we have

$$E_t^*[S(T)] = E_t^*[S(t)\exp[Y(T)]]. \qquad (27)$$

It is shown in chapter three on page 28, that if the variable *Y(T)* is normal, then the spot price is log-normal with a mean equal to

$$E_t^*[S(T)] = S(t)\exp\left[E_t^*[Y(T)] + \frac{1}{2}V_t^*[Y(T)]\right]. \qquad (28)$$

Substituting the mean and variance of *Y(T)* into the above yields

$$E_t^*[S(T)] = S(t)\exp\left[\left(r - \frac{1}{2}\sigma^2\right)\tau + \frac{1}{2}\sigma^2\tau\right],$$

$$E_t^*[S(T)] = S(t)e^{r\tau}. \qquad (29)$$

This is agrees with expression (26).

Under an equivalent martingale measure, the price of a European call option written on the risk-adjusted spot price can be expressed as

$$C(S,t) = e^{-r(T-t)} E_t^*[C(S,T)], \tag{30}$$

where $C(S,t)$ is the price of the call option today, and $C(S,T)$ is the call option at a terminal date T. Letting $t = 0$ and substituting in for the risk-adjusted terminal call price we get

$$\begin{aligned}
C(S,0) &= e^{-rT} E_0^*[\max(S(T) - X, 0)], \\
&= e^{-rT} E_0^*\left[\max\left(S(0)e^{Y(T)} - X, 0\right)\right] \\
&= e^{-rT}\left\{E_0^*\left[S(0)e^{Y(T)} - X \mid S(T) \geq X\right] + E_0^*[0 \mid S(T) < X]\right\} \\
&= e^{-rT}\left\{E_0^*\left[S(0)e^{Y(T)} - X \mid S(0)e^{Y(T)} \geq X\right] + E_0^*\left[0 \mid S(0)e^{Y(T)} < X\right]\right\} \\
&= e^{-rT}\left\{E_0^*\left[S(0)e^{Y(T)} - X \mid e^{Y(T)} \geq \frac{X}{S(0)}\right] + E_0^*\left[0 \mid e^{Y(T)} < \frac{X}{S(0)}\right]\right\} \\
&= e^{-rT}\left\{E_0^*\left[S(0)e^{Y(T)} - X \mid Y(T) \geq \ln\left(\frac{X}{S(0)}\right)\right] + E_0^*\left[0 \mid Y(T) < \ln\left(\frac{X}{S(0)}\right)\right]\right\}
\end{aligned} \tag{31}$$

Expression (31) shows the price of the call option to be a linear combination of partial expectations. The first partial expectation operator in expression (31) considers the probability that the spot price is greater than the exercise price at maturity. Formally, this is

$$P^*\{X \leq S(T) \leq \infty\},$$

where P^* is an equivalent martingale probability. Transforming the variable then constitutes the following

$$P^*\left\{\frac{X}{S(0)} \leq \frac{S(T)}{S(0)} \leq \infty\right\},$$

$$P^*\left\{\ln\left(\frac{X}{S(0)}\right)\le \ln\left(\frac{S(T)}{S(0)}\right)\le\infty\right\}.$$

Since $\ln\left(\frac{S(T)}{S(0)}\right)=Y(T)$, the above becomes

$$P^*\left\{\ln\left(\frac{X}{S(0)}\right)\le Y(T)\le\infty\right\}. \tag{32a}$$

The second partial expectation operator in expression (31) considers the probability that the spot price is less than the exercise price at maturity. Formally, this is

$$P^*\{-\infty\le S(T)\le X\},$$

where P^* is an equivalent martingale probability. Again, transforming the variable then constitutes the following

$$P^*\left\{-\infty\le\frac{S(T)}{S(0)}\le\frac{X}{S(0)}\right\},$$

$$P^*\left\{-\infty\le \ln\left(\frac{S(T)}{S(0)}\right)\le \ln\left(\frac{X}{S(0)}\right)\right\},$$

$$P^*\left\{-\infty\le Y(T)\le \ln\left(\frac{X}{S(0)}\right)\right\}. \tag{32b}$$

Given the probability statements in expressions (32a) and (32b), we can rewrite the linear combination of two partial expectation operators in expression (31) as

$$C(S,0)=e^{-rT}\int_{-\infty}^{\ln\left(\frac{X}{S(0)}\right)}(0)dP^*+e^{-rT}\int_{\ln\left(\frac{X}{S(0)}\right)}^{\infty}\left(S(0)e^{y(T)}-X\right)dP^* \tag{33}$$

where dP^* is the risk-neutral probability measure and is equal to

$$dP^{*} = \frac{1}{\sqrt{2\pi\sigma^2 T}} \exp\left[-\frac{\left(y(T) - \left(\left(r - \frac{1}{2}\sigma^2 \right) T \right) \right)^2}{2\sigma^2 T} \right] dy(T). \quad (34)$$

Equation (33) expresses the option payoff as the discounted expected value of the terminal option price. The expectation in expression (33), however, is given with nontraditional notation. Typically, when one takes the expected value of a random variable Z they express this quantity as

$$E[Z] = \int_{-\infty}^{\infty} z f(z) dz, \quad (35)$$

where $f(z)$ is the probability density function of the random variable Z. This is not what we have in expression (33). Expression (33) is equivalent to saying

$$E[Z] = \int_{-\infty}^{\infty} z dP. \quad (36)$$

The question is what is dP? Consider a normally distributed random variable Z at a fixed time t with a mean of zero and unit variance. Formally,

$$Z \sim N(0,1).$$

The probability density function $f(z)$ of this random variable is given by

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}. \quad (37)$$

We are interested in the probability when Z falls near a specific value $\bar{z}$. Since the normal distribution is a continuous distribution the probability of Z taking on a specific value is zero. Therefore, if we want to find the likelihood of witnessing a specific value for Z, we need to choose a small interval, Δ, around a value $\bar{z}$ and then calculate the integral of the density function over this region. This is

$$P\left(\bar{z}-\frac{1}{2}\Delta < Z(=\bar{z}) < \bar{z}+\frac{1}{2}\Delta\right) = \int_{\bar{z}-\frac{1}{2}\Delta}^{\bar{z}+\frac{1}{2}\Delta} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} dz. \qquad (38)$$

Now if the region around $\bar{z}$ is small then *f*(*z*) will not change very much as Z varies from $\bar{z}-\frac{1}{2}\Delta$ to $\bar{z}+\frac{1}{2}\Delta$. This means we can approximate *f*(*z*) by $f(\bar{z})$ during this small interval and write the integral above as

$$\int_{\bar{z}-\frac{1}{2}\Delta}^{\bar{z}+\frac{1}{2}\Delta} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} dz \cong \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\bar{z}^2} \int_{\bar{z}-\frac{1}{2}\Delta}^{\bar{z}+\frac{1}{2}\Delta} dz$$

$$\cong \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\bar{z}^2} \Delta. \qquad (39)$$

The probability above is a mass and described this way corresponds to a measure that is associated with possible values of *Z* in small intervals. For an infinitesimal interval, which we write Δ as *dz*, these measures are denoted by the symbol *dP*. Given the distributional characteristics of the random variable Z, the probability measure is denoted as

$$dP = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(z)^2} dz.$$

We have a normally distributed variable, Z, with a probability measure dP.

The probability measure above is for the actual probability distribution of *Z*. The question remains, how do we transform the probability distribution of Z to alter its mean? Consider the function

$$\xi(z) = e^{z\mu - \frac{1}{2}\mu^2}.$$

If we multiply *dP* by $\xi(z)$, we obtain

$$dP^{*} = dP\xi(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^{2} + z\mu - \frac{1}{2}\mu^{2}} dz\,,$$

$$dP^{*} = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(z-\mu)^{2}} dz\,.$$

Integrating over expression above yields

$$\int_{-\infty}^{\infty} dP^{*} = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(z-\mu)^{2}} dz = 1\,.$$

From the integral above, we see that dP^{*} is also a probability measure. It turns out that by multiplying dP by $\xi(z)$, and then switching to P^{*}, we succeeded in changing the mean of *Z*. Note, that in this particular case, the multiplication by $\xi(z)$ preserved the shape of the probability measure. In fact, dP^{*} is still a bell shaped, Gaussian curve with the same variance as dP, but dP and dP^{*} are different measures. They have different means and they assign different weights to intervals on the *z*-axis.

Returning to expression (34), we see that this expression is the probability measure for the variable *Y(T)*. This measure is risk adjusted. This is understood given the mean for *Y(T)*. That is, under the synthetic probability measure the mean for *Y(T)* is $\left(r - \frac{1}{2}\sigma^{2}\right)\tau$, where as the actual mean of *Y(T)* given the true probability distribution is $\left(\mu - \frac{1}{2}\sigma^{2}\right)\tau$.

Continuing with expression (33), we substitute into the risk adjusted probability measure and obtain

$$C(S,0) = e^{-rT} \int_{-\infty}^{\ln\left(\frac{X}{S(0)}\right)} (0) \frac{1}{\sqrt{2\pi\sigma^2 T}} \exp\left[-\frac{\left(y(T) - \left(\left(r - \frac{1}{2}\sigma^2\right)T\right)\right)^2}{2\sigma^2 T} \right] dy(T)$$

$$+ e^{-rT} \int_{\ln\left(\frac{X}{S(0)}\right)}^{\infty} \frac{\left(S(0)e^{y(T)} - X\right)}{\sqrt{2\pi\sigma^2 T}} \exp\left[-\frac{\left(y(T) - \left(\left(r - \frac{1}{2}\sigma^2\right)T\right)\right)^2}{2\sigma^2 T} \right] dy(T). \quad (40)$$

The first term in expression (40) is equal to zero leaving us with

$$C(S,0) = e^{-rT} \int_{\ln\left(\frac{X}{S(0)}\right)}^{\infty} \left(S(0)e^{y(T)} - X\right) \frac{1}{\sqrt{2\pi\sigma^2 T}} \exp\left[-\frac{\left(y(T) - \left(\left(r - \frac{1}{2}\sigma^2\right)T\right)\right)^2}{2\sigma^2 T} \right] dy(T).$$

(41)

Expression (41) can be split up into two different integrals. For ease of exposition, the brackets of the exponential in expression (41) are acknowledged with a dot

$$\int_{\ln\left(\frac{X}{S(0)}\right)}^{\infty} e^{-rT} S(0)e^{y(T)} \frac{1}{\sqrt{2\pi\sigma^2 T}} \exp[\cdot]dy(T) - \int_{\ln\left(\frac{X}{S(0)}\right)}^{\infty} e^{-rT} X \frac{1}{\sqrt{2\pi\sigma^2 T}} \exp[\cdot]dy(T). \quad (42)$$

We start with the second integral in expression (42). Consider the transformation of $Y(T)$

$$Z = \frac{Y(T) - \left(r - \frac{1}{2}\sigma^2\right)T}{\sigma\sqrt{T}}. \quad (43)$$

If we wish to write the integral in (41) as a function of Z, we must adjust the limits. Recall that the likelihood the option will be in the money is given as

$$P^{*}\left\{\ln\left(\frac{X}{S(0)}\right) \le Y(T) \le \infty\right\}.$$

If we subtract the risk adjusted mean of Y and divide by the standard deviation this yields

$$P^{*}\left\{\frac{\ln\left(\frac{X}{S(0)}\right)-\left(r-\frac{1}{2}\sigma^{2}\right)T}{\sigma\sqrt{T}} \le \frac{Y(T)-\left(r-\frac{1}{2}\sigma^{2}\right)T}{\sigma\sqrt{T}} \le \infty\right\}.$$

From equation (43) we have

$$P^{*}\left\{\frac{\ln\left(\frac{X}{S(0)}\right)-\left(r-\frac{1}{2}\sigma^{2}\right)T}{\sigma\sqrt{T}} \le Z \le \infty\right\}. \tag{44}$$

If we change the variable of integration in the above to dz, then the limits of integration for this problem become $\frac{\ln\left(\frac{X}{S(0)}\right)-\left(r-\frac{1}{2}\sigma^{2}\right)T}{\sigma\sqrt{T}}$ and ∞, making the integral

$$-e^{-rT}X \int_{\frac{\ln\left(\frac{X}{S(0)}\right)-\left(r-\frac{1}{2}\sigma^{2}\right)T}{\sigma\sqrt{T}}}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{1}{2}z^{2}\right] dz. \tag{45}$$

Expression (45) is the risk-adjusted partial expectation for the second term in expression (42). We know this expectation is risk adjusted by inspection of the lower limit. Here we see the mean value for the log return is $\left(r-\frac{1}{2}\sigma^{2}\right)T$. The actual mean value is $\left(\mu-\frac{1}{2}\sigma^{2}\right)T$.

The normal distribution is a symmetrical distribution. A property any normal distribution states that we can write

$$\int_{L}^{\infty} f(z)dz = \int_{-\infty}^{-L} f(z)dz. \tag{46}$$

Using the above property, expression (45) may be rewritten as

$$-e^{-rT}X \int_{-\infty}^{\frac{-\left(\ln\left(\frac{X}{S(0)}\right)-\left(r-\frac{1}{2}\sigma^2\right)T\right)}{\sigma\sqrt{T}}} \frac{1}{\sqrt{2\pi}}\exp\left[-\frac{1}{2}z^2\right]dz.$$

Consequently, the upper limit in expression (45) can be rewritten. Rewriting yields

$$-\frac{\ln\left(\frac{X}{S(0)}\right)-\left(r-\frac{1}{2}\sigma^2\right)T}{\sigma\sqrt{T}},$$

$$-\frac{\ln(X)-\ln S(0)-\left(r-\frac{1}{2}\sigma^2\right)T}{\sigma\sqrt{T}},$$

$$\frac{\ln S(0)-\ln(X)+\left(r-\frac{1}{2}\sigma^2\right)T}{\sigma\sqrt{T}},$$

$$\frac{\ln\left(\frac{S(0)}{X}\right)+\left(r-\frac{1}{2}\sigma^2\right)T}{\sigma\sqrt{T}}. \tag{47}$$

We define the above result as

$$d_2 = \frac{\ln\left(\frac{S(0)}{X}\right)+\left(r-\frac{1}{2}\sigma^2\right)T}{\sigma\sqrt{T}} \tag{48}$$

Thus, equation (31) can be expressed as

$$-e^{-rT}X\int_{-\infty}^{d_2}\frac{1}{\sqrt{2\pi}}\exp\left[-\frac{1}{2}z^2\right]dz. \tag{49}$$

Furthermore, expression (45) becomes

$$-e^{-rT}XN(d_2) \tag{50}$$

We now turn to the first integral in expression (41), which is

$$\int_{\ln\left(\frac{X}{S(0)}\right)}^{\infty} e^{-rT}S(0)e^{y(T)}\frac{1}{\sqrt{2\pi\sigma^2 T}}\exp[\cdot]dy(T).$$

We know from expression (43) that

$$Z = \frac{Y(T)-\left(r-\frac{1}{2}\sigma^2\right)T}{\sigma\sqrt{T}}.$$

The integral above may be written in terms of the variable Z. The limits for the integral in terms of Z are $\frac{\ln\left(\frac{X}{S(0)}\right)-\left(r-\frac{1}{2}\sigma^2\right)T}{\sigma\sqrt{T}}$ and ∞. Substituting the limits into the integral above and changing the probability density function we have

$$\int_{\frac{\ln\left(\frac{X}{S(0)}\right)-\left(r-\frac{1}{2}\sigma^2\right)T}{\sigma\sqrt{T}}}^{\infty} e^{-rT}S(0)e^{y(T)}\frac{1}{\sqrt{2\pi}}\exp\left[-\frac{1}{2}z^2\right]dz.$$

The limits for the integral above is written in terms of the variable Z, however, the variable $Y(T)$ is still in the integral. We need to standardize this variable. First we recognize that the lower limit in the expression above is equal to

$$-d_2 = \frac{\ln\left(\frac{X}{S(0)}\right)-\left(r-\frac{1}{2}\sigma^2\right)T}{\sigma\sqrt{T}}, \tag{51}$$

Therefore, we can express the integral as

$$\int_{-d_2}^{\infty} e^{-rT}S(0)e^{y(T)}\frac{1}{\sqrt{2\pi}}\exp\left[-\frac{1}{2}z^2\right]dz. \tag{52}$$

Now working on standardizing the variable $Y(T)$ we multiply the above expression by 1,

$$\exp\left(\left(r-\frac{1}{2}\sigma^2\right)T\right)\exp\left(-\left(r-\frac{1}{2}\sigma^2\right)T\right),$$

which yields

$$e^{-rT}e^{\left(\left(r-\frac{1}{2}\sigma^2\right)T\right)}S(0)\int_{-d_2}^{\infty} e^{y(T)}e^{\left(-\left(r-\frac{1}{2}\sigma^2\right)T\right)}\frac{1}{\sqrt{2\pi}}\exp\left[-\frac{1}{2}z^2\right]dz.$$

This is equal to

$$e^{-rT}e^{\left(\left(r-\frac{1}{2}\sigma^2\right)T\right)}S(0)\int_{-d_2}^{\infty} e^{\left(y(T)-\left(r-\frac{1}{2}\sigma^2\right)T\right)}\frac{1}{\sqrt{2\pi}}\exp\left[-\frac{1}{2}z^2\right]dz. \tag{53}$$

We know

$$Z=\frac{Y(T)-\left(r-\frac{1}{2}\sigma^2\right)T}{\sigma\sqrt{T}},$$

which is equal to

$$Z\sigma\sqrt{T}=Y(T)-\left(r-\frac{1}{2}\sigma^2\right)T. \tag{54}$$

Substituting (54) into (53) yields

$$e^{-rT}e^{\left(r-\frac{1}{2}\sigma^2\right)T}S(0)\int_{-d_2}^{\infty} e^{z\sigma\sqrt{T}}\frac{1}{\sqrt{2\pi}}\exp\left[-\frac{1}{2}z^2\right]dz. \tag{55}$$

Working on the term inside the integral in equation (55), we combine the exponents. This yields

$$e^{-rT}e^{\left(r-\frac{1}{2}\sigma^2\right)T}S(0)\int_{-d_2}^{\infty}\frac{1}{\sqrt{2\pi}}e^{z\sigma\sqrt{T}}\exp\left[-\frac{1}{2}z^2\right]dz,$$

$$e^{-rT}e^{\left(r-\frac{1}{2}\sigma^2\right)T}S(0)\int_{-d_2}^{\infty}\frac{1}{\sqrt{2\pi}}e^{\frac{1}{2}(2)z\sigma\sqrt{T}}\exp\left[-\frac{1}{2}z^2\right]dz,$$

$$e^{-rT}e^{\left(r-\frac{1}{2}\sigma^2\right)T}S(0)\int_{-d_2}^{\infty}\frac{1}{\sqrt{2\pi}}\exp\left[-\frac{1}{2}z^2+\frac{1}{2}(2)z\sigma\sqrt{T}\right]dz,$$

$$e^{-rT}e^{\left(r-\frac{1}{2}\sigma^2\right)T}S(0)\int_{-d_2}^{\infty}\frac{1}{\sqrt{2\pi}}\exp\left[-\frac{1}{2}\left(z^2-2z\sigma\sqrt{T}\right)\right]dz. \qquad (56)$$

The exponential term inside the integral in expression (56) can be completed into a square. To do so we need to multiply (56) by

$$\exp\left[\frac{\sigma^2T}{2}\right]\exp\left[-\frac{\sigma^2T}{2}\right]. \qquad (57)$$

This yields

$$e^{\frac{\sigma^2T}{2}}e^{-\frac{\sigma^2T}{2}}e^{-rT}e^{\left(r-\frac{1}{2}\sigma^2\right)T}S(0)\int_{-d_2}^{\infty}\frac{1}{\sqrt{2\pi}}\exp\left[-\frac{1}{2}\left(z^2-2z\sigma\sqrt{T}\right)\right]dz,$$

$$e^{-rT}e^{\frac{\sigma^2T}{2}}e^{\left(r-\frac{1}{2}\sigma^2\right)T}S(0)\int_{-d_2}^{\infty}e^{\frac{-\sigma^2T}{2}}\frac{1}{\sqrt{2\pi}}\exp\left[-\frac{1}{2}\left(z^2-2z\sigma\sqrt{T}\right)\right]dz,$$

$$e^{-rT}e^{\frac{\sigma^2T}{2}}e^{\left(r-\frac{1}{2}\sigma^2\right)T}S(0)\int_{-d_2}^{\infty}\frac{1}{\sqrt{2\pi}}\exp\left[-\frac{1}{2}\sigma^2T-\frac{1}{2}\left(z^2-2z\sigma\sqrt{T}\right)\right]dz,$$

$$e^{-rT}S(0)e^{\frac{\sigma^2T}{2}+\left(r-\frac{1}{2}\sigma^2\right)T}\int_{-d_2}^{\infty}\frac{1}{\sqrt{2\pi}}\exp\left[-\frac{1}{2}\left(z^2-2z\sigma\sqrt{T}+\sigma^2T\right)\right]dz,$$

or

$$e^{-rT}S(0)e^{rT}\int_{-d_2}^{\infty}\frac{1}{\sqrt{2\pi}}\exp\left[-\frac{1}{2}\left(z-\sigma\sqrt{T}\right)^2\right]dz. \qquad (58)$$

Examining expression (58), we see that the integral is no longer for a standard normal variable. The mean has been moved to the left by $\sigma\sqrt{T}$ units. If we let $W=Z-\sigma\sqrt{T}$, then we see that $\frac{dw}{dz}=1$. In addition, when $z=\infty$ then $w=\infty$, and when

$$z = \frac{\ln\left(\frac{X}{S(0)}\right) - \left(r - \frac{1}{2}\sigma^2\right)T}{\sigma\sqrt{T}}$$

then

$$w = \frac{\ln\left(\frac{X}{S(0)}\right) - \left(r - \frac{1}{2}\sigma^2\right)T}{\sigma\sqrt{T}} - \sigma\sqrt{T}.$$

Therefore, using *dw* as the variable of integration makes expression (58) equal to

$$e^{-rT} S(0) e^{rT} \int_{-d_2-\sigma\sqrt{T}}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{1}{2}w^2\right] dw,$$

where the probability density function is now for a standard normal variable. Using the symmetry property of the normal distribution allows us to write the above as

$$e^{-rT} S(0) e^{rT} \int_{-\infty}^{d_2+\sigma\sqrt{T}} \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{1}{2}w^2\right] dw,$$

$$e^{-rT} S(0) e^{rT} \int_{-\infty}^{d_1} \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{1}{2}w^2\right] dw. \tag{59}$$

The upper limit in expression (59) is equal to

$$d_1 = d_2 + \sigma\sqrt{T},$$

$$d_1 = \frac{\ln\left(\frac{S(0)}{X}\right) + \left(r - \frac{1}{2}\sigma^2 T\right)}{\sigma\sqrt{T}} + \sigma\sqrt{T},$$

$$d_1 = \frac{\ln\left(\frac{S(0)}{X}\right) + \left(r - \frac{1}{2}\sigma^2 T\right)}{\sigma\sqrt{T}} + \frac{\sigma^2 T}{\sigma\sqrt{T}},$$

$$d_1 = \frac{\ln\left(\frac{S(0)}{X}\right) + \left(r + \frac{1}{2}\sigma^2 T\right)}{\sigma\sqrt{T}}. \tag{60}$$

Given the above, expression (59) reduces to

$$e^{-rT}S(0)e^{rT}N(d_1),$$

$$S(0)N(d_1). \tag{61}$$

Now replacing the integrals in expression (41), with expressions (50) and (61) yields

$$C(S,0) = S(0)N(d_1) - e^{-rT}XN(d_2), \tag{62}$$

where

$$d_1 = \frac{\ln\left(\frac{S(0)}{X}\right) + \left(r + \frac{1}{2}\sigma^2 T\right)}{\sigma\sqrt{T}}, \tag{63}$$

and

$$d_2 = d_1 - \sigma\sqrt{T}. \tag{64}$$

Expressions (62), (63), and (64) are identical to the Black-Scholes option pricing model. The unique result of the analysis is that the derivation is conducted under a partial expectation model as opposed to the standard arbitrage methodology. Notice, during the derivation there is no explicit use of an arbitrage formula and no partial differential equation needs to be solved. This result is not a complete surprise in that we know from the Feynman-Kac theorem that a correspondence exists between a certain class of conditional expectations and a set of partial differential equations.

4.2 One-factor model

In order to price a European call option written on a commodity futures contract, we must first understand the dynamics of the futures price. Recall from chapter three, in the one-factor model we use a mean reverting spot price to price a futures contract. The spot price process is given by expression (23) and is

$$\frac{dS}{S} = k(\mu - \ln S)dt + \sigma dZ_s .$$

This process may be altered by applying the Girsanov theorem. The risk adjusted process becomes

$$\frac{dS}{S} = k(\mu - \ln S)dt + \sigma\left(dZ_s^* - \lambda dt\right),$$

$$= k(\mu - \sigma\lambda - \ln S)dt + \sigma dZ_s^*,$$

$$= k\left(\alpha^* - \ln S\right)dt + \sigma dZ_s^*, \qquad (65)$$

where $\alpha^* = \mu - \sigma\lambda$. To determine the process followed by a futures price under a mean reverting spot price, we assume the futures price is a twice continuously differentiable function with respect to *S* and time *t*. Next, we use Ito's lemma and the risk-adjusted differential spot price to determine the futures price differential. This is,

$$dF = F_s dS + \frac{1}{2} F_{ss} dS^2 + F_t dt \quad , \qquad (66)$$

where F_s F_{ss} and F_t are the partial derivatives of the futures price with respect to the spot price and time. Substituting in for dS and $(dS)^2$, equation (66) becomes

$$dF = F_s\left[k\left(\alpha^* - \ln S\right)Sdt + \sigma S dZ_s^*\right] + \frac{1}{2} F_{ss}\left[\sigma^2 S^2 dt\right] + F_t dt \quad . \qquad (67)$$

Rearranging the terms in equation (67) yields

$$dF = \left[F_s\left(k\left(\alpha^* - \ln S\right)S\right) + \frac{1}{2} F_{ss} S^2 \sigma^2 + F_t\right]dt + F_s S \sigma dZ_s^* . \qquad (68)$$

Expression (68) is the risk-neutral dynamics for the futures price. If there are no arbitrage opportunities, the equivalent martingale measure converts the futures price dynamic into a martingale. Implicitly, this no arbitrage constraint implies that the instantaneous drift in expression (68) equals zero. Formally this is

$$F_s\left(k\left(\alpha^* - \ln S\right)S\right) + \frac{1}{2}F_{ss}S^2\sigma^2 + F_t = 0. \qquad (69)$$

Closer inspection of the equation above shows that this expression is equal to the fundamental partial differential equation used in the arbitrage free method of Black-Scholes.[4] This is not a surprise. A no arbitrage condition in the Black-Scholes analysis explicitly constrains the expected movements in the futures price to equal zero. Under a risk neutral probability measure, the expected movements in the futures price is implicitly constrained by a no arbitrage condition to equal zero. Since both models are pricing the same asset, we should expect the two constraints to be identical, and they are. Furthermore, the Feynman-Kac theorem shows us a correspondence between a class of condition expectations and a set of partial differential equations.

Given the drift term in expression (68) is zero we have[5]

$$dF = F_s S\sigma dZ_s^*. \qquad (70)$$

Now recall from expression (40) in chapter three, the solution to the one-factor futures model is

$$F\left[S(T),\tau\right] = \exp\left[e^{-k\tau}\ln S(t) + \alpha^*\left(1 - e^{-k\tau}\right) + \frac{\sigma^2}{4k}\left[1 - e^{-2k\tau}\right]\right],$$

$$= S(t)^{e^{-k\tau}}\exp\left[\alpha^*\left(1 - e^{-k\tau}\right) + \frac{\sigma^2}{4k}\left[1 - e^{-2k\tau}\right]\right]. \qquad (71)$$

Expression (71) is the solution to the martingale process stated in expression (70).

Taking the partial derivative of $F\left[S(T),T\right]$ with respect S equals

$$F_s = e^{-k\tau} S(t)^{e^{-k\tau}-1} \exp\left[\alpha^*\left(1-e^{-k\tau}\right)+\frac{\sigma^2}{4k}\left[1-e^{-2k\tau}\right]\right] \quad . \tag{72}$$

Substituting expression (72) into equation (70) yields

$$dF = e^{-k\tau} S(t)^{e^{-k\tau}-1} \exp\left[\alpha^*\left(1-e^{-k\tau}\right)+\frac{\sigma^2}{4k}\left[1-e^{-2k\tau}\right]\right] S\sigma dZ_s^*,$$

$$= e^{-k\tau} S(t)^{e^{-k\tau}} \exp\left[\alpha^*\left(1-e^{-k\tau}\right)+\frac{\sigma^2}{4k}\left[1-e^{-2k\tau}\right]\right] \sigma dZ_s^*,$$

$$= Fe^{-k\tau}\sigma dZ_s^*. \tag{73}$$

Let us define a parameter σ_F such that

$$\sigma_F dZ_F^* = e^{-k\tau}\sigma dZ_s^*. \tag{74}$$

Note, that the Weiner processes in expression (74) are identical. That is, $dZ_F^* \equiv dZ_s^*$. This is true since the only source of uncertainty in the futures price comes from the spot price. The distinction is made here to correlate the innovation term with the dynamics of the futures price and not the spot price.

Equation (74) represents the term structure of the volatility for a commodity's futures price. In general, expression (74) states that the futures volatility is the dampened spot price volatility over the investment horizon of the contract.[6] The degree to which the spot price volatility is diminished depends on the time to maturity and the spot price's reversion parameter. Should this be expected? Consider Figure 3 on page 289 in chapter five. This figure shows the behavior of the one-factor futures price at different times to maturity and levels of mean reversion, when the spot price is below its long run mean. The figure shows that as the time to maturity increases the difference between the futures price and the spot price's long run mean diminishes. The reason, given a relatively long time horizon, the spot price has the time to revert to its mean. Since the futures price is equal to the risk-adjusted expected spot price, any turbulence in the spot price today should not significantly alter this value. Thus, implying the volatility of spot price movement should be greater than the volatility of the futures price. Furthermore, if the propensity for the spot price to revert to its mean increases, the spot price trends to its mean faster. Therefore, one should naturally expect that as the speed of reversion increases, the futures price should become less volatile as well. In a final note, at maturity the futures price volatility equals σ, the spot price volatility. This is true, since by definition the futures price equals the spot price when $\tau = 0$.[7]

Given equation (74), expression (73) simplifies to

$$dF = F\sigma_F dZ_F^*. \tag{75}$$

Equation (75) is the risk-neutral return dynamics for a commodity's futures price. The stochastic differential is shown to follow a geometric Brownian with no drift. Consequently, the futures price is a martingale process, and the best forecast for the terminal futures price is today's futures price.

We have the dynamics for the futures price process and now want to consider valuing options written on these futures contracts. Expression (75) shows that the futures price follows a geometric Brownian motion. Upon casual observation, one should expect that we may invoke the Black-Scholes model to price these options. There is one fundamental difference. In the Black-Scholes model, the options are priced on underlying spot price process, which is given exogenously. In our analysis, however, the futures price dynamic is constructed endogenously. Given this complication, we do note that the futures price dynamics in expression (75) are free from the level of the spot price.

In addition, the futures price dynamics are explicitly only a function of the spot price volatility, speed of adjustment parameter and time to maturity. This is the case because

$$\sigma_F dZ_F^* = e^{-k\tau}\sigma dZ_s^*.$$

Now, these parameters are deterministic as well as given exogenously. Therefore, for the purpose of pricing options on futures, we may treat the futures price dynamics as an exogenous process and apply the Black-Scholes methodology to price the options.

From expression (75), we may summarize the futures price as follows. Let $H = \ln F$, where H is twice differentiable in F. With this transformation and Ito's lemma, the increment of H is expressed as

$$dH = H_F dF + \frac{1}{2} H_{FF} dF^2$$

$$= \frac{1}{F}\left(Fe^{-k\tau}\sigma dZ_s^*\right) - \frac{1}{2F^2}\left(F^2 e^{-2k\tau}\sigma^2 dt\right),$$

$$= -\frac{1}{2}e^{-2k\tau}\sigma^2 dt + e^{-k\tau}\sigma dZ_s^*. \tag{76}$$

Integrating over equation (76) yields

$$\int_t^T dH = \frac{1}{2}\int_t^T - e^{-2k(T-s)}\sigma^2 ds + \int_t^T e^{-k(T-s)}\sigma dZ_s^*,$$

$$H(T) - H(t) = -\frac{1}{4k}\left[e^{-2k(T-T)}\sigma^2 - e^{-2k(T-t)}\sigma^2\right] + \sigma\int_t^T e^{-k(T-s)} dZ_s^*,$$

$$H(T) = H(t) - \frac{1}{4k}\left[\sigma^2 - e^{-2k(T-t)}\sigma^2\right] + \sigma\int_t^T e^{-k(T-s)} dZ_s^*,$$

$$H(T) = H(t) - \frac{1}{4k}\sigma^2\left[1 - e^{-2k(T-t)}\right] + \sigma\int_t^T e^{-k(T-s)} dZ_s^*,$$

$$F(T) = F(t)\exp\left[-\frac{1}{4k}\sigma^2\left[1-e^{-2k(T-t)}\right]+\sigma\int_t^T e^{-k(T-s)}dZ_s^*\right],$$

$$F(T) = F(t)\exp[Y(T)], \tag{77}$$

where $Y(T) = -\frac{1}{4k}\sigma^2\left[1-e^{-2k(T-t)}\right]+\sigma\int_t^T e^{-k(T-s)}dZ_s^*$. Expression (77) is the alternative expression for the dynamic expressed in equation (75), and thereby, honors the constraint in expression (69). The expected value of $F(T)$ is given as

$$E_t^*[F(T)] = F(t)\exp\left[E_t^*[Y(T)]+\frac{1}{2}V_t^*[Y(T)]\right]. \tag{78}$$

The expected value of $F(T)$ is a function of the mean and variance of $Y(T)$. Therefore, we need to evaluate the both the mean and the variance of $Y(T)$. The expected value of $Y(T)$ is

$$E_t^*[Y(T)] = -\frac{1}{4k}\sigma^2\left[1-e^{-2k(T-t)}\right]. \tag{79}$$

The variance of $Y(T)$ is

$$V_t^*[Y(T)] = \sigma^2\int_t^T e^{-2k(T-s)}ds,$$

$$= \frac{\sigma^2}{2k}\left[e^{-2k(T-T)} - e^{-2k(T-t)}\right],$$

$$= \frac{\sigma^2}{2k}\left[1-e^{-2k(T-t)}\right]. \tag{80}$$

Substituting equations (79) and (80) into expression (78) yields

$$E_t^*[F(T)] = F(t)\exp\left[-\frac{\sigma^2}{4k}\left[1-e^{-2k(T-t)}\right]+\frac{\sigma^2}{4k}\left[1-e^{-2k(T-t)}\right]\right],$$

$$= F(t)\exp[0].$$

$$= F(t). \tag{81}$$

We can see that the expected value of the terminal futures price, $F(T)$, is simply today's futures price, as we expected.[8]

Before we start to price the options we need to make one final note. Options on futures often expire prior to the futures contract expiration. For energy contracts on the New York Mercantile Exchange, the futures options expire three business days prior to the futures expiration date. This has important implications for pricing the options. Recall in the Black-Scholes model, volatility is a parameter we need to value the options. If the futures option expires prior to the futures contract, then the volatility of the futures contract over the life of the futures option is to be used when pricing the option. The futures price volatility is

$$V_t[Y(T)] = \frac{\sigma^2}{2k}\left[1 - e^{-2k(T-t)}\right].$$

Let t equal current time, T_1 equal the maturity date of a call option, and T equal the maturity date of a futures contract, with $T_1 \leq T$. The volatility of the futures contract over the maturity of the option contract is

$$V_t^*[Y(T_1)] = \sigma^2 \int_t^{T_1} e^{-2k(T-s)} ds,$$

$$= \frac{\sigma^2}{2k} e^{-2k(T-s)} \Big|_t^{T_1},$$

$$= \frac{\sigma^2}{2k}\left(e^{-2k(T-T_1)} - e^{-2k(T-t)}\right). \tag{82}$$

Now define $V_t^*[Y(T_1)] \equiv v^2$. v^2 is the appropriate volatility for pricing options written on commodity futures. The expected value of $Y(T_1)$ is

$$E_t^*[Y(T_1)] = \frac{1}{2}\int_t^{T_1} -e^{-2k(T-s)}\sigma^2 ds + \sigma\int_t^{T_1} e^{-k(T-s)} E_t^*[dZ_s^*],$$

$$= \frac{1}{2}\left[-\frac{\sigma^2}{2k}\left(e^{-2k(T-T_1)} - e^{-2k(T-t)}\right)\right],$$

$$= -\frac{1}{2}v^2. \tag{83}$$

The expected futures price at the options expiration is

$$E_t[F(T_1)] = F(t)\exp\left[-\frac{1}{2}v^2 + \frac{1}{2}v^2\right],$$

$$= F(t)\exp[0],$$

$$= F(t). \tag{84}$$

We know the futures price is a martingale process and that today's price is the best forecast for the terminal futures price. The difference between expressions (84) and (81) is the time horizon, where $T_1 < T$. Expression (84) states the value for the expected futures price at time T_1, and expression (81) states the value for the expected futures price at time T. Both equations yield the same result. The expected futures price at both time periods is equal to today's futures price. This is not surprising since the futures price follows a martingale process.

Under an equivalent martingale measure the price of a European option written on the risk-adjusted futures contract can be expressed as

$$C(S,t) = e^{-r(T_1-t)}E_t^*[C(S,T_1)], \tag{85}$$

where $C(S,t)$ is the price of the option today, and $C(S,T_1)$ is the risk adjusted option at a terminal date T_1. Substituting in for the risk-adjusted terminal call price we get

$$C(F,t) = e^{-rT_1}E_t^*[\max(F(T_1) - X, 0)],$$

$$= e^{-r\tau_1} E_t^*\left[\max\left(F(t)e^{Y(T_1)} - X, 0\right)\right].$$

$$= e^{-r\tau_1}\left\{E_t^*\left[F(t)e^{Y(T_1)} - X \mid F(T_1) \geq X\right] + E_t^*\left[0 \mid F(T_1) < X\right]\right\}$$

$$= e^{-r\tau_1}\left\{E_t^*\left[F(t)e^{Y(T_1)} - X \mid F(t)e^{Y(T_1)} \geq X\right] + E_t^*\left[0 \mid F(t)e^{Y(T_1)} < X\right]\right\}$$

$$= e^{-r\tau_1}\left\{E_t^*\left[F(t)e^{Y(T_1)} - X \mid e^{Y(T_1)} \geq \frac{X}{F(t)}\right] + E_t^*\left[0 \mid e^{Y(T_1)} < \frac{X}{F(t)}\right]\right\}$$

$$= e^{-r\tau_1}\left\{E_t^*\left[F(t)e^{Y(T_1)} - X \mid Y(T_1) \geq \ln\left(\frac{X}{F(t)}\right)\right] + E_t^*\left[0 \mid Y(T_1) < \ln\left(\frac{X}{F(t)}\right)\right]\right\}$$

(86)

Expression (86) shows the call price is a linear combination of two partial expectations. The first partial expectation in expression (86) considers the probability that the spot price is greater than the exercise price at maturity. Formally, this is

$$P^*\{X \leq F(T_1) \leq \infty\}$$

where P^* is an equivalent martingale measure. Transforming the variable then constitutes the following

$$P^*\left\{\frac{X}{F(t)} \leq \frac{F(T_1)}{F(t)} \leq \infty\right\},$$

$$P^*\left\{\ln\left(\frac{X}{F(t)}\right) \leq \ln\left(\frac{F(T_1)}{F(t)}\right) \leq \infty\right\}.$$

Since $\ln\left(\frac{F(T_1)}{F(t)}\right) = Y(T_1)$, the above expression becomes

$$P^*\left\{\ln\left(\frac{X}{F(t)}\right) \leq Y(T_1) \leq \infty\right\}. \qquad \textbf{(87a)}$$

The second partial expectation in expression (86) considers the probability that the spot price is less than the exercise price at maturity. Formally, this is

$$P^{*}\{\infty \leq F(T_1) \leq X\}.$$

Transforming the variable then constitutes the following

$$P^{*}\left\{-\infty \leq \frac{F(T_1)}{F(t)} \leq \frac{X}{F(t)}\right\},$$

$$P^{*}\left\{-\infty \leq \ln\left(\frac{F(T_1)}{F(t)}\right) \leq \ln\left(\frac{X}{F(t)}\right)\right\}.$$

$$P^{*}\left\{-\infty \leq Y(T_1) \leq \ln\left(\frac{X}{F(t)}\right)\right\}. \tag{87b}$$

Given the statements in expressions (87a) and (87b), we can rewrite the expectation in (86) as

$$C(F,t) = e^{-r\tau_1} \int_{-\infty}^{\ln\left(\frac{X}{F(t)}\right)} (0) dP^{*} + e^{-r\tau_1} \int_{\ln\left(\frac{X}{F(t)}\right)}^{\infty} \left(F(t)e^{y(T_1)} - X\right) dP^{*} \tag{88}$$

where dP^{*} is the risk-adjusted probability measure denoted as

$$dP^{*} = \frac{1}{\sqrt{2\pi v^2}} \exp\left[-\frac{\left(y(T_1) + \frac{1}{2}v^2\right)^2}{2v^2}\right] dy(T_1). \tag{89}$$

τ_1 is the time to maturity for the option contract. Again, note in expression (88), we are using an equivalent probability measure to evaluate the expectation in equation (86) (this is discussed on page 130). The first term in expression (88) is equal to zero leaving us with

$$C(F,t) = e^{-r\tau_1} \int_{\ln\left(\frac{X}{F(t)}\right)}^{\infty} \left(F(t)e^{y(T_1)} - X\right)\frac{1}{\sqrt{2\pi v^2}}\exp\left[-\frac{\left(y(T_1)+\frac{1}{2}v^2\right)^2}{2v^2}\right]dy(T_1). \qquad (90)$$

Expression (90) can be split up into two different integrals. For ease of exposition, the brackets of the exponential in expression (90) are acknowledged with a dot.

$$\int_X^{\infty} e^{-r\tau_1} F(t)e^{y(T_1)}\frac{1}{\sqrt{2\pi v^2}}\exp[\cdot]dy(T_1) - \int_X^{\infty} e^{-r\tau_1} X\frac{1}{\sqrt{2\pi v^2}}\exp[\cdot]dy(T_1) \qquad (91)$$

We start with the second integral in expression (91). Consider the transformation of $Y(T_1)$

$$Z = \frac{Y(T_1)+\frac{1}{2}v^2}{v}. \qquad (92)$$

If we wish to write the integrals in (91) as a function of Z, we must adjust the limits. Now subtract the mean of Y and divide by the standard deviation yields

$$P^*\left\{\frac{\ln\left(\frac{X}{F(t)}\right)+\frac{1}{2}v^2}{v} \leq \frac{Y(T_1)+\frac{1}{2}v^2}{v} \leq \infty\right\}.$$

From equation (92) we have

$$P^*\left\{\frac{\ln\left(\frac{X}{F(t)}\right)+\frac{1}{2}v^2}{v} \leq Z \leq \infty\right\}. \qquad (93)$$

Thus, if we change the variable of integration in the above to dz, then the limits of integration for this problem become $\frac{\ln\left(\frac{X}{F(t)}\right)+\frac{1}{2}v^2}{v}$ and ∞, making the integral

$$-e^{-r\tau_1}X\int_{\frac{\ln\left(\frac{X}{F(t)}\right)+\frac{1}{2}v^2}{v}}^{\infty}\frac{1}{\sqrt{2\pi}}\exp\left[-\frac{1}{2}z^2\right]dz. \tag{94}$$

Recall the symmetry property of the standard normal distribution. This is

$$\int_{L}^{\infty}f(z)dz=\int_{-\infty}^{-L}f(z)dz. \tag{95}$$

Thus, expression (94) is reduced to

$$-e^{-r\tau_1}X\int_{\frac{\ln\left(\frac{X}{F(t)}\right)+\frac{1}{2}v^2}{v}}^{\infty}\frac{1}{\sqrt{2\pi}}\exp\left[-\frac{1}{2}z^2\right]dz$$

$$-e^{-r\tau_1}X\int_{-\infty}^{\frac{-\left(\ln\left(\frac{X}{F(t)}\right)+\frac{1}{2}v^2\right)}{v}}\frac{1}{\sqrt{2\pi}}\exp\left[-\frac{1}{2}z^2\right]dz. \tag{96}$$

The upper limit in expression (96) can be alternatively expressed as

$$-\frac{\ln\left(\frac{X}{F(t)}\right)+\frac{1}{2}v^2}{v},$$

$$\frac{\ln\left(\frac{F(t)}{X}\right)-\frac{1}{2}v^2}{v}. \tag{97}$$

We define the upper limit as

$$d_2=\frac{\ln\left(\frac{F(t)}{X}\right)-\frac{1}{2}v^2}{v}. \tag{98}$$

Thus, expression (96) reduces to

$$-e^{-r\tau_1}X\int_{-\infty}^{d_2}\frac{1}{\sqrt{2\pi}}\exp\left[-\frac{1}{2}z^2\right]dz, \tag{99}$$

which equals

$$-e^{-r\tau_1} XN(d_2)\,. \tag{100}$$

Now working on the first integral in expression (91), we start with

$$e^{-r\tau_1} \int_{\ln\left(\frac{X}{F(t)}\right)}^{\infty} F(t)e^{y(T_1)} \frac{1}{\sqrt{2\pi v^2}} \exp\left[-\frac{\left(y(T_1)+\frac{1}{2}v^2\right)^2}{2v^2}\right] dy(T_1)\,.$$

We may change the variable of integration. Changing the variable the above integral may be rewritten as

$$\int_{\frac{\ln\left(\frac{X}{F(t)}\right)+\frac{1}{2}v^2}{v}}^{\infty} e^{-r\tau_1} F(t)e^{y(T_1)} \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{1}{2}z^2\right] dz\,. \tag{101}$$

The limits for the integral in expression (101) are in terms of Z. We still have $Y(T_1)$ in this expression and must account for this. The lower limit for the above integral is simply $-d_2$. Therefore we my reduce the limit to

$$\int_{-d_2}^{\infty} e^{-r\tau_1} F(t)e^{y(T_1)} \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{1}{2}z^2\right] dz\,. \tag{102}$$

Transforming the variable $Y(T_1)$ we multiply the above expression by 1,

$$\exp\left(-\frac{1}{2}v^2\right)\exp\left(\frac{1}{2}v^2\right), \tag{103}$$

which yields

$$e^{-r\tau_1} e^{\left(-\frac{1}{2}v^2\right)} F(t) \int_{-d_2}^{\infty} e^{y(T_1)} e^{\left(\frac{1}{2}v^2\right)} \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{1}{2}z^2\right] dz\,.$$

This is equal to

$$e^{-r\tau_1} e^{\left(-\frac{1}{2}v^2\right)} F(t) \int_{-d_2}^{\infty} e^{\left(y(T_1)+\frac{1}{2}v^2\right)} \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{1}{2}z^2\right] dz\,. \tag{104}$$

We know

$$Z = \frac{Y(T_1) + \frac{1}{2}v^2}{v},$$

which is equal to

$$Zv = Y(T_1) + \frac{1}{2}v^2. \tag{105}$$

Substituting (105) into (104) yields

$$e^{-r\tau_1} e^{\left(-\frac{1}{2}v^2\right)} F(t) \int_{-d_2}^{\infty} e^{zv} \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{1}{2}z^2\right] dz. \tag{106}$$

Combining the exponents inside the integral above yields

$$e^{-r\tau_1} e^{\left(-\frac{1}{2}v^2\right)} F(t) \int_{-d_2}^{\infty} \frac{1}{\sqrt{2\pi}} e^{zv} \exp\left[-\frac{1}{2}z^2\right] dz,$$

$$e^{-r\tau_1} e^{\left(-\frac{1}{2}v^2\right)} F(t) \int_{-d_2}^{\infty} \frac{1}{\sqrt{2\pi}} e^{\frac{1}{2}(2)zv} \exp\left[-\frac{1}{2}z^2\right] dz,$$

$$e^{-r\tau_1} e^{\left(-\frac{1}{2}v^2\right)} F(t) \int_{-d_2}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{1}{2}z^2 + \frac{1}{2}(2)zv\right] dz,$$

$$e^{-r\tau_1} e^{\left(-\frac{1}{2}v^2\right)} F(t) \int_{-d_2}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(z^2 - 2zv\right)\right] dz. \tag{107}$$

The exponential term inside the integral in expression (107) can be completed into a square. To do so we push

$$\exp\left[-\frac{v^2}{2}\right], \tag{108}$$

inside the integral in expression (107). This yields

$$e^{-r\tau_1} F(t) \int_{-d_2}^{\infty} e^{\left(-\frac{1}{2}v^2\right)} \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(z^2 - 2zv\right)\right] dz,$$

$$e^{-r\tau_1} F(t) \int_{-d_2}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{1}{2}v^2 - \frac{1}{2}\left(z^2 - 2zv\right)\right] dz,$$

$$e^{-r\tau_1} F(t) \int_{-d_2}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(z^2 - 2zv + v^2\right)\right] dz,$$

or

$$e^{-r\tau_1} F(t) \int_{-d_2}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{1}{2}(z - v)^2\right] dz. \tag{109}$$

The integral in expression (109) is no longer for a standard normal variable. We see that then mean of Z has been moved to the left by v units. If we let $W = Z - v$ then we see that $\frac{dw}{dz} = 1$. When $z = \infty$ then $w = \infty$, and when

$$z = \frac{\ln\left(\frac{F(t)}{X}\right) + \frac{1}{2}v^2}{v},$$

then

$$w = \frac{\ln\left(\frac{F(t)}{X}\right) + \frac{1}{2}v^2}{v} - v,$$

$$w = -d_2 - v.$$

Therefore, using dw as the variable of integration makes expression (109) equal to

$$e^{-r\tau_1} F(t) \int_{-d_2 - v}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{1}{2}w^2\right] dw. \tag{110}$$

Using the symmetry property of the normal distribution allows us to write the above as

$$e^{-r\tau_1} F(t) \int_{-\infty}^{d_2 + v} \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{1}{2}w^2\right] dw,$$

$$e^{-r\tau_1}F(t)\int_{-\infty}^{d_1}\frac{1}{\sqrt{2\pi}}\exp\left[-\frac{1}{2}w^2\right]dw. \tag{111}$$

The upper limit is equal to

$$\begin{aligned} d_1 &= d_2 + v, \\ &= \frac{\ln\left(\frac{F(t)}{X}\right)-\frac{1}{2}v^2}{v}+v, \\ &= \frac{\ln\left(\frac{F(t)}{X}\right)-\frac{1}{2}v^2}{v}+\frac{v^2}{v}, \\ &= \frac{\ln\left(\frac{F(t)}{X}\right)+\frac{1}{2}v^2}{v}. \end{aligned} \tag{112}$$

Given the above, expression (111) reduces to

$$e^{-r\tau_1}F(t)N(d_1). \tag{113}$$

Now replacing the integrals in expression (91), with expressions (100) and (113) yields

$$C(F,t) = e^{-r\tau_1}F(t)N(d_1) - e^{-r\tau_1}XN(d_2), \tag{114}$$

where

$$d_1 = \frac{\ln\left(\frac{F(t)}{X}\right)+\frac{1}{2}v^2}{v}, \tag{115}$$

and

$$d_2 = d_1 - v. \tag{116}$$

Equation (116) is the price of an option written on a commodity futures contract, whose price is governed by the process stated in expression (75).

4.3 Two-factor model

In the above analysis, we were concerned with pricing an option contract written on a futures contract, whose price was entirely summarized by a mean reverting spot price. Our next step is to price options on futures, where the futures contract is a function of two state variables, the spot price and convenience yield. Let $F(S(T),\delta(T),t)$. If F is a twice differentiable function with respect to the spot price, the convenience yield and time, we may use Ito's lemma to write the increment of F as

$$dF = F_t dt + F_s dS(t) + \frac{1}{2}F_{ss}[dS(t)]^2 + F_{s\delta}dS(t)d\delta(t) + \frac{1}{2}F_{\delta\delta}[d\delta(t)]^2 + F_\delta d\delta(t). \quad (117)$$

We need to substitute the stochastic differentials for the spot price and convenience yield into expression (117). In chapter three on page 42, we see that the system of stochastic differential equations for the two-factor futures model is

$$dS(t) = \mu S(t)dt + \sigma_s S(t)dZ_s(t) \quad (118)$$

$$d\delta(t) = k(\alpha - \delta(t))dt + \sigma_c dZ_c(t). \quad (119)$$

Substituting the dynamics for the state variables into equation (117) yields

$$dF = F_t dt + F_s(\mu S(t)dt + \sigma_s S(t)dZ_s(t)) + \frac{1}{2}F_{ss}\sigma_s^2[S(t)]^2 dt$$

$$+ F_{s\delta}\sigma_c\sigma_s\rho_{cs}S(t)dt + \frac{1}{2}F_{\delta\delta}\sigma_c^2 dt + F_\delta(k(\alpha - \delta(t))dt + \sigma_c dZ_c(t)). \quad (120)$$

The expression in (120) is not a risk adjusted process. The Girsanov theorem provides the risk neutral transformations for the driving Wiener processes in (120). These are

$$dZ_s(t) = dZ_s^*(t) - \frac{\mu + \delta(t) - r}{\sigma_s}dt, \quad (121)$$

$$dZ_c(t) = dZ_c^*(t) - \lambda dt. \quad (122)$$

The coefficients in front of the dt terms in expression (121) and (122) are the market prices of risk for the spot price and convenience yield respectively. The asymmetric treatment between the two coefficients is due to the convenience yield being a non-traded good. That is no market exists for this state variable. The convenience yield is a latent variable that is subsumed in the spot price, and since there is no market for the convenience, we do not have an analytical description of its market price of risk. In this case, the market price of risk is denoted simply by λ, and is presumed to be constant. Substituting (121) and (122) into (120) yields

$$dF = F_t dt + F_s\left(\mu S(t)dt + \sigma_s S(t)\left(dZ_s^*(t) - \frac{\mu + \delta(t) - r}{\sigma_s}dt\right)\right)$$
$$+\frac{1}{2}F_{ss}\sigma_s^2[S(t)]^2 dt + F_{s\delta}\sigma_c\sigma_s\rho_{cs}S(t)dt + \frac{1}{2}F_{\delta\delta}\sigma_c^2 dt$$
$$F_\delta\left(k(\alpha - \delta(t))dt + \sigma_c\left(dZ_c^*(t) - \lambda dt\right)\right) . \quad (123)$$

Rearranging terms in (123) yields

$$dF = \left(\frac{1}{2}F_{ss}\sigma_s^2[S(t)]^2 + F_{s\delta}S(t)\sigma_c\sigma_s\rho_{cs} + \frac{1}{2}F_{\delta\delta}\sigma_c^2\right.$$
$$\left. + F_s S(t)(r - \delta(t)) + F_\delta k\left(\alpha - \delta(t) - \frac{\lambda\sigma_c}{k}\right) + F_t\right)dt$$
$$+ \sigma_s S(t)F_s dZ_s^*(t) + \sigma_c F_\delta dZ_c^*(t) \quad . \quad (124)$$

Expression (124) is the risk adjusted dynamic for the futures price. In a risk neutral world the futures price is a martingale process. Again, this implies that the drift term must equal zero. That is,

$$\frac{1}{2}F_{ss}\sigma_s^2[S(t)]^2 + F_{s\delta}S(t)\sigma_c\sigma_s\rho_{cs} + \frac{1}{2}F_{\delta\delta}\sigma_c^2 + F_s S(t)(r - \delta(t))$$

$$+F_\delta k\left(\alpha-\delta(t)-\frac{\lambda\sigma_c}{k}\right)+F_t=0\ . \qquad (125)$$

We see that the expression above is identical to the partial differential equation given by the arbitrage method. This is not a surprise since both methods use a no arbitrage condition to constrain the expected movements of the futures price. Furthermore, the Feynman-Kac theorem illustrates the correspondence between a risk-adjusted expectation and the Black-Scholes particular partial differential equation.

The dynamics of the futures price follows the stochastic differential equation below

$$dF=\sigma_s S(t)F_s dZ_s^*(t)+\sigma_c F_\delta dZ_c^*(t)\,. \qquad (126)$$

This expression may be simplified by substituting in for the partial derivatives of the futures price. Recall the two-factor futures price from chapter three on page 56 equals

$$F(S(T),\delta(T),T)=S(t)\exp\left\{\left[-\alpha+\frac{1}{k}(\sigma_c\lambda-\sigma_c\sigma_s\rho_{cs})+\frac{1}{2}\left(\frac{1}{k}\right)^2\sigma_c^2\right](T-t)\right.$$
$$-\frac{1}{k}\left(\delta(t)-\alpha+\frac{1}{k}(\sigma_c\lambda-\sigma_c\sigma_s\rho_{cs})+\left(\frac{1}{k}\right)^2\sigma_c^2\right)(1-\theta)$$
$$\left.+\left(\frac{1}{k}\right)^2\frac{\sigma_c^2}{4k}\left(1-\theta^2\right)\right\},$$

where $\theta=e^{-k\tau}$. The partial derivative of the futures price with respect to the spot price is

$$F_s=\exp\left\{\left[-\alpha+\frac{1}{k}(\sigma_c\lambda-\sigma_c\sigma_s\rho_{cs})+\frac{1}{2}\left(\frac{1}{k}\right)^2\sigma_c^2\right](T-t)\right.$$
$$-\frac{1}{k}\left(\delta(t)-\alpha+\frac{1}{k}(\sigma_c\lambda-\sigma_c\sigma_s\rho_{cs})+\left(\frac{1}{k}\right)^2\sigma_c^2\right)(1-\theta)$$

$$+\left(\frac{1}{k}\right)^2 \frac{\sigma_c^2}{4k}\left(1-\theta^2\right)\Bigg\}. \tag{127}$$

The partial derivative of the futures price with respect to the convenience yield is

$$F_\delta = -\left(\frac{1}{k}\right)(1-\theta)F = -H_c(\tau)F, \tag{128}$$

where $H_c(\tau) = \frac{\left(1-e^{-k\tau}\right)}{k}$. Substituting expressions (127) and (128) into (126) yields

$$dF(t) = \sigma_s F dZ_s^*(t) - \sigma_c H_c(\tau) F dZ_c^*(t). \tag{129}$$

We define another standard Weiner process and a diffusion parameter, σ_F, such that

$$\sigma_F dZ_F^*(t) \equiv \sigma_s dZ_s^*(t) - \sigma_c H_c(\tau) dZ_c^*(t). \tag{130}$$

Then equation (129) can be written as

$$dF(t) = \sigma_F F dZ_F^*(t), \tag{131}$$

where $\sigma_F = \sqrt{\sigma_s^2 + \sigma_c^2 H_c^2(\tau) - 2\sigma_s\sigma_c\rho_{sc}H_c(\tau)}$. Equation (130) represents the term structure of the volatility for a commodity futures price. Here the volatility depends on the volatility of the spot price, the volatility of the convenience yield, the correlation between the spot price and convenience yield, the speed of adjustment parameter and the time to maturity. The speed of adjustment parameter enters the volatility term for the futures because of the volatility for the convenience yield is dampened by the mean reversion process. That is, over time we should expect the volatility of the convenience yield to decrease because the state variable reverts to its long run mean. Furthermore, as the speed of adjustment increases the volatility should decrease. This occurs because the convenience yield will begin revert back towards its mean much faster.

Looking at the volatility in expression (131) it is interesting to note, that in this model, the volatility does not depend on the level of the spot price or the convenience yield. The dynamics of the futures price is only a function of the deterministic parameters for the state variables. As a consequence, we may treat the futures price as an exogenous process when pricing options written on the futures contract.

The first step to pricing options written on commodity futures is to determine what process the futures price follows. We have this process and it is given by expression (131). This process may be alternatively written. Consider the following. Let $H = \ln F$, where H is twice differentiable in F. With this transformation and Ito's lemma, the increment of H is expressed as

$$\begin{aligned} dH &= H_F dF + \frac{1}{2} H_{FF} dF^2 \\ &= \frac{1}{F}\left(\sigma_F F dZ_F^*(t)\right) - \frac{1}{2F^2}\left(\sigma_F^2 F^2 dt\right) \\ &= -\frac{1}{2}\sigma_F^2 dt + \sigma_F dZ_F^*(t). \end{aligned} \tag{132}$$

We want to find an expression for the futures price at the time of maturity for the option contract. To do this, we integrate over equation (132) from t to T_1. This yields

$$\int_t^{T_1} dH = -\frac{1}{2}\int_t^{T_1} \sigma_F^2 ds + \int_t^{T_1} \sigma_F dZ_F^*,$$

$$H(T_1) - H(t) = -\frac{1}{2}\int_t^{T_1} \sigma_F^2 ds + \int_t^{T_1} \sigma_F dZ_F^*,$$

$$H(T_1) = H(t) - \frac{1}{2}\int_t^{T_1} \sigma_F^2 ds + \int_t^{T_1} \sigma_F dZ_F^*,$$

$$F(T_1) = F(t)\exp\left[-\frac{1}{2}\int_t^{T_1}\sigma_F^2 ds + \int_t^{T_1}\sigma_F dZ_F^*\right],$$

$$F(T_1) = F(t)\exp[Y(T_1)], \tag{133}$$

where $Y(T_1) = -\frac{1}{2}\int_t^{T_1}\sigma_F^2 ds + \int_t^{T_1}\sigma_F dZ_F^*$. Expression (133) is the alternative expression for the dynamic expressed in equation (131). This expression is identical in form to expression (77). An important difference between these expression (77) and (133) is the treatment of the volatility term. Currently, the diffusion is a function of the spot price and the convenience yield, where as, the diffusion term in expression (77) is a function of the spot price volatility only. The reason for the difference is due to the set up of the model. The first model presumes all the uncertainty in the futures price is summarized by a mean reverting spot price and the second model presumes two factor influence the futures price.

Following the derivation of the option contract in the one factor model, we may find the expected futures price at the option's maturity. To determine this value, we need the mean and variance of $H(T)$. The mean is

$$E_t[H(T_1)] = H(t) - \frac{1}{2}\int_t^{T_1}\sigma_F^2 ds \ . \tag{134}$$

The variance equals

$$V_t[H(T_1)] = \int_t^{T_1}\sigma_F^2 ds. \tag{135}$$

Let $v^2 \equiv \int_t^{T_1}\sigma_F^2 ds$, then the terminal futures price at the option's expiration equals

$$E_t^*[F(T_1)] = \exp\left[E_t^*[H(T_1)] + \frac{1}{2}V_t^*[H(T_1)]\right],$$

$$= \exp\left[H(t) - \frac{1}{2}v^2 + \frac{1}{2}v^2\right],$$

$$= \exp[H(t)],$$

$$= F(t). \qquad (136)$$

Again, the reader should note that the expected futures price at time T_1 is today's futures price.

Now, if we can evaluate the integral of the diffusion coefficient for the futures price process, then we can price the option contract following the procedure above. Solving the integral, $v^2 \equiv \int_t^{T_1} \sigma_F^2 ds$ becomes

$$v^2 \equiv \int_t^{T_1} \sigma_F^2 ds,$$

$$= \int_t^{T_1} \left(\sigma_s^2 + \sigma_c^2 H_c^2(\tau) - 2\rho_{sc}\sigma_s\sigma_c H_c(\tau)\right) ds,$$

$$= \int_t^{T_1} \sigma_s^2 ds + \int_t^{T_1} \sigma_c^2 H_c^2(\tau) ds - \int_t^{T_1} 2\rho_{sc}\sigma_s\sigma_c H_c(\tau) ds. \qquad (137)$$

Solving the individual integrals we obtain

$$\int_t^{T_1} \sigma_s^2 ds = \sigma_s^2 (T_1 - t), \qquad (138)$$

$$\int_t^{T_1} \sigma_c^2 H_c^2(\tau) ds = \int_t^{T_1} \left[\frac{\sigma_c}{k_c}\left(1 - e^{-k_c(T-s)}\right)\right]^2 ds,$$

$$= \int_t^{T_1} \left[\frac{\sigma_c^2}{k_c^2} + \frac{\sigma_c^2}{k_c^2} e^{-2k_c(T-s)} - \frac{2\sigma_c^2}{k_c^2} e^{-k_c(T-s)}\right] ds,$$

$$= \int_t^{T_1} \frac{\sigma_c^2}{k_c^2} ds + \int_t^{T_1} \frac{\sigma_c^2}{k_c^2} e^{-2k_c(T-s)} ds - \int_t^{T_1} \frac{2\sigma_c^2}{k_c^2} e^{-k_c(T-s)} ds,$$

$$= \frac{\sigma_c^2}{k_c^2}(T_1 - t) + \frac{\sigma_c^2}{2k_c^3}\left(e^{-2k_c(T-T_1)} - e^{-2k_c(T-t)}\right) - \frac{2\sigma_c^2}{k_c^3}\left(e^{-k_c(T-T_1)} - e^{-k_c(T-t)}\right),$$

$$= \frac{\sigma_c^2}{k_c^2}\left[(T_1 - t) + \frac{1}{2k_c}\left(e^{-2k_c(T-T_1)} - e^{-2k_c(T-t)}\right) - \frac{2}{k_c}\left(e^{-k_c(T-T_1)} - e^{-k_c(T-t)}\right)\right]. \tag{139}$$

$$\int_t^{T_1} 2\rho_{sc}\sigma_s\sigma_c H_c(\tau) ds = \int_t^{T_1} \frac{2\rho_{sc}\sigma_s\sigma_c}{k_c}\left(1 - e^{-k_c(T-s)}\right) ds,$$

$$= \frac{2\rho_{sc}\sigma_s\sigma_c}{k_c}(T_1 - t) - \frac{2\rho_{sc}\sigma_s\sigma_c}{k_c^2}\left(e^{-k_c(T-T_1)} - e^{-k_c(T-t)}\right),$$

$$= \frac{2\rho_{sc}\sigma_s\sigma_c}{k_c}\left[(T_1 - t) - \frac{1}{k_c}\left(e^{-k_c(T-T_1)} - e^{-k_c(T-t)}\right)\right]. \tag{140}$$

Substituting expressions (138)-(140) into (137) yields

$$v^2 = \sigma_s^2(T_1 - t) + \frac{\sigma_c^2}{k_c^2}\left[(T_1 - t) + \frac{1}{2k_c}\left(e^{-2k_c(T-T_1)} - e^{-2k_c(T-t)}\right) - \frac{2}{k_c}\left(e^{-k_c(T-T_1)} - e^{-k_c(T-t)}\right)\right]$$

$$- \frac{2\rho_{sc}\sigma_s\sigma_c}{k_c}\left[(T_1 - t) - \frac{1}{k_c}\left(e^{-k_c(T-T_1)} - e^{-k_c(T-t)}\right)\right]. \tag{141}$$

We have the expected terminal futures price, therefore we are ready to price an option contract written on a commodity futures contract. The option price is

$$C(F,t) = e^{-r\tau_1} F(t) N(d_1) - e^{-r\tau_1} X N(d_2), \tag{142}$$

where

$$d_1 = \frac{\ln\left(\frac{F(t)}{X}\right) + \frac{1}{2}v^2}{v}, \tag{143}$$

and

$$d_2 = d_1 - \nu. \tag{144}$$

Expression (142) is the price of an option contract written on a commodity futures contract, whose price is influence by a stochastic spot price and convenience yield. We see the difference between the one- and two-factor option models is the volatility term used in expression (142). For the one-factor model the volatility term is given by expression (82) in section 4.2 above. This is

$$\nu^2 = \frac{\sigma^2}{2k}\left(e^{-2k(T-T_1)} - e^{-2k(T-t)}\right).$$

The two-factor volatility is given by expression (141) above, and is more complex than expression (82). That is, the introduction of a stochastic convenience yield increases the sophistication of the option model by adding another source of uncertainty. The first model is not a special case of the second model. The one factor model shows that the uncertainty in the futures market comes entirely from a mean reverting spot price. The second model allows uncertainty from two sources. One is the spot price which follows a regular geometric Brownian motion, and the other source is the convenience yield which is said to mean revert. The advantage of the second model is it allows for greater flexibility and realism in modeling commodity prices.

4.4 Three-factor model

Following the practice of chapter three, further rigor is added to the analysis by considering the addition of a stochastic interest rate to the two-factor model. Recall from chapter three, the following system of stochastic differential equations is used to price a futures contract

$$dS(t) = (r(t) - \delta(t))S(t)dt + \sigma_s S(t)dZ_s^*(t), \tag{145}$$

$$d\delta(t) = \left(k_c(\alpha - \delta(t)) - \lambda\sigma_c\right)dt + \sigma_c dZ_c^*(t), \tag{146}$$

$$dr(t) = \left(f_t(s,t) + k_r f(s,t) + \frac{\sigma_r^2}{2k_r}\left(1 - e^{-2k_r(t-s)}\right) - k_r r(t)\right)dt + \sigma_r dZ_r^*(t). \tag{147}$$

Note, that we start with the risk-adjusted processes. We posit that the futures price is a twice-continuously differentiable function with respect to the spot price, convenience yield, interest rate and time, namely $F(S,\delta,r,t)$. Using Ito's lemma, the dynamics of the futures price may be expressed as

$$dF(t) = F_s dS(t) + \frac{1}{2}F_{ss}[dS(t)]^2 + F_\delta d\delta(t) + \frac{1}{2}F_{\delta\delta}[d\delta(t)]^2 + F_r dr(t) + \frac{1}{2}F_{rr}[dr(t)]^2$$

$$F_t dt + F_{s\delta} dSd\delta + F_{sr} dSdr + F_{r\delta} drd\delta. \tag{148}$$

Substituting the stochastic differentials into (148) yields

$$dF(t) = \frac{1}{2}F_{ss}\sigma_s^2[S(t)]^2 dt + \frac{1}{2}F_{\delta\delta}\sigma_c^2 dt + \frac{1}{2}F_{rr}\sigma_r^2 dt$$

$$+ F_s\left((r(t) - \delta(t))S(t)dt + \sigma S(t)dZ_s^*(t)\right)$$

$$+ F_\delta\left(\left(k_c(\alpha - \delta(t)) - \lambda\sigma_c\right)dt + \sigma_c dZ_c^*(t)\right)$$

$$+ F_r\left(\left(f_t(s,t) + k_r f(s,t) + \frac{\sigma_r^2}{2k_r}\left(1 - e^{-2k_r(t-s)}\right) - k_r r(t)\right)dt + \sigma_r dZ_r^*(t)\right)$$

$$+ F_{s\delta}S(t)\sigma_s\sigma_c\rho_{sc}dt + F_{sr}S(t)\sigma_s\sigma_r\rho_{sr}dt + F_{r\delta}\sigma_r\sigma_c\rho_{rc}dt$$

$$+ F_t dt.$$

Rearranging the above yields

$$dF(t) = \left[\frac{1}{2}F_{ss}\sigma_s^2[S(t)]^2 + \frac{1}{2}F_{\delta\delta}\sigma_c^2 + F_{rr}\frac{1}{2}\sigma_r^2 + F_s\left((r(t) - \delta(t))S(t)\right) + F_\delta\left(k_c(\alpha - \delta(t)) - \lambda\sigma_c\right)\right.$$

$$+F_r\left(f_t(s,t)+k_r f(s,t)+\frac{\sigma_r^2}{2k_r}\left(1-e^{-2k_r(t-s)}\right)-k_r r(t)\right)+F_{s\delta}S(t)\sigma_s\sigma_c\rho_{sc}$$

$$+F_{sr}S(t)\sigma_s\sigma_r\rho_{sr}+F_{r\delta}\sigma_r\sigma_c\rho_{rc}+F_t\Big]dt$$

$$+F_s\sigma S(t)dZ_s^*(t)+F_c\sigma_c dZ_c^*(t)+F_r\sigma_r dZ_r^*(t) \tag{149}$$

Expression (149) is the risk-adjusted stochastic differential for the three-factor futures price. In a risk neutral world the futures price is a martingale process. From page 142, we see this implies that the drift term above equals zero, thereby reducing expression (149) to

$$dF(t)=F_s\sigma_s S(t)dZ_s^*(t)+F_\delta\sigma_c dZ_c^*(t)+F_r\sigma_r dZ_r^*(t)\,. \tag{150}$$

Furthermore, we may simplify the above expression by substituting in for the partial derivatives of $F(t)$. Recall from chapter three section three page 88, we found the futures price to equal

$$F(S(T),\delta(T),r(T),T)=S(t)A(\tau)D_1(\tau)D_2(\tau)D_3(\tau)e^{-H_c(\tau)\delta(t)}\frac{1}{P(t,T)}, \tag{151}$$

where $$A(\tau)=\exp\left(\frac{\left(H_c(\tau)-\tau\right)\left(k_c^2\alpha-k_c\lambda\sigma_c-\frac{\sigma_c^2}{2}+\rho_{cs}\sigma_s\sigma_c k_c\right)}{k_c^2}-\frac{\sigma_c^2H_c^2(\tau)}{4k_c}\right), \tag{152}$$

$$D_1(\tau)=\exp\left[\frac{\sigma_c\sigma_r\rho_{rc}}{k_r+k_c}\left[\frac{\left(H_r(\tau)-\tau\right)}{k_r}+\frac{\left(H_c(\tau)-\tau\right)}{k_c}+H_r(\tau)H_c(\tau)\right]\right], \tag{153}$$

$$D_2(\tau)=\exp\left[\left(\frac{\sigma_s\sigma_r\rho_{rs}}{k_r}\left(\tau-H_r(\tau)\right)\right)\right], \tag{154}$$

$$D_3(\tau)=\exp\left[-\left(H_r(\tau)-\tau\right)\frac{\sigma_r^2}{k_r^2}-\frac{\sigma_r^2H_r^2(\tau)}{2k_r}\right], \tag{155}$$

where $H_c(\tau)=\frac{\left(1-e^{-k_c\tau}\right)}{k_c}$ and $H_r(\tau)=\frac{\left(1-e^{-k_r\tau}\right)}{k_r}$. The partial derivative of the futures price with respect to the spot price is

$$F_s = A(\tau)D_1(\tau)D_2(\tau)D_3(\tau)\exp\left(-H_c(\tau)\delta(t)\right)\frac{1}{P(t,T)}. \tag{156}$$

The partial derivative of the futures price with respect to the convenience yield is

$$F_\delta = -H_c(\tau)S(t)A(\tau)D_1(\tau)D_2(\tau)D_3(\tau)\exp\left(-H_c(\tau)\delta(t)\right)\frac{1}{P(t,T)},$$

$$F_\delta = -H_c(\tau)F(S(T),\delta(T),r(T),T) \tag{157}$$

The partial derivative of the futures price with respect to the interest rate is

$$F_r = H_r(\tau)S(t)A(\tau)D_1(\tau)D_2(\tau)D_3(\tau)\exp\left(-H_c(\tau)\delta(t)\right)\frac{1}{P(t,T)}.$$

$$F_r = H_r(\tau)F(S(T),\delta(T),r(T),T) \tag{158}$$

Substituting (156)-(158) into (155) yields

$$dF(t) = \sigma_s S(t)dZ_s^*(t)A(\tau)D_1(\tau)D_2(\tau)D_3(\tau)\exp\left(-H_c(\tau)\delta(t)\right)\frac{1}{P(t,T)}$$

$$+\sigma_c dZ_c^*(t)\left(-H_c(\tau)F(S(T),\delta(T),r(T),T)\right)$$

$$+\sigma_r dZ_r^*(t)\left(H_r(\tau)F(S(T),\delta(T),r(T),T)\right). \tag{159}$$

The first term in expression (159) may be reduced to the diffusion coefficient, the risk-adjusted Brownian motion, and the futures price. For ease of exposition the notation for the futures price is reduced to *F*(*t*). Expression (159) becomes

$$dF(t) = F(t)\sigma_s dZ_s^*(t) - F(t)H_c(\tau)\sigma_c dZ_c^*(t) + F(t)H_r(\tau)\sigma_r dZ_r^*(t)^*. \tag{160}$$

Define $\sigma_F(\tau)dZ_F^* \equiv \sigma_s dZ_s^*(t) - H_c(\tau)\sigma_c dZ_c^*(t) + H_r(\tau)\sigma_r dZ_r^*(t)$. The futures price volatility is the linear combination of the normally distributed innovation terms for the state variables underlying the futures price. The above becomes

$$\frac{dF(t)}{F(t)} = \sigma_F(\tau)dZ_F^*(t). \tag{161}$$

Let $H = \ln F(t)$, where H is twice differentiable in $F(t)$. With this transformation and Ito's lemma, the increment of H is expressed as

$$dH = H_F dF(t) + \frac{1}{2}H_{FF}[dF(t)]^2$$

$$= \frac{1}{F(t)}\left(\sigma_F F(t)dZ_F^*(t)\right) - \frac{1}{2[F(t)]^2}\left(\sigma_F^2[F(t)]^2 dt\right)$$

$$= -\frac{1}{2}\sigma_F^2 dt + \sigma_F dZ_F^*(t). \tag{162}$$

Working with expression (162), we may find the futures price at the maturity of the option contract. That is, we integrate over equation (162) from t to T_1. This yields

$$\int_t^{T_1} dH = -\frac{1}{2}\int_t^{T_1}\sigma_F^2 ds + \int_t^{T_1}\sigma_F dZ_F^*,$$

$$H(T_1) - H(t) = -\frac{1}{2}\int_t^{T_1}\sigma_F^2 ds + \int_t^{T_1}\sigma_F dZ_F^*,$$

$$H(T_1) = H(t) - \frac{1}{2}\int_t^{T_1}\sigma_F^2 ds + \int_t^{T_1}\sigma_F dZ_F^*,$$

$$F(T_1) = F(t)\exp\left[-\frac{1}{2}\int_t^{T_1}\sigma_F^2 ds + \int_t^{T_1}\sigma_F dZ_F^*\right],$$

$$F(T_1) = F(t)\exp[Y(T_1)], \tag{163}$$

where $Y(T_1) = -\frac{1}{2}\int_t^{T_1}\sigma_F^2 ds + \int_t^{T_1}\sigma_F dZ_F^*$. Expression (163) gives the terminal value of a futures price at the maturity of the option contract.

Following the derivation of the option contract in the one factor model, we need the expected futures price at the option's expiration. This is equal to

$$E_t^*[F(T_1)] = F(t)\exp\left[E_t^*[Y(T_1)] + \frac{1}{2}V_t^*[Y(T_1)]\right]. \tag{164}$$

The risk-adjusted conditional mean of $Y(T_1)$ is

$$E_t^*[Y(T_1)] = -\frac{1}{2}\int_t^{T_1}\sigma_F^2 ds , \tag{165}$$

and the risk-adjusted conditional variance is

$$V_t^*[Y(T_1)] = \int_t^{T_1}\sigma_F^2 ds . \tag{166}$$

The expected futures price is $E_t^*[F(T_1)] = F(t)$. The volatility of the futures price is $v^2 \equiv \int_t^{T_1}\sigma_F^2 ds$. Therefore, all we need is the expression for $v^2 \equiv \int_t^{T_1}\sigma_F^2 ds$, and we may solve for the option price. Solving the integral, $v^2 \equiv \int_t^{T_1}\sigma_F^2 ds$ becomes

$$v^2 \equiv \int_t^{T_1}\sigma_F^2 ds ,$$

$$= \int_t^{T_1}\Big(\sigma_s^2 + \sigma_c^2 H_c^2(\tau) + \sigma_r^2 H_r^2(\tau) - 2\rho_{sc}\sigma_s\sigma_c H_c(\tau)$$

$$+ 2\sigma_r\sigma_s\rho_{sr}H_r(\tau) - 2\sigma_r\sigma_c\rho_{cr}H_r(\tau)H_c(\tau)\Big)ds$$

$$= \int_t^{T_1}\sigma_s^2 ds + \int_t^{T_1}\sigma_c^2 H_c^2(\tau)ds + \int_t^{T_1}\sigma_r^2 H_r^2(\tau)$$

$$-2\int_t^{T_1}\rho_{sc}\sigma_s\sigma_c H_c(\tau)ds+2\int_t^{T_1}\rho_{sr}\sigma_s\sigma_r H_r(\tau)ds$$

$$-\int_t^{T_1}2\sigma_r\sigma_c\rho_{cr}H_r(\tau)H_c(\tau)ds. \qquad (167)$$

From our analysis of the volatility term in the two-factor option model on page 164, we know

$$\int_t^{T_1}\sigma_s^2 ds=\sigma_s^2(T_1-t), \qquad (168)$$

$$\int_t^{T_1}\sigma_c^2 H_c^2(\tau)ds=\frac{\sigma_c^2}{k_c^2}\left[(T_1-t)+\frac{1}{2k_c}\left(e^{-2k_c(T-T_1)}-e^{-2k_c(T-t)}\right)-\frac{2}{k_c}\left(e^{-k_c(T-T_1)}-e^{-k_c(T-t)}\right)\right],$$

(169)

$$\int_t^{T_1}2\rho_{sc}\sigma_s\sigma_c H_c(\tau)ds=\frac{2\rho_{sc}\sigma_s\sigma_c}{k_c}\left[(T_1-t)-\frac{1}{k_c}\left(e^{-k_c(T-T_1)}-e^{-k_c(T-t)}\right)\right]. \qquad (170)$$

Solving the remaining integrals in expression (167) yields

$$\int_t^{T_1}\sigma_r^2 H_r^2(\tau)ds=\int_t^{T_1}\left[\frac{\sigma_r}{k_r}\left(1-e^{-k_r(T-s)}\right)\right]^2 ds,$$

$$=\int_t^{T_1}\left[\frac{\sigma_r^2}{k_r^2}+\frac{\sigma_r^2}{k_r^2}e^{-2k_r(T-s)}-\frac{2\sigma_r^2}{k_r^2}e^{-k_r(T-s)}\right]ds,$$

$$=\int_t^{T_1}\frac{\sigma_r^2}{k_r^2}ds+\int_t^{T_1}\frac{\sigma_r^2}{k_r^2}e^{-2k_r(T-s)}ds-\int_t^{T_1}\frac{2\sigma_r^2}{k_r^2}e^{-k_r(T-s)}ds,$$

$$=\frac{\sigma_r^2}{k_r^2}(T_1-t)+\frac{\sigma_r^2}{2k_r^3}\left(e^{-2k_r(T-T_1)}-e^{-2k_r(T-t)}\right)-\frac{2\sigma_r^2}{k_r^3}\left(e^{-k_r(T-T_1)}-e^{-k_r(T-t)}\right),$$

$$=\frac{\sigma_r^2}{k_r^2}\left[(T_1-t)+\frac{1}{2k_r}\left(e^{-2k_r(T-T_1)}-e^{-2k_r(T-t)}\right)-\frac{2}{k_r}\left(e^{-k_r(T-T_1)}-e^{-k_r(T-t)}\right)\right].$$

(171)

$$\int_t^{T_1} 2\rho_{sr}\sigma_s\sigma_r H_r(\tau)ds = \int_t^{T_1} \frac{2\rho_{sr}\sigma_s\sigma_r}{k_r}\left(1-e^{-k_r(T-s)}\right)ds,$$

$$= \frac{2\rho_{sr}\sigma_s\sigma_r}{k_r}(T_1-t) - \frac{2\rho_{sr}\sigma_s\sigma_r}{k_r^2}\left(e^{-k_r(T-T_1)} - e^{-k_r(T-t)}\right),$$

$$= \frac{2\rho_{sr}\sigma_s\sigma_r}{k_r}\left[(T_1-t) - \frac{1}{k_r}\left(e^{-k_r(T-T_1)} - e^{-k_r(T-t)}\right)\right]. \tag{172}$$

$$\int_t^{T_1} 2\sigma_r\sigma_c\rho_{cr}H_r(\tau)H_c(\tau)ds = \int_t^{T_1}\frac{2\sigma_r\sigma_c\rho_{cr}}{k_r k_c}\left(1-e^{-k_r(T-s)}\right)\left(1-e^{-k_c(T-s)}\right)ds$$

$$= \frac{\sigma_c\sigma_r\rho_{rc}}{k_c k_r}\int_t^{T_1}\left(1-e^{-k_r(T-s)} - e^{-k_c(T-s)} + e^{-k_r(T-s)}e^{-k_c(T-s)}\right)ds,$$

$$= \frac{2\sigma_c\sigma_r\rho_{rc}}{k_c k_r}\left[(T_1-t) - \frac{\left(e^{-k_r(T-T_1)} - e^{-k_r(T-t)}\right)}{k_r} - \frac{\left(e^{-k_c(T-T_1)} - e^{-k_c(T-t)}\right)}{k_c}\right.$$

$$\left. + \frac{1}{k_r+k_c}\left(e^{-k_r(T-T_1)}e^{-k_c(T-T_1)} - e^{-k_r(T-t)}e^{-k_c(T-t)}\right)\right]. \tag{173}$$

Substituting expressions (168)-(173) into (167) yields

$$v^2 = \sigma_s^2(T_1-t) + \frac{\sigma_c^2}{k_c^2}\left[(T_1-t) + \frac{1}{2k_c}\left(e^{-2k_c(T-T_1)} - e^{-2k_c(T-t)}\right) - \frac{2}{k_c}\left(e^{-k_c(T-T_1)} - e^{-k_c(T-t)}\right)\right]$$

$$+ \frac{\sigma_r^2}{k_r^2}\left[(T_1-t) + \frac{1}{2k_r}\left(e^{-2k_r(T-T_1)} - e^{-2k_r(T-t)}\right) - \frac{2}{k_r}\left(e^{-k_r(T-T_1)} - e^{-k_r(T-t)}\right)\right]$$

$$- \frac{2\rho_{sc}\sigma_s\sigma_c}{k_c}\left[(T_1-t) - \frac{1}{k_c}\left(e^{-k_c(T-T_1)} - e^{-k_c(T-t)}\right)\right]$$

$$+ \frac{2\rho_{sr}\sigma_s\sigma_r}{k_r}\left[(T_1-t) - \frac{1}{k_r}\left(e^{-k_r(T-T_1)} - e^{-k_r(T-t)}\right)\right]$$

$$- \frac{2\sigma_c\sigma_r\rho_{rc}}{k_c k_r}\left[(T_1-t) - \frac{\left(e^{-k_r(T-T_1)} - e^{-k_r(T-t)}\right)}{k_r} - \frac{\left(e^{-k_c(T-T_1)} - e^{-k_c(T-t)}\right)}{k_c}\right.$$

$$+\frac{1}{k_r+k_c}\left(e^{-k_r(T-T_1)}e^{-k_c(T-T_1)}-e^{-k_r(T-t)}e^{-k_c(T-t)}\right)\Bigg]. \tag{174}$$

Equation (161) shows the futures price follows a geometric Brownian motion. In addition, we have the expected value of the two-factor futures price. Therefore, we are ready to price the option written on the futures contract.

Unlike the first two option models, we are unable to use the Black-Scholes methodology to price an option contract written on the three-factor futures price. For the one- and two-factor option models we proceed as follows. Under an equivalent martingale measure the price of a European call option written on the risk-adjusted futures contract can be expressed as

$$C(F,t)=e^{-r(T_1-t)}E_t^*\left[C(F,T_1)\right], \tag{175}$$

where $C(F,t)$ is the price of the call option today, and $C(F,T_1)$ is the risk adjusted call option at a terminal date T_1. Substituting in for the risk-adjusted terminal call price we get

$$\begin{aligned}
C(F,t) &= e^{-r\tau_1}E_t^*\left[\max\left(F(T_1)-X,0\right)\right],\\
&= e^{-r\tau_1}E_t^*\left[\max\left(F(t)e^{Y(T_1)}-X,0\right)\right].\\
&= e^{-r\tau_1}\left\{E_t^*\left[F(t)e^{Y(T_1)}-X \mid F(T_1)\geq X\right]+E_t^*\left[0 \mid F(T_1)<X\right]\right\}\\
&= e^{-r\tau_1}\left\{E_t^*\left[F(t)e^{Y(T_1)}-X \mid F(t)e^{Y(T_1)}\geq X\right]+E_t^*\left[0 \mid F(t)e^{Y(T_1)}<X\right]\right\}\\
&= e^{-r\tau_1}\left\{E_t^*\left[F(t)e^{Y(T_1)}-X \mid e^{Y(T_1)}\geq \frac{X}{F(t)}\right]+E_t^*\left[0 \mid e^{Y(T_1)}<\frac{X}{F(t)}\right]\right\}\\
&= e^{-r\tau_1}\left\{E_t^*\left[F(t)e^{Y(T_1)}-X \mid Y(T_1)\geq \ln\left(\frac{X}{F(t)}\right)\right]+E_t^*\left[0 \mid Y(T_1)<\ln\left(\frac{X}{F(t)}\right)\right]\right\}.
\end{aligned} \tag{176}$$

The linear combination of partial expectations in expression (176) may be rewritten as

$$C(F,t) = e^{-rT_1} \int_{-\infty}^{\ln\left(\frac{X}{F(t)}\right)} (0)dP^* + e^{-rT_1} \int_{\ln\left(\frac{X}{F(t)}\right)}^{\infty} \left(F(t)e^{y(T_1)} - X\right)dP^* \tag{177}$$

where dP^* is the risk-adjusted probability measure denoted as

$$dP^* = \frac{1}{\sqrt{2\pi v^2}} \exp\left[-\frac{\left(y(T_1) + \frac{1}{2}v^2\right)^2}{2v^2}\right] dy(T_1). \tag{178}$$

With the introduction of a stochastic interest rate to the model the risk-neutral expectation of the option becomes

$$C(F,t) = E_t^*\left[e^{\int_t^{T_1} -r(s)ds} \max\left(F(t)e^{Y(T_1)} - X, 0\right)\right]. \tag{179}$$

The expectation above cannot be evaluated because the interest process is correlated to the options payout. Casual observation of expression (160) shows that the futures price movements are a function of the interest rate process, thereby making the futures price process correlated with the interest rate process. If the two components are independent then the expectation could be evaluated. To illustrate this problem consider the following example.

Consider to random variables X and Y. Each random variable has two observations, which we denote as $[x_1, x_2]$, and $[y_1, y_2]$, with marginal probabilities denoted by $f(x_i)$ and $f(y_i)$; where $i = 1,2$. If X and Y are independent then

$$E(XY) = E(X)E(Y). \tag{180}$$

To see this we know that

$$E(XY) = \sum_{y}\sum_{y} xyf(x,y), \tag{181}$$

where $f(x,y)$ is the joint probability density function for X and Y. We may rewrite the double sum as

$$= x_1y_1f(x_1,y_1) + x_2y_1f(x_2,y_1) + x_1y_2f(x_1,y_2) + x_2y_2f(x_2,y_2). \tag{182}$$

Since X and Y are independent the joint density function can be rewritten as the product of the marginal density functions. This is

$$= x_1f(x_1)y_1f(y_1) + x_2f(x_2)y_1f(y_1) + x_1y_2f(x_1)f(y_2) + x_2f(x_2)y_2f(y_2). \tag{183}$$

Simplifying expression (183) yields

$$= [x_1f(x_1) + x_2f(x_2)]y_1f(y_1) + [x_1f(x_1) + x_2f(x_2)]y_2f(y_2).$$

$$= E[X]y_1f(y_1) + E[X]y_2f(y_2). \tag{184}$$

Factoring out the expected value of X yields

$$= (E[X])[y_1f(y_1) + y_2f(y_2)]$$

$$= E[X]E[Y]. \tag{185}$$

We see that if X and Y are independent then the expectation of the product of X and Y is equal to the product of the individual expectations. If X is analogous to $e^{\int_t^{T_1} -r(s)ds}$ and Y to the futures price, then we may use the above result to evaluate the expectation. The result is good only if the two variables are independent, and for our model this is not applicable. The option price, which is a function of the futures price, and interest rates are correlated. Therefore, we need to consider the expectation for two correlated variables. Let X and Y be two dependent variables. What is $E(XY)$ equal to? By definition

$$E(XY) = \sum_{y}\sum_{y} xyf(x,y). \tag{186}$$

We may rewrite the double sum as

$$= x_1 y_1 f(x_1, y_1) + x_2 y_1 f(x_2, y_1) + x_1 y_2 f(x_1, y_2) + x_2 y_2 f(x_2, y_2). \quad (187)$$

We may further rewrite the above by rewriting the joint densities as follows

$$= x_1 y_1 f(x_1) f(y_1 \mid x_1) + x_2 y_1 f(x_2) f(y_1 \mid x_2)$$

$$+ x_1 y_2 f(x_1) f(y_2 \mid x_1) + x_2 y_2 f(x_2) f(y_2 \mid x_2). \quad (188)$$

Simplifying the expression above yields

$$= x_1 y_1 f(x_1) f(y_1 \mid x_1) + x_1 y_2 f(x_1) f(y_2 \mid x_1)$$

$$+ x_2 y_1 f(x_2) f(y_1 \mid x_2) + x_2 y_2 f(x_2) f(y_2 \mid x_2). \quad (189)$$

Factoring yields

$$= x_1 f(x_1)[y_1 f(y_1 \mid x_1) + y_2 f(y_2 \mid x_1)] + x_2 f(x_2)[y_1 f(y_1 \mid x_2) + y_2 f(y_2 \mid x_2)].$$

$$= x_1 f(x_1) E[Y \mid X = x_1] + x_2 f(x_2) E[Y \mid X = x_2]. \quad (190)$$

The expression above cannot be simplified any further. The expectation for the multiplicative interaction between two correlated random variable X and Y does not simplify to simple expression. This is why we may not find a closed for expression for the Black-Scholes model when the interests is stochastic.

4.5 Merton's Model

In section 4.1 of this chapter, we use the Black-Scholes method of solution to price options written on commodity futures. The Black-Scholes methodology presumes that the only source of uncertainty is the underlying security, whose price is said to follow a geometric Brownian motion. In consideration of this observation, if we want to invoke the Black-Scholes methodology, we must first show that the stochastic process for the futures price is indeed a geometric Brownian motion. Using Ito's lemma and the fact that the futures price is twice continuously differentiable function of time and other state variables, we do show the futures price follows a geometric Brownian motion. This result allows us to use the Black-Scholes model to price options written on commodity futures. The question we investigate now is whether or not the assumption about the price movements for commodities is a fair characterization. If not, we need an alternative method for pricing options.

Casual observation of the time series behavior of commodity spot prices show these prices typically do not resemble a geometric Brownian motion *all* the time. That is, commodity spot prices often move in small increments but occasionally exhibit random price spikes. These price spikes are usually due to aberrant market conditions causing sudden changes in either market supply or demand schedules. Therefore, these random price jumps create another source of uncertainty in the spot market.

Another interesting side effect of random price spikes in spot prices is the impact they have on futures prices. Futures prices are a function of spot prices and are necessarily influenced by these discrete price movements. If we witness spot price jumps, we can therefore believe that the futures price will not fit the assumed dynamics of the Black-Scholes model. In light of these casual observations, we need to find a model capable of pricing options that incorporates random discontinuous jumps.

Merton (1976) develops an option pricing model for stock options that incorporates random jumps. His model is a simple modification of the Black-Scholes model. Merton's model represents the asset's price dynamic with an equation that allows unpredictable changes from two categories: normal events and rare events. Normal price vibrations occur in a continuous fashion and are modeled by the increment of a standard Wiener process, $dZ(t)$. Rare events, however, occur sporadically and induce a discrete random jump in the spot price. To model these abnormal price vibrations, Merton uses a Poisson process, which he defines as $dq(t)$.

We can be more specific. It is clear from Merton that changes in the asset's price are due to a mixture of normal and rare events. If we witness a rare event (a jump) the stock price experiences a discrete jump. Typically, the size and time of this event are unknowns. For now, however, if a jump occurs we assume that jump in the commodity's price is of size 1. The jump size can be characterized by any type of distribution, and later we will change this assumption, but for ease of exposition we analyze the case for jump size of one. At any instant $t-1$, one has

$$q(t) - q(t-1) = \begin{cases} 1 & \text{with probability} \quad \lambda dt \\ 0 & \text{with probability} \quad 1 - \lambda dt \end{cases}$$

where dt is an infinitesimal interval of time and λ is the mean number of arrivals per unit of time. λ does not depend on the information set available at time $t-1$. We let

$$dq(t) = q(t) - q(t-1), \tag{191}$$

where $dq(t)$ represents the number of jumps that occur over an infinitesimal interval dt. The size of each of these jumps is 1, and they occur with a constant rate λ. Remember, the size of the jump need not be restricted to jump size 1, but for this example the number of jumps is restricted to one over the interval dt.

We are interested in modeling discrete random jumps in asset prices that occur over time. It seems natural to model this behavior with a Poisson counting process.[9] A Poisson process has the following properties. One, during a small interval of time dt, at most one event can occur with probability very close to 1. Two, the information set up to time t does not help to predict the occurrence (or the nonoccurrence) of the event in the next time interval. Lastly, the events occur at a constant rate λ. In fact, the Poisson process is the only process that satisfies all these conditions simultaneously, thereby, making it a good candidate for modeling jump discontinuities.[10] We do, however, need to make some modifications.

In the absence of arbitrage, market equilibrium suggests that we can find a stochastic process such that all properly discounted asset prices behave as martingales. Because of this, martingales have a fundamental role to play in practical asset pricing. Currently, we are interested in modeling discrete random jumps in asset prices, while maintaining the property of a martingale. When modeling asset prices in continuous time we presume that the total change in an asset's price is composed of two terms. One is the conditional expected drift the other is the unpredicted movements. These unpredictable movements are labeled innovations and their expected value is equal to zero. In the earlier models, these innovations are modeled with a Weiner process, $dZ(t)$. Now we need a process to model innovation terms stemming from discrete random jumps. The process $dq(t)$ in expression (191) seems to be an appropriate candidate for modeling jumps but it has a nonzero mean. That is,

$$E_t[dq(t)] = \lambda dt 1 + (1 - \lambda dt)0 = \lambda dt . \qquad (192)$$

We do not know when the jump will occur, and sometimes we will not know the size of the jump, but this does not mean that we do not have some expectation about the jump. In the expression above, we see there is an expected movement in the asset price associated with the jump. This means $dq(t)$ is not completely unpredictable, thereby; it is not a candidate for being an innovation term used in pricing assets. Therefore, if we plan to use a jump process $dq(t)$ to model random jumps in the spot price we need to modify $dq(t)$ to take out its trend. Consider the modified variable

$$N(t) = \left(q(t) - \lambda t\right). \tag{193}$$

The increments $dN(t)$ will have a zero mean. For instance, if

$$dN(t) = \left(dq(t) - \lambda dt\right). \tag{194}$$

then the expected value is

$$\begin{aligned} E_t\left[dN(t)\right] &= E_t\left[dq(t) - \lambda dt\right], \\ &= E_t\left[dq(t)\right] - E_t\left[\lambda dt\right], \\ &= E_t\left[dq(t)\right] - \lambda dt, \\ &= \left[1\lambda dt + 0(1 - \lambda dt)\right] - \lambda dt, \\ &= \lambda dt - \lambda dt. \end{aligned} \tag{195}$$

Hence, $dN(t)$ is an appropriate candidate to represent unexpected jumps in asset prices.

The example above considers the impact of a Poisson process characterized by jumps of size one. Let us relax this assumption and consider the influence of a random jump size. We assume the following structure for jumps. Between jumps, $N(t)$ remains constant. At jump times $t = 1,2,3,\cdots$ it varies by some discrete and random amount. We assume that there are k possible types of jumps with sizes denoted by $a_i, i = 1,2,\cdots,k$. Again the jumps occur at a rate of λ. Once a jump occurs the jump size is selected randomly and independently. The probability of selecting a particular jump size is given by p_i.

The expected random jump size is stated as

$$E[A] = \sum_{i=1}^{k} a_i p_i \,. \tag{196}$$

Now, the modified variable $N(t)$ is redefined and given as

$$N(t) = q(t) - \lambda\left(\sum_{i=1}^{k} a_i p_i\right)t \,, \tag{197}$$

and the increment of $N(t)$ is given by

$$dN(t) = dq(t) - \lambda \sum_{i=1}^{k} a_i p_i dt \,. \tag{198}$$

Taking the expected value of expression (198), yields

$$E_t[dN(t)] = E_t[dq(t)] - \lambda \sum_{i=1}^{k} a_i p_i dt \,.$$

The expected movement for $N(t)$ is equal to the linear combination of the expected change in $q(t)$ and a constant. The expected change in $q(t)$, $E_t\,[dq(t)]$, is a function of the expected arrival of the jump, λ, and the expected size of the jump, $E[A]$. Expression (196) gives the expected size of the jump for $q(t)$. Considering these observations, we may rewrite the expectation $E_t[dq(t)]$. Rewriting and substituting into the above yields

$$E_t[dN(t)] = [\lambda E[A]dt + 0(1 - \lambda dt)] - \lambda \sum_{i=1}^{k} a_i p_i dt$$

$$= \lambda \sum_{i=1}^{k} a_i p_i dt - \lambda \sum_{i=1}^{k} a_i p_i dt \,. \tag{199}$$

We see from above, that the introduction of a random jump size alters the form of the analysis, but not the essence of the analysis. The variable $dN(t)$ is still a viable candidate for modeling discrete random jumps in asset prices, and in particular, it allows for random jump sizes. On a last note, the random jump size need not follow a discrete distribution. We could allow the random jump sizes to follow a continuous distribution without changing the analysis as well.

Given the above, we can observe how Merton constructs a model for options that incorporates both normal and rare events. The normal events are governed by the continuous Gauss Wiener process and the rare events by a Poisson process. Formally, we state this dynamic as

$$\frac{dS(t)}{S(t)} = \alpha dt + \sigma dZ(t) + dN(t). \tag{200}$$

Substituting in for $dN(t)$ we get

$$\frac{dS(t)}{S(t)} = \alpha dt + \sigma dZ(t) + dq(t) - \lambda\bar{k}dt,$$

$$= (\alpha - \lambda\bar{k})dt + \sigma dZ(t) + dq(t), \tag{201}$$

where α is the instantaneous expected return on the stock. σ is the diffusion parameter for stock price conditional on no jumps occurring. dZ is the increment of a standard Brownian motion. q is a Poisson process. λ is the mean number of arrivals per unit of time and $\bar{k}$ denotes the expected percentage change in the asset's price due to the occurrence of a jump. In terms of the examples above, the jump sizes are in returns and $\bar{k} = E[A]$, where $a_i \ni A$ for $i = 1,2,\cdots k$. In addition, the jump process and the Wiener process are statistically independent at every instant t.

To understand expression (201), we may look at the equation a little differently. We may write (201) in a more cumbersome form as

$$\frac{dS(t)}{S(t)} = (\alpha - \lambda\bar{k})dt + \sigma dZ(t) \qquad \text{if a Poisson event does not occur,} \tag{202a}$$

$$= (\alpha - \lambda\bar{k})dt + \sigma dZ(t) + (Y-1) \quad \text{if a Poisson event does occur.} \tag{202b}$$

Note the last term in equation (202b). Merton labels (Y-1) as the impulse response function. This expression characterizes the percentage change in an asset's price the instance a jump is witnessed. We recall that the expected movements in the price returns for $S(t)$ are composed of both α, the instantaneous drift and $\lambda\bar{k}$, the expected discrete change. As we move through time the spot price is trending according to α and $\lambda\bar{k}$. The movements in the spot price given by the last term in (202b) is the actual percentage change in $S(t)$ specific to a Poisson event occurring.

To illustrate the notion above, consider the following. Let X be a function of the spot price. Applying Ito's lemma for a jump-diffusion, the increment of X may be written as

$$dX = X_s dS_{dq=0} + \frac{1}{2} X_{ss} dS^2_{dq=0} + X(Y) - X \,.$$

Let $X = \ln S$, then the above becomes

$$dX = \frac{1}{S}\left((\alpha - \lambda\bar{k})Sdt + \sigma S dZ(t)\right) - \frac{1}{2S^2}\sigma^2 S^2 dt + \ln SY - \ln S \,,$$

$$dX = (\alpha - \lambda\bar{k})dt + \sigma dZ(t) - \frac{1}{2}\sigma^2 dt + \ln\left(\frac{SY}{S}\right),$$

$$dX = \left(\left(\alpha - \frac{1}{2}\sigma^2\right) - \lambda\bar{k}\right)dt + \sigma dZ(t) + \ln(Y),$$

$$\int_t^T dX = \int_t^T \left(\left(\alpha - \frac{1}{2}\sigma^2\right) - \lambda\bar{k}\right)ds + \sigma\int_t^T dZ(s) + \sum_{j=1}^{n}\ln(Y_j),$$

$$X(T) - X(t) = \left(\left(\alpha - \frac{1}{2}\sigma^2\right) - \lambda\bar{k}\right)(T-t) + \sigma\int_t^T dZ(s) + \sum_{j=1}^{n}\ln(Y_j),$$

$$X(T) = X(t) + \left(\left(\alpha - \frac{1}{2}\sigma^2\right) - \lambda\bar{k}\right)(T-t) + \sigma\int_t^T dZ(s) + \sum_{j=1}^{n}\ln(Y_j),$$

$$S(T)=S(t)\exp\left\{\left(\left(\alpha-\frac{1}{2}\sigma^{2}\right)-\lambda\bar{k}\right)(T-t)+\sigma\int_{t}^{T}dZ(s)+\sum_{j=1}^{n}\ln\left(Y_{j}\right)\right\}. \quad (203)$$

The above analysis considers the movements in the spot price given n jumps have occurred. Now lets consider the movement of the stock price over just one interval of time. That is, let $t = 0$ and $T = 1$. The expression above becomes

$$S(1)=S(0)\exp\left\{\left(\left(\alpha-\frac{1}{2}\sigma^{2}\right)-\lambda\bar{k}\right)(1-0)+\sigma\int_{0}^{1}dZ(s)+\ln Y\right\}. \quad (204)$$

The one period return is

$$\frac{dS(0)}{S(0)}=\frac{S(0)\exp\left\{\left(\left(\alpha-\frac{1}{2}\sigma^{2}\right)-\lambda\bar{k}\right)(1-0)+\sigma\int_{0}^{1}dZ(s)+\ln Y\right\}-S(0)}{S(0)}. \quad (205)$$

The one period return is influence by the instantaneous movements in the spot price and the discrete random jump. If we restrict the instantaneous movements to be zero then we have

$$=\frac{S(0)Y-S(0)}{S(0)},$$

$$=\frac{S(0)(Y-1)}{S(0)},$$

$$=(Y-1). \quad (206)$$

Expression (206) shows the percentage change in the spot price due to a random jump occurring over an interval of time. Now if $Y = 1$ then expression (206) would equal zero indicating no jump has occurred over the interval, thereby, leaving the spot process to trend according to it instantaneous drift and diffusion. If, however, a jump has occurred Y will take a value other than one and the actual percentage change in the spot price is given by $(Y–1)$. Therefore we may define k as $k \equiv Y-1$ and $\bar{k}$ as $\bar{k} \equiv E[Y-1]$.

Given the return dynamics of the spot price, we want to investigate the impact it has on a call option written on the underlying security. If the option is twice-continuously differentiable function of the stock price and time, namely, $C(SY,t)$, then the option return dynamics can be written in a similar form:

$$\frac{dC}{C(SY,t)} = (\alpha_c - \lambda\bar{k}_c)dt + \sigma_c dZ + dq_c . \tag{207}$$

In expression (207), the expectation on the jump component is taken with respect to the size and λdt represents the probability that the jump occurs over the infinitesimal interval *dt*.

Equation (207) is a general expression for the dynamics of an option's price. We now formally state this price dynamic. Using Ito's lemma for the continuous part and an analogous lemma for the jump component, the increment of the option price is written as

$$dC = \frac{1}{2}C_{ss}\left[dS_{dq=0}\right]^2 + C_s dS_{dq=0} + C_t dt + C(SY,t) - C(S,t) . \tag{208}$$

The first three expressions in equation (208) are the changes in $C(SY,t)$ due to continuous movements in the spot price. The last two terms in equation (208) represent the discrete jump in $C(SY,t)$. Now, substituting in for $\left[dS_{dq=0}\right]^2$ and rearranging yields

$$dC = \left[\frac{1}{2}C_{ss}\sigma^2 S^2 + C_t\right]dt + C_s dS_{dq=0} + C(SY,t) - C(S,t) . \tag{209}$$

Our goal is to formally state the stochastic differential equation for an option contract. This process will have a drift term and two innovation terms. The first innovation term introduces small random movements from continuous price vibrations the other innovation term introduces large price vibrations from discrete jumps. Observing expression (209), the last two terms illustrate the actual level change in the option's price due to a jump. We can convert this change into an innovation by taking the trend out of the jump term in equation (209). To do this we add and subtract the mean jump size to expression (209) and obtain

$$dC = \left[\frac{1}{2}C_{ss}\sigma^2 S^2 + C_t\right]dt + C_s dS_{dq=0}$$

$$+ C(SY,t) - C(S,t) - \lambda E\left[C(SY,t) - C(S,t)\right]dt + \lambda E\left[C(SY,t) - C(S,t)\right]dt. \quad (210)$$

Rearranging the above yields

$$dC = \left[\frac{1}{2}C_{ss}\sigma^2 S^2 + \lambda E\left[C(SY,t) - C(S,t)\right]dt + C_t\right]dt + C_s dS_{dq=0}$$

$$+ C(SY,t) - C(S,t) - \lambda E\left[C(SY,t) - C(S,t)\right]dt \quad (211)$$

Expression (211) is becoming more representative of stochastic process for the option price. That is, the last three terms in expression (211) represent the unpredictable changes in an option's price due to witnessing a jump occur. These movements are called the jump innovations. Note, that even though we took the trend out of the actual change in the option's price caused by the jump the expected change in the option's price is still considered in the analysis. This expression is now in the drift term. Finishing the stochastic differential we substitute in for $dS_{dq=0}$ to obtain

$$dC = \left[\frac{1}{2}C_{ss}\sigma^2 S^2 + \lambda E\left[C(SY,t) - C(S,t)\right] + C_t\right]dt + C_s\left[(\alpha - \lambda\bar{k})Sdt + \sigma S dZ(t)\right]$$

$$+ C(SY,t) - C(S,t) - \lambda E\left[C(SY,t) - C(S,t)\right]dt.$$

$$= \left[\frac{1}{2}C_{ss}\sigma^2 S^2 + C_s(\alpha - \lambda\bar{k})S + \lambda E\left[C(SY,t) - C(S,t)\right] + C_t\right]dt + C_s\sigma S dZ(t)$$

$$+ C(SY,t) - C(S,t) - \lambda E\left[C(SY,t) - C(S,t)\right]dt. \quad (212)$$

Dividing the left-hand side and the right-hand side of the above by $C(SY,t)$ yields

$$\frac{dC}{C(SY,t)} = \frac{\left[\frac{1}{2}C_{ss}\sigma^2 S^2 + C_s(\alpha - \lambda\bar{k})S + \lambda E[C(SY,t) - C(S,t)] + C_t\right]}{C(SY,t)}dt + \frac{C_s\sigma S}{C(SY,t)}dZ(t)$$

$$+ \frac{C(SY,t) - C(S,t)}{C(SY,t)} - \frac{\lambda E[C(SY,t) - C(S,t)]}{C(SY,t)}dt. \qquad (213)$$

Define

$$\alpha_c = \frac{\left[\frac{1}{2}C_{ss}\sigma^2 S^2 + C_s(\alpha - \lambda\bar{k})S + \lambda E[C(SY,t) - C(S,t)] + C_t\right]}{C(SY,t)}, \qquad (214)$$

$$\sigma_c = \frac{C_s\sigma S}{C(SY,t)}, \qquad (215)$$

$$\bar{k}_c = \frac{E[C(SY,t) - C(S,t)]}{C(SY,t)} \qquad (216)$$

$$dq_c = \frac{C(SY,t) - C(S,t)}{C(SY,t)}. \qquad (217)$$

Substituting expressions (214) – (217) into (213) yields

$$\frac{dC}{C(SY,t)} = \alpha_c dt + \sigma_c dZ + dq_c - \lambda\bar{k}_c dt,$$

$$= (\alpha_c - \lambda\bar{k}_c)dt + \sigma_c dZ + dq_c. \qquad (218)$$

Equation (218) is equal to equation (207), and we have formally defined the coefficients.

Expression (218) shows the drift term for the option price is broken into two parts. The first component, α_c, is expected instantaneous movements in the stock price generated by the Gaussian process and the second component, $\lambda\bar{k}_c$, is the expected discrete change due to the jump process. Furthermore, the expected discrete change from the jump involves two parts. The first is the expected arrival time of the jump, λ, and the second is the mean size of the jump denoted by $\bar{k}_c$. Therefore, to determine the expected movement in an option's price, we need the instantaneous drift, the expected time of the jump, and the expected size of the jump.

Now consider the following investment strategy where we hold the stock, the option and a risk-free asset as we did in the Black-Scholes analysis. The return dynamics for the portfolio are

$$\frac{dA}{A} = (\alpha_A - \lambda\bar{k}_A)dt + \sigma_A dZ_A + dq_A. \tag{219}$$

Again, since the portfolio is a linear combination of the assets, the instantaneous return to the portfolio is equal to the linear combination of those assets' instantaneous returns. That is,

$$\alpha_A = w_1\alpha + w_2\alpha_c + w_3 r.$$

Since the weights of a portfolio must sum to one, we rewrite the above as

$$\alpha_A = w_1(\alpha - r) + w_2(\alpha_c - r) + r. \tag{220}$$

The diffusion coefficient is equal to

$$\sigma_A = w_1\sigma + w_2\sigma_c. \tag{221}$$

Equations (220) and (221) give the instantaneous drift and diffusion for the portfolio. The last term we need to specify is the jump term in equation (219). From expression (202b) we see that when a Poisson event occurs we have the term (Y-1). In similar form to expressions (202a) and (202b), we may express equation (218) as

$$\frac{dC}{C} = (\alpha_c - \lambda\bar{k}_c)dt + \sigma_c dZ(t) \tag{222a}$$

$$= (\alpha_c - \lambda\bar{k}_c)dt + \sigma_c dZ(t) + (Y_c - 1), \tag{222b}$$

where equation (222a) states the option dynamics if no Poisson event occurs and equation (222b) states the dynamics if a Poisson event occurs. In addition, the portfolio's dynamic may also be written in the same form as the spot and option price. This is,

$$\frac{dA}{A} = (\alpha_A - \lambda\bar{k}_A)dt + \sigma_A dZ_A(t) \tag{223a}$$

$$= (\alpha_A - \lambda\bar{k}_A)dt + \sigma_A dZ_A(t) + (Y_A - 1), \tag{223b}$$

where equation (223a) states the portfolio's dynamics if no Poisson event occurs and equation (223b) states the dynamics if a Poisson event occurs. From expression (202b) and (223b), the jump component for the portfolio equals

$$(Y_A - 1) = w_1(Y - 1) + w_2(Y_c - 1) \ . \tag{224}$$

Substituting in for $(Y_c - 1)$ yields

$$(Y_A - 1) = w_1(Y - 1) + w_2\big(C(SY,t) - C(S,t)\big)/C(SY,t). \tag{225}$$

We have characterized the terms for the dynamics of the portfolio and now want to find an investment strategy that makes the portfolio risk-free. If we make the portfolio risk-free then the return to the portfolio must equal the risk-free rate of return. This is the method we followed in the Black-Scholes analysis.

Recall, we obtained an expression for the portfolio's instantaneous return and showed that it equals

$$\mu_A - r = w_1(\mu - r) + w_2(\mu_c - r) = 0.$$

The diffusion term for the portfolio equals

$$\sigma_A = w_1\sigma + w_2\sigma_c = 0.$$

This homogeneous system of equations gives us the Black-Scholes weighting scheme, where the weights equal

$$w_1^* = \frac{\mu - r}{\sigma} \text{ and } w_2^* = \frac{\mu_c - r}{\sigma_c}.$$

Under current assumptions, the Black-Scholes weighting scheme, unfortunately, will not make the current portfolio with jumps riskless. To see this, we need to take a closer look at equation (224). Inspection of expression (224)

$$(Y_A - 1) = w_1(Y - 1) + w_2\big(C(SY,t) - C(S,t)\big)/C(SY,t), \tag{226}$$

shows there does not exist a set of portfolio weights that will eliminate the 'jump' risk. For example, suppose the weights are set equal to the Black-Scholes model, then the return characteristics for equation (219) would be

$$\frac{dA^*}{A^*} = (\alpha_A^* - \lambda \bar{k}_A^*)dt + \sigma_A^* dZ_A + dq_A^*.$$

The asterisks indicates the portfolio under the Black-Scholes weighting scheme. With the Black-Scholes weights, w_1^* and w_2^*, the diffusion term, σ_A, is equal to zero. Therefore, we have

$$\frac{dA^*}{A^*} = (\alpha_A^* - \lambda \bar{k}_A^*)dt + dq_A^* \,. \tag{227}$$

Expression (227) is void of the Brownian motion term because the Black-Scholes weights eliminate the uncertainty associated with the diffusion term in equation (201). These weights, however, do not eliminate the uncertainty associated with the jump process. Recall that

$$w_1^*\sigma + w_2^*\sigma_c = 0.$$

This implies

$$w_1^* = -\frac{w_2^*\sigma_c}{\sigma}. \tag{228}$$

Using equations (169) in (167) yields

$$(Y_A^* - 1) = -\frac{w_2^* \sigma_c}{\sigma}(Y-1) + w_2^*\big(C(SY,t) - C(S,t)\big)/C(SY,t). \tag{229}$$

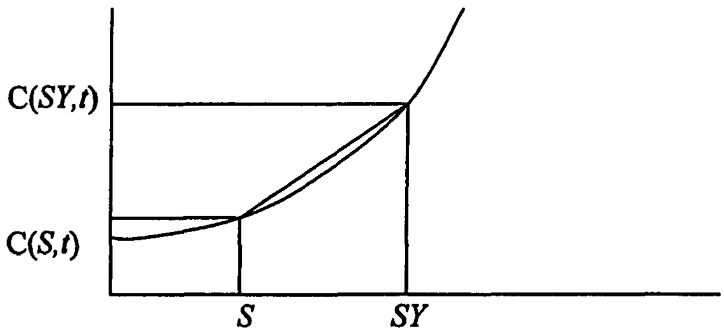
Substituting the expression for the option's diffusion term in expression (215) into the above yields

$$\begin{aligned}(Y_A^* - 1) &= -\frac{w_2^*\big(\sigma S C_s / C(SY,t)\big)}{\sigma}(Y-1) + w_2^*\big(C(SY,t) - C(S,t)\big)/C(SY,t),\\ &= -w_2^*\big(SC_s/C(SY,t)\big)(Y-1) + w_2^*\big(C(SY,t) - C(S,t)\big)/C(SY,t),\\ &= -w_2^*\big(SC_s(Y-1)/C(SY,t)\big) + w_2^*\big(C(SY,t) - C(S,t)\big)/C(SY,t),\\ &= w_2^*\big(\big(C(SY,t) - C(S,t)\big) - SC_s(Y-1)\big)/C(SY,t),\\ &= w_2^*\big(\big(C(SY,t) - C(S,t)\big) - SC_s(Y-1)\big)/C(SY,t). \end{aligned} \tag{230}$$

Now by strict convexity of the option price in the stock price,

$$C(SY,t) - C(S,t) - C_s(SY - S),$$

is positive for every value of Y. That is,



The condition for strict convexity states

$$\frac{C(SY,t) - C(S,t)}{S(Y-1)} > C_s(S,t),$$

$$C(SY,t) - C(S,t) > C_s(S,t)S(Y-1),$$

$$C(SY,t) - C(S,t) - C_s(S,t)(SY - S) > 0 . \tag{231}$$

This is the statement above. If w_2^* is positive, then $(Y_A^* - 1)$ will be positive and the unanticipated return on the portfolio will always be positive. If w_2^* is negative then the unanticipated return will always be negative. In short, there will always be some arbitrage opportunity. If we cannot eliminate the risk associated with the jump diffusion we cannot make the return to the portfolio equal to the risk-free rate of return.

Currently the problem we are faced with is to find a partial differential equation for a contingent claim written on an underlying asset whose price follows a mixed jump-diffusion process. We can construct a partial differential equation for a call option, however, we need the trend for the equation to equal the risk-free rate of return.[11] To do this, Merton turns towards the original Black and Scholes (1973) model. In their original work Black and Scholes show that their result is obtainable from both a partial equilibrium model, or in a general equilibrium under the Capital Asset Pricing Model (CAPM). Merton assumes that the CAPM holds, thereby, taking advantage of this model. The CAPM analysis presumes that stock price movements are a function of two components, idiosyncratic and non-idiosyncratic risks. Identifying this result, Merton posits that information from a market arrives one of two ways. The first is a continuous stream of news where this information has a marginal impact on stock prices. The second is the introduction of sudden non-trivial news that has an instantaneous, non-marginal impact on stock prices. If the latter type information is firm specific, then it may have little impact on stocks in general. These sudden jumps in prices from firm specific news are then deemed to be idiosyncratic and can therefore be diversified away.

In return, the jump component will be uncorrelated with the market portfolio, and therefore does not need to be priced.

Returning to the portfolio in expression (168), the only source of uncertainty in the return dynamics is the jump component. Under the CAPM, the jump represents only non-systematic risk, and therefore, is ignored. Thus, under the CAPM, the return dynamics of the portfolio are

$$\frac{dA}{A} = \alpha_A dt . \tag{232}$$

The portfolio is risk-free and has a beta of zero. Zero beta portfolios have a return equal to the risk-free rate of return. This implies $\alpha_A = r$. Substituting in for α_A yields

$$\alpha_A = w_1 \alpha + w_2 \alpha_c + w_3 r = r .$$

Rearranging the above yields

$$\alpha_A = w_1 + w_2 \alpha_c + [1 - w_1 - w_2] r = r$$

$$\alpha_A = w_1 (\alpha - r) + w_2 (\alpha_c - r) + r = r$$

$$\alpha_A - r = w_1 (\alpha - r) + w_2 (\alpha_c - r) = 0 . \tag{233}$$

Recall from the Black-Scholes methodology in section 4.1 of this chapter, the portfolio's diffusion term is given by equation (11) and is

$$\sigma_A = w_1 \sigma + w_2 \sigma_c = 0 .$$

Thus, solving the homogeneous system of equations, we obtain portfolio weights equal to

$$w_1^* = \frac{\alpha - r}{\sigma}, \text{ and } w_2^* = \frac{\alpha_c - r}{\sigma_c},$$

Black and Scholes show that these portfolio weights are equal. This is

$$\frac{\alpha - r}{\sigma} = \frac{\alpha_c - r}{\sigma_c} \tag{234}$$

Making the necessary substitutions into equation (234), we obtain

$$\frac{\alpha - r}{\sigma} = \frac{\left[\left[\frac{1}{2}C_{ss}\sigma^2 S^2 + C_s(\alpha - \lambda\bar{k})S + C_t + \lambda E[C(SY,t) - C(S,t)]\right] \Big/ C(SY,t)\right] - r}{\sigma S C_s / C(SY,t)},$$

$$\frac{(\alpha - r)\sigma S C_s / C(SY,t)}{\sigma} = \frac{\left[\frac{1}{2}C_{ss}\sigma^2 S^2 + C_s(\alpha - \lambda\bar{k})S + C_t + \lambda E[C(SY,t) - C(S,t)]\right]}{C(SY,t)} - r,$$

$$(\alpha - r)SC_s = \frac{1}{2}C_{ss}\sigma^2 S^2 + C_s(\alpha - \lambda\bar{k})S + C_t + \lambda E[C(SY,t) - C(S,t)] - rC(SY,t)$$

$$\frac{1}{2}C_{ss}\sigma^2 S^2 + C_s(r - \lambda\bar{k})S + C_t + \lambda E[C(SY,t) - C(S,t)] - rC(SY,t) = 0. \tag{235}$$

Equation (235) is the partial differential equation for a contingent claim written on a stock whose price follows a mixed-diffusion process. This expression does not depend on the true mean return for the stock. Instead, as in the Black-Scholes case, the mean is equal to the known rate of interest, r. Moreover, (233) reduces to the Black-Scholes equation if $\lambda = 0$. It is important to note that even though the jumps represent pure non-systematic risk, the jump component does affect the equilibrium option price. This is seen by the fourth term in expression (235). Inspection of the stochastic process in expression (201) indicates that this is true. That is, the expression shows the stock price follows a mixed jump diffusion process. Press (1967) shows that the distribution for any stock price at some time T will have a mixed Poisson-Gaussian distribution as opposed to having just a Gaussian distribution like the Black-Scholes model. We should expect the option price to be influenced by these jumps and from expression (235) we see this is correct.

Merton has developed the partial differential equation for an option whose price incorporates idiosyncratic jumps. To find the value of the call option we need the boundary conditions. These are

$$C(0,\tau) = 0, \tag{236}$$

$$C(SY,0) = \max[0, SY - X]. \tag{237}$$

Given (174), (175) and (176) the value of call option is

$$C(SY,t) = \sum_{n=0}^{\infty} \frac{e^{-\lambda\tau}(\lambda\tau)^n}{n!} [E_n(C(SY_n,t))], \tag{238}$$

where $E_n(C(SY_n,t))$ is the solution to the Black-Scholes option model conditional on n jumps occurring over the life of the option, and τ is the time to maturity.[12] The random variable, Y_n, is defined as the product of n independently and identically distributed random variables, each being identically distributed to the random variable Y in expression (202b). E_n is the expectation operator over the distribution of Y_n.

Expression (238) is a general solution for an option written on a stock whose price follows a mixed jump diffusion process. There is a special case for Merton's model where a closed form solution exists. The closed form solution is found when Y_n (where $Y_n = \prod_{j=1}^{n} Y_j$) is jointly log-normally distributed. Under this assumption we have the following. Recall $\bar{k} \equiv E[Y-1]$. This implies $Y = 1 + k$. If we let Y be log-normally distributed with mean $\left(\gamma - \frac{1}{2}\eta^2\right)$ and a variance of η^2 then the log of Y, $\ln Y = \ln(1+k)$, is said to be normally distributed with mean γ and a variance of η^2.[13] The product of n independently distributed log-normal variables $Y_n = \prod_{j=1}^{n} Y_j$ is jointly log-normal with a mean $n\left(\gamma-\frac{1}{2}\eta^2\right)$ and a variance $n\eta^2$. Given these distributional characteristics, we can find a closed form solution for expression (238).

Expression (238) is a Poisson weighted sum of the Black-Scholes option prices conditional on n jumps occurring. If we can find a solution to the Black-Scholes expression then we can find a closed form solution for expression (238). Consider the following

$$\frac{dS(t)}{S(t)}=(r-\lambda\bar{k})dt+\sigma dZ(t)+dq(t). \tag{239}$$

Note, the instantaneous drift for the spot price has changed from α to r. This is due to Merton treating the jumps as idiosyncratic risk and making the return to the arbitrage portfolio equal to the riskless rate of return. Let $X=\ln S$. With this transformation and applying Ito's lemma for a jump diffusion the stochastic differential for X may be written as

$$dX=X_s dS_{dq=0}+\frac{1}{2}X_{ss}dS^2_{dq=0}+\ln SY-\ln S,$$

$$=\frac{1}{S}\left((r-\lambda\bar{k})Sdt+\sigma SdZ(t)\right)-\frac{1}{2S^2}\sigma^2S^2dt+\ln SY-\ln S,$$

$$=(r-\lambda\bar{k})dt+\sigma dZ(t)-\frac{1}{2}\sigma^2dt+\ln\left(\frac{SY}{S}\right),$$

$$=\left(\left(r-\frac{1}{2}\sigma^2\right)-\lambda\bar{k}\right)dt+\sigma dZ(t)+\ln(Y),$$

$$\int_t^T dX=\int_t^T\left(\left(r-\frac{1}{2}\sigma^2\right)-\lambda\bar{k}\right)ds+\sigma\int_t^T dZ(s)+\sum_{j=1}^{n}\ln(Y_j),$$

$$X(T)-X(t)=\left(\left(r-\frac{1}{2}\sigma^2\right)-\lambda\bar{k}\right)(T-t)+\sigma\int_t^T dZ(s)+\sum_{j=1}^{n}\ln(Y_j),$$

$$X(T)=X(t)+\left(\left(r-\frac{1}{2}\sigma^2\right)-\lambda\bar{k}\right)(T-t)+\sigma\int_t^T dZ(s)+\sum_{j=1}^{n}\ln(Y_j),$$

$$S(T)=S(t)\exp\left\{\left(\left(r-\frac{1}{2}\sigma^2\right)-\lambda\bar{k}\right)(T-t)+\sigma\int_t^T dZ(s)+\sum_{j=1}^{n}\ln(Y_j)\right\},$$

$$S(T)=S(t)\exp\{B(T)\}, \qquad (240)$$

where $B(T)=\left(\left(r-\frac{1}{2}\sigma^2\right)-\lambda\bar{k}\right)(T-t)+\sigma\int_t^T dZ(s)+\sum_{j=1}^{n}\ln(Y_j)$.

The terminal stock price depends on the instantaneous movements and the discrete random jumps. The expected value of the spot price will then depend on the instantaneous return, the average jump size and the probability of n jumps occurring. The infinite sum in expression (238) considers the probability of n possible jumps occurring in the spot price over the time interval considered. Therefore, we are left with finding the expected terminal stock price given the distributional characteristics of the instantaneous movements and for the size of the jump. This is

$$E_t^*[S(T)\mid n\ jumps]=\exp\left(E_t^*[X(T)\mid n\ jumps]+\frac{1}{2}V_t^*[X(T)\mid n\ jumps]\right). \qquad (241)$$

Thus, the mean and variance of $X(T)$ are

$$E_t^*[X(T)\mid n\ jumps]=X(t)+\left(\left(r-\frac{1}{2}\sigma^2\right)-\lambda\bar{k}\right)(T-t)+\sum_{j=1}^{n}E_n[\ln(Y_j)],$$

$$=X(t)+\left(\left(r-\frac{1}{2}\sigma^2\right)-\lambda\bar{k}\right)(T-t)+\sum_{j=1}^{n}\left(\gamma-\frac{1}{2}\eta^2\right),$$

$$=X(t)+\left(\left(r-\frac{1}{2}\sigma^2\right)-\lambda\bar{k}\right)(T-t)+n\left(\gamma-\frac{1}{2}\eta^2\right),$$

$$= X(t) + \left(r - \lambda \bar{k} + \frac{n\gamma}{\tau} \right)\tau - \frac{1}{2}\left(\sigma^2 \tau + n\eta^2\right) \qquad (242)$$

$$V_t^*[X(T) \mid n\ jumps] = \sigma^2 \int_t^T ds + \sum_{j=1}^{n} V_n\left[\ln(Y_j)\right],$$

$$= \sigma^2 (T - t) + \sum_{j=1}^{n} \eta^2 ,$$

$$= \sigma^2 \tau + n\eta^2 . \qquad (243)$$

Plugging the expected value and variance for $X(T)$ back into the expected spot price we get

$$E_t^*[S(T) \mid n\ jumps] = \exp\left(X(t) + \left(r - \lambda \bar{k} + \frac{n\gamma}{\tau} \right)\tau - \frac{1}{2}\left(\sigma^2\tau + n\eta^2\right) + \frac{1}{2}\left(\sigma^2\tau + n\eta^2\right) \right),$$

$$= \exp\left(X(t) + \left(r - \lambda \bar{k} + \frac{n\gamma}{\tau} \right)\tau \right),$$

$$= S(t)\exp\left(\left(r - \lambda \bar{k} + \frac{n\gamma}{\tau} \right)\tau \right). \qquad (244)$$

Equation (244) is the expected terminal stock price conditional on n jumps occurring.

Following the derivation in section 4.1, we can price can find the Black-Scholes solution for this model. We know from equation (20) of this chapter that the Black-Scholes formula is

$$C(S,t) = S(t)N(d_1) - e^{-r\tau} XN(d_2),$$

where

$$d_1 = \frac{\ln\left(\frac{S(t)}{X}\right) + \left(r + \frac{1}{2}\sigma^2\right)\tau}{\sigma\sqrt{\tau}},$$

and

$$d_2 = d_1 - \sigma\sqrt{\tau} .$$

We may tailor the Black-Scholes solution above to fit the characteristics of the spot price for the current model. Under a risk-neutrality the Black-Scholes option price may be expressed as

$$C(S,t \mid n) = e^{-r\tau} E_t^*\left[C(SY,T) \mid n\ jumps\right], \tag{245}$$

where $C(S,t \mid n)$ is the price of the Black-Scholes call option today, and $C(S,T \mid n)$ is the price of Black-Scholes call option at a terminal date T. Substituting in for the risk-adjusted terminal call price we get

$$\begin{aligned}
C(S,t \mid n) &= e^{-r\tau} E_t^*\left[\max(S(T) - X, 0) \mid n\ jumps\right], \\
&= e^{-r\tau} E_t^*\left[\max\left(S(t)e^{B(T)} - X, 0\right) \mid n\ jumps\right], \\
&= e^{-r\tau}\left\{E_t^*\left[S(t)e^{B(T)} - X \mid S(T) \geq X, n\ jumps\right] + E_t^*\left[0 \mid S(T) < X, n\ jumps\right]\right\} \\
&= e^{-r\tau}\left\{E_t^*\left[S(t)e^{B(T)} - X \mid S(t)e^{B(T)} \geq X, n\ jumps\right] + E_t^*\left[0 \mid S(t)e^{B(T)} < X, n\ jumps\right]\right\} \\
&= e^{-r\tau}\left\{E_t^*\left[S(t)e^{B(T)} - X \mid e^{B(T)} \geq \frac{X}{S(t)}, n\ jumps\right] + E_t^*\left[0 \mid e^{B(T)} < \frac{X}{S(t)}, n\ jumps\right]\right\} \\
&= e^{-r\tau}\left\{E_t^*\left[S(t)e^{B(T)} - X \mid B(T) \geq \ln\left(\frac{X}{S(t)}\right), n\ jumps\right] + E_t^*\left[0 \mid B(T) < \ln\left(\frac{X}{S(t)}\right), n\ jumps\right]\right\}
\end{aligned} \tag{246}$$

From expression (246), we see that the price of a call today is simply the linear combination of two partial expectations conditioned on n jumps occurring. The first partial expectation considers the option when it is in the money and the second is for the option when it is out of the money. Presuming n jumps have occurred, we may formally write the probability of the option expiring in the money. This is

$$P^*\{X \leq S(T) \leq \infty\},$$

where P^* is an equivalent martingale probability. Transforming the variable then constitutes the following

$$P^*\left\{\frac{X}{S(t)} \leq \frac{S(T)}{S(t)} \leq \infty\right\},$$

$$P^*\left\{\ln\left(\frac{X}{S(t)}\right) \leq \ln\left(\frac{S(T)}{S(t)}\right) \leq \infty\right\}.$$

Since $\ln\left(\frac{S(T)}{S(t)}\right) = B(T)$, the above becomes

$$P^*\left\{\ln\left(\frac{X}{S(t)}\right) \leq B(T) \leq \infty\right\}.$$

The second partial expectation operator in expression (246) considers the probability that the spot price is less than the exercise price at maturity. Formally, this is

$$P^*\{-\infty \leq S(T) \leq X\},$$

where P^* is an equivalent martingale probability. Again, transforming the variable then constitutes the following

$$P^*\left\{-\infty \leq \frac{S(T)}{S(t)} \leq \frac{X}{S(t)}\right\},$$

$$P^*\left\{-\infty \leq \ln\left(\frac{S(T)}{S(t)}\right) \leq \ln\left(\frac{X}{S(t)}\right)\right\},$$

$$P^*\left\{-\infty \leq B(T) \leq \ln\left(\frac{X}{S(t)}\right)\right\}.$$

Given the probability statements above, the linear combination of the partial expectations in expression (246) may be written as

$$C(S,t\mid n)=e^{-r\tau}\int_{-\infty}^{\ln\left(\frac{X}{S(t)}\right)}(0)dP^{*}+e^{-r\tau}\int_{\ln\left(\frac{X}{S(t)}\right)}^{\infty}\left(S(t)e^{b(T)}-X\right)dP^{*} \tag{247}$$

where dP^{*} is the probability measure and is equal to

$$dP^{*}=\frac{1}{\sqrt{2\pi\left(\sigma^{2}\tau+n\eta^{2}\right)}}\exp\left[-\frac{\left(b(T)-\left(\left(r-\lambda\bar{k}+\frac{n\gamma}{\tau}\right)\tau-\frac{1}{2}\left(\sigma^{2}+\frac{n\eta^{2}}{\tau}\right)\tau\right)\right)^{2}}{2\left(\sigma^{2}\tau+n\eta^{2}\right)}\right]db(T).$$

The use of the probability measure dP^{*} in expression (247) is the same as the analysis in section 4.1 expression (33), with one fundamental difference. The probability measure in expression (247) is not completely risk adjusted. The spot price is said to trend at the known riskless rate of return because of the assumptions made by Merton. Merton presumes that the jumps are diversifiable, thereby obtaining a zero beta portfolio. Since this portfolio trends at the known riskless rate we develop the model in a pseudo risk neutral world. The jump risk is ignored but still impacts the options equilibrium price. In the probability measure, we see that the actual jump parameters are used and not the risk-adjusted parameters. Hence, we are working with a pseudo risk-neutral measure.

Continuing with expression (247), we substitute in the probability measure and obtain (for ease of exposition, the brackets of the exponential are acknowledged with a dot)

$$C(S,t\mid n)=e^{-r\tau}\int_{-\infty}^{\ln\left(\frac{X}{S(t)}\right)}(0)\frac{1}{\sqrt{2\pi\left(\sigma^{2}\tau+n\eta^{2}\right)}}\exp[\cdot]db(T)$$

$$+e^{-r\tau}\int_{\ln\left(\frac{X}{S(t)}\right)}^{\infty}\frac{\left(S(t)e^{b(T)}-X\right)}{\sqrt{2\pi\left(\sigma^2\tau+n\eta^2\right)}}\exp[\cdot]db(T) \tag{248}$$

The first term in expression (248) is equal to zero leaving us with

$$C(SY,t)=e^{-r\tau}\int_{\ln\left(\frac{X}{S(t)}\right)}^{\infty}\left(S(t)e^{b(T)}-X\right)\frac{1}{\sqrt{2\pi\left(\sigma^2\tau+n\eta^2\right)}}\exp[\cdot]db(T)\,. \tag{249}$$

Expression (249) can be split up into two different integrals.

$$\int_{\ln\left(\frac{X}{S(t)}\right)}^{\infty} e^{-r\tau}S(t)e^{b(T)}\frac{1}{\sqrt{2\pi\left(\sigma^2\tau+n\eta^2\right)}}\exp[\cdot]db(T)-\int_{\ln\left(\frac{X}{S(t)}\right)}^{\infty} e^{-r\tau}X\frac{1}{\sqrt{2\pi\left(\sigma^2\tau+n\eta^2\right)}}\exp[\cdot]db(T)$$

(250)

We start with the second integral in expression (250). Consider the transformation of $B(T)$

$$Z=\frac{B(T)-\left(\left(r-\lambda\bar{k}+\frac{n\gamma}{\tau}\right)-\frac{1}{2}\left(\sigma^2+\frac{n\eta^2}{\tau}\right)\right)\tau}{\sqrt{\left(\sigma^2\tau+n\eta^2\right)}}\,. \tag{251}$$

If we wish to write the integral in (250) as a function of Z, we must adjust the limits. Currently, we are considering the probability that the spot price is greater than the exercise price at maturity.

$$P^{*}\left\{\ln\left(\frac{X}{S(t)}\right)\leq B(T)\leq\infty\right\}. \tag{252}$$

Recall that $B(T)$ has a mean of $\left(\left(r-\lambda\bar{k}+\frac{n\gamma}{\tau}\right)-\frac{1}{2}\left(\sigma^2+\frac{n\eta^2}{\tau}\right)\right)\tau$ and a variance of $\left(\sigma^2+\frac{n\eta^2}{\tau}\right)\tau$. Subtracting the mean of $B(T)$ and dividing by the standard deviation yields

$$P^*\left\{\frac{\ln\left(\frac{X}{S(t)}\right)-E[B(T)]}{\sqrt{(\sigma^2\tau+n\eta^2)}}\leq\frac{B(T)-E[B(T)]}{\sqrt{(\sigma^2\tau+n\eta^2)}}\leq\infty\right\}. \quad (253)$$

From equation (251) we have

$$P^*\left\{\frac{\ln\left(\frac{X}{S(t)}\right)-\left(\left(r-\lambda\bar{k}+\frac{n\gamma}{\tau}\right)-\frac{1}{2}\left(\sigma^2+\frac{n\eta^2}{\tau}\right)\right)\tau}{\sqrt{(\sigma^2\tau+n\eta^2)}}\leq Z\leq\infty\right\}. \quad (254)$$

If we change the variable of integration in the above to dz, then the limits of integration for this problem become $\frac{\ln\left(\frac{X}{S(t)}\right)-\left(\left(r-\lambda\bar{k}+\frac{n\gamma}{\tau}\right)-\frac{1}{2}\left(\sigma^2+\frac{n\eta^2}{\tau}\right)\right)\tau}{\sqrt{(\sigma^2\tau+n\eta^2)}}$ and ∞, making the integral

$$-e^{-r\tau}X\int_{\frac{\ln\left(\frac{X}{S(t)}\right)-\left(\left(r-\lambda\bar{k}+\frac{n\gamma}{\tau}\right)-\frac{1}{2}\left(\sigma^2+\frac{n\eta^2}{\tau}\right)\right)\tau}{\sqrt{(\sigma^2\tau+n\eta^2)}}}^{\infty}\frac{1}{\sqrt{2\pi}}\exp\left[-\frac{1}{2}z^2\right]dz, \quad (255)$$

The normal distribution is a symmetrical distribution. A property of the standard normal distribution states that we can write

$$\int_L^{\infty}f(z)dz=\int_{-\infty}^{-L}f(z)dz. \quad (256)$$

Using the above property, expression (255) may be rewritten as

$$-e^{-r\tau}X\int_{-\infty}^{\frac{-\left(\ln\left(\frac{X}{S(t)}\right)-\left(\left(r-\lambda\bar{k}+\frac{n\gamma}{\tau}\right)-\frac{1}{2}\left(\sigma^2+\frac{n\eta^2}{\tau}\right)\right)\tau\right)}{\sqrt{\left(\sigma^2\tau+n\eta^2\right)}}}\frac{1}{\sqrt{2\pi}}\exp\left[-\frac{1}{2}z^2\right]dz \tag{257}$$

Consequently, the upper limit in expression (255) can be rewritten. Rewriting yields

$$-\frac{\ln\left(\frac{X}{S(t)}\right)-\left(\left(r-\lambda\bar{k}+\frac{n\gamma}{\tau}\right)-\frac{1}{2}\left(\sigma^2+\frac{n\eta^2}{\tau}\right)\right)\tau}{\sqrt{\left(\sigma^2\tau+n\eta^2\right)}},$$

$$-\frac{\ln X-\ln S(t)-\left(\left(r-\lambda\bar{k}+\frac{n\gamma}{\tau}\right)-\frac{1}{2}\left(\sigma^2+\frac{n\eta^2}{\tau}\right)\right)\tau}{\sqrt{\left(\sigma^2\tau+n\eta^2\right)}},$$

$$\frac{\ln S(t)-\ln X+\left(\left(r-\lambda\bar{k}+\frac{n\gamma}{\tau}\right)-\frac{1}{2}\left(\sigma^2+\frac{n\eta^2}{\tau}\right)\right)\tau}{\sqrt{\left(\sigma^2\tau+n\eta^2\right)}},$$

$$\frac{\ln\left(\frac{S(t)}{X}\right)+\left(\left(r-\lambda\bar{k}+\frac{n\gamma}{\tau}\right)-\frac{1}{2}\left(\sigma^2+\frac{n\eta^2}{\tau}\right)\right)\tau}{\sqrt{\left(\sigma^2\tau+n\eta^2\right)}}. \tag{258}$$

We define the above result as

$$d_2=\frac{\ln\left(\frac{S(t)}{X}\right)+\left(\left(r-\lambda\bar{k}+\frac{n\gamma}{\tau}\right)-\frac{1}{2}\left(\sigma^2+\frac{n\eta^2}{\tau}\right)\right)\tau}{\sqrt{\left(\sigma^2\tau+n\eta^2\right)}} \tag{259}$$

Thus, equation (255) can be expressed as

$$-e^{-r\tau}X\int_{-\infty}^{d_2}\frac{1}{\sqrt{2\pi}}\exp\left[-\frac{1}{2}z^2\right]dz. \tag{260}$$

Furthermore, expression (260) becomes

$$-e^{-r\tau}XN(d_2) \tag{261}$$

Now working on the first integral in expression (250), we start with the variable

Z. The limits for the integral are $\dfrac{\ln\left(\dfrac{X}{S(t)}\right)-\left(\left(r-\lambda\bar{k}+\dfrac{n\gamma}{\tau}\right)-\dfrac{1}{2}\left(\sigma^2+\dfrac{n\eta^2}{\tau}\right)\right)\tau}{\sqrt{\left(\sigma^2\tau+n\eta^2\right)}}$ and ∞,

therefore we have

$$\int_{\frac{\ln\left(\frac{X}{S(t)}\right)-\left(\left(r-\lambda\bar{k}+\frac{n\gamma}{\tau}\right)-\frac{1}{2}\left(\sigma^2+\frac{n\eta^2}{\tau}\right)\right)\tau}{\sqrt{\left(\sigma^2\tau+n\eta^2\right)}}}^{\infty} e^{-r\tau}S(t)e^{b(T)}\frac{1}{\sqrt{2\pi}}\exp\left[-\frac{1}{2}z^2\right]dz. \tag{262}$$

The limits for expression (262) are in terms of Z, however we still have $b(T)$ inside the integral. We need to transform the variable $B(T)$. First we recognize that the lower limit in the expression above is equal to

$$-d_2=\frac{\ln\left(\dfrac{X}{S(t)}\right)-\left(\left(r-\lambda\bar{k}+\dfrac{n\gamma}{\tau}\right)-\dfrac{1}{2}\left(\sigma^2+\dfrac{n\eta^2}{\tau}\right)\right)\tau}{\sqrt{\left(\sigma^2\tau+n\eta^2\right)}}.$$

Therefore we can express the integral as

$$\int_{-d_2}^{\infty} e^{-r\tau}S(t)e^{b(T)}\frac{1}{\sqrt{2\pi}}\exp\left[-\frac{1}{2}z^2\right]dz. \tag{263}$$

Now we start working on $B(T)$. Multiplying expression (263) by 1,

$$\exp\left(\left(\left(r-\lambda\bar{k}+\frac{n\gamma}{\tau}\right)-\frac{1}{2}\left(\sigma^2+\frac{n\eta^2}{\tau}\right)\right)\tau\right)\exp\left(-\left(\left(r-\lambda\bar{k}+\frac{n\gamma}{\tau}\right)-\frac{1}{2}\left(\sigma^2+\frac{n\eta^2}{\tau}\right)\right)\tau\right),$$

yields

$$e^{-r\tau}e^{\left(\left(\left(r-\lambda\bar{k}+\frac{n\gamma}{\tau}\right)-\frac{1}{2}\left(\sigma^2+\frac{n\eta^2}{\tau}\right)\right)\tau\right)}S(t)\int_{-d_2}^{\infty} e^{b(T)}e^{\left(-\left(\left(r-\lambda\bar{k}+\frac{n\gamma}{\tau}\right)-\frac{1}{2}\left(\sigma^2+\frac{n\eta^2}{\tau}\right)\right)\tau\right)}\frac{1}{\sqrt{2\pi}}\exp\left[-\frac{1}{2}z^2\right]dz.$$

This is equal to

$$e^{-r\tau}e^{\left(\left(\left(r-\lambda\bar{k}+\frac{n\gamma}{\tau}\right)-\frac{1}{2}\left(\sigma^2+\frac{n\eta^2}{\tau}\right)\right)\tau\right)}S(t)\int_{-d_2}^{\infty}e^{\left(b(T)-\left(\left(r-\lambda\bar{k}+\frac{n\gamma}{\tau}\right)-\frac{1}{2}\left(\sigma^2+\frac{n\eta^2}{\tau}\right)\right)\tau\right)}\frac{1}{\sqrt{2\pi}}\exp\left[-\frac{1}{2}z^2\right]dz. \quad (264)$$

We know

$$Z=\frac{B(T)-\left(\left(r-\lambda\bar{k}+\frac{n\gamma}{\tau}\right)-\frac{1}{2}\left(\sigma^2+\frac{n\eta^2}{\tau}\right)\right)\tau}{\sqrt{\left(\sigma^2\tau+n\eta^2\right)}},$$

which is equal to

$$Z\sqrt{\left(\sigma^2\tau+n\eta^2\right)}=B(T)-\left(\left(r-\lambda\bar{k}+\frac{n\gamma}{\tau}\right)-\frac{1}{2}\left(\sigma^2+\frac{n\eta^2}{\tau}\right)\right)\tau. \quad (265)$$

Substituting (265) into (264) yields

$$e^{-r\tau}e^{\left(\left(\left(r-\lambda\bar{k}+\frac{n\gamma}{\tau}\right)-\frac{1}{2}\left(\sigma^2+\frac{n\eta^2}{\tau}\right)\right)\tau\right)}S(t)\int_{-d_2}^{\infty}e^{z\sqrt{\left(\sigma^2\tau+n\eta^2\right)}}\frac{1}{\sqrt{2\pi}}\exp\left[-\frac{1}{2}z^2\right]dz. \quad (266)$$

Working on the term inside the integral in equation (266), we combine the exponents. This yields

$$e^{-r\tau}e^{\left(\left(\left(r-\lambda\bar{k}+\frac{n\gamma}{\tau}\right)-\frac{1}{2}\left(\sigma^2+\frac{n\eta^2}{\tau}\right)\right)\tau\right)}S(t)\int_{-d_2}^{\infty}\frac{1}{\sqrt{2\pi}}e^{z\sqrt{\left(\sigma^2\tau+n\eta^2\right)}}\exp\left[-\frac{1}{2}z^2\right]dz,$$

$$e^{-r\tau}e^{\left(\left(\left(r-\lambda\bar{k}+\frac{n\gamma}{\tau}\right)-\frac{1}{2}\left(\sigma^2+\frac{n\eta^2}{\tau}\right)\right)\tau\right)}S(t)\int_{-d_2}^{\infty}\frac{1}{\sqrt{2\pi}}e^{\frac{1}{2}(2)z\sqrt{\left(\sigma^2\tau+n\eta^2\right)}}\exp\left[-\frac{1}{2}z^2\right]dz,$$

$$e^{-r\tau}e^{\left(\left(\left(r-\lambda\bar{k}+\frac{n\gamma}{\tau}\right)-\frac{1}{2}\left(\sigma^2+\frac{n\eta^2}{\tau}\right)\right)\tau\right)}S(t)\int_{-d_2}^{\infty}\frac{1}{\sqrt{2\pi}}\exp\left[-\frac{1}{2}z^2+\frac{1}{2}(2)z\sqrt{\left(\sigma^2\tau+n\eta^2\right)}\right]dz,$$

$$e^{-r\tau}e^{\left(\left(\left(r-\lambda\bar{k}+\frac{n\gamma}{\tau}\right)-\frac{1}{2}\left(\sigma^2+\frac{n\eta^2}{\tau}\right)\right)\tau\right)}S(t)\int_{-d_2}^{\infty}\frac{1}{\sqrt{2\pi}}\exp\left[-\frac{1}{2}\left(z^2-2z\sqrt{\left(\sigma^2\tau+n\eta^2\right)}\right)\right]dz. \quad (267)$$

The exponential term inside the integral in expression (267) can be completed into a square. To do so we need to multiply (267) by

$$\exp\left[\frac{\left(\sigma^2\tau+n\eta^2\right)}{2}\right]\exp\left[-\frac{\left(\sigma^2\tau+n\eta^2\right)}{2}\right]. \tag{268}$$

This yields

$$e^{\frac{\left(\sigma^2\tau+n\eta^2\right)}{2}}e^{\frac{-\left(\sigma^2\tau+n\eta^2\right)}{2}}e^{-r\tau}e^{\left(\left(r-\lambda\bar{k}+\frac{n\gamma}{\tau}\right)-\frac{1}{2}\left(\sigma^2+\frac{n\eta^2}{\tau}\right)\right)\tau}S(t)\int_{-d_2}^{\infty}\frac{1}{\sqrt{2\pi}}\exp\left[-\frac{1}{2}\left(z^2-2z\sqrt{\left(\sigma^2\tau+n\eta^2\right)}\right)\right]dz$$

$$e^{-r\tau}e^{\left(\left(r-\lambda\bar{k}+\frac{n\gamma}{\tau}\right)-\frac{1}{2}\left(\sigma^2+\frac{n\eta^2}{\tau}\right)\right)\tau+\frac{1}{2}\left(\sigma^2\tau+n\eta^2\right)}S(t)\int_{-d_2}^{\infty}e^{-\frac{\left(\sigma^2\tau+n\eta^2\right)}{2}}\frac{1}{\sqrt{2\pi}}\exp\left[-\frac{1}{2}\left(z^2-2z\sqrt{\left(\sigma^2\tau+n\eta^2\right)}\right)\right]dz,$$

$$e^{-r\tau}e^{\left(\left(r-\lambda\bar{k}+\frac{n\gamma}{\tau}\right)\right)\tau}S(t)\int_{-d_2}^{\infty}\frac{1}{\sqrt{2\pi}}\exp\left[-\frac{1}{2}\left(\sigma^2\tau+n\eta^2\right)-\frac{1}{2}\left(z^2-2z\sqrt{\left(\sigma^2\tau+n\eta^2\right)}\right)\right]dz,$$

$$e^{-r\tau}S(t)e^{\left(r-\lambda\bar{k}+\frac{n\gamma}{\tau}\right)\tau}\int_{-d_2}^{\infty}\frac{1}{\sqrt{2\pi}}\exp\left[-\frac{1}{2}\left(z^2-2z\sqrt{\left(\sigma^2\tau+n\eta^2\right)}+\left(\sigma^2\tau+n\eta^2\right)\right)\right]dz,$$

or

$$e^{-r\tau}S(t)e^{\left(r-\lambda\bar{k}+\frac{n\gamma}{\tau}\right)\tau}\int_{-d_2}^{\infty}\frac{1}{\sqrt{2\pi}}\exp\left[-\frac{1}{2}\left(z-\sqrt{\left(\sigma^2\tau+n\eta^2\right)}\right)^2\right]dz. \tag{269}$$

Examining expression (269), we see that the integral in equation (269) is no longer for a standard normal variable. The mean has been moved to the left by $\sqrt{\left(\sigma^2\tau+n\eta^2\right)}$ units. To express the probability density function in terms of a standard normal variable we need to redefine the variable of integration. If we let $W=Z-\sqrt{\left(\sigma^2\tau+n\eta^2\right)}$, then we see that $\frac{dw}{dz}=1$. In addition, when $z=\infty$ then $w=\infty$, and when

$$z=\frac{\ln\left(\frac{X}{S(t)}\right)-\left(\left(r-\lambda\bar{k}+\frac{n\gamma}{\tau}\right)-\frac{1}{2}\left(\sigma^2+\frac{n\eta^2}{\tau}\right)\right)\tau}{\sqrt{\left(\sigma^2\tau+n\eta^2\right)}}$$

then

$$w=\frac{\ln\left(\frac{X}{S(t)}\right)-\left(\left(r-\lambda\bar{k}+\frac{n\gamma}{\tau}\right)-\frac{1}{2}\left(\sigma^2+\frac{n\eta^2}{\tau}\right)\right)\tau}{\sqrt{\left(\sigma^2\tau+n\eta^2\right)}}-\sqrt{\left(\sigma^2\tau+n\eta^2\right)}.$$

Therefore, using dw as the variable of integration makes expression (269) equal to

$$e^{-r\tau}S(t)e^{\left(r-\lambda\bar{k}+\frac{n\gamma}{\tau}\right)\tau}\int_{-d_2-\sqrt{\left(\sigma^2\tau+n\eta^2\right)}}^{\infty}\frac{1}{\sqrt{2\pi}}\exp\left[-\frac{1}{2}w^2\right]dw,$$

where the probability density function is now for a standard normal variable. Using the symmetry property of the normal distribution allows us to write the above as

$$e^{-r\tau}S(t)e^{\left(r-\lambda\bar{k}+\frac{n\gamma}{\tau}\right)\tau\tau}\int_{-\infty}^{d_2+\sqrt{\left(\sigma^2\tau+n\eta^2\right)}}\frac{1}{\sqrt{2\pi}}\exp\left[-\frac{1}{2}w^2\right]dw,$$

$$e^{-r\tau}S(t)e^{\left(r-\lambda\bar{k}+\frac{n\gamma}{\tau}\right)\tau}\int_{-\infty}^{d_1}\frac{1}{\sqrt{2\pi}}\exp\left[-\frac{1}{2}w^2\right]dw. \tag{270}$$

The upper limit in expression (270) is equal to

$$d_1=d_2+\sqrt{\left(\sigma^2\tau+n\eta^2\right)},$$

$$=\frac{\ln\left(\frac{S(t)}{X}\right)+\left(\left(r-\lambda\bar{k}+\frac{n\gamma}{\tau}\right)-\frac{1}{2}\left(\sigma^2+\frac{n\eta^2}{\tau}\right)\right)\tau}{\sqrt{\left(\sigma^2\tau+n\eta^2\right)}}+\sqrt{\left(\sigma^2\tau+n\eta^2\right)},$$

$$=\frac{\ln\left(\frac{S(t)}{X}\right)+\left(\left(r-\lambda\bar{k}+\frac{n\gamma}{\tau}\right)-\frac{1}{2}\left(\sigma^2+\frac{n\eta^2}{\tau}\right)\right)\tau}{\sqrt{\left(\sigma^2\tau+n\eta^2\right)}}+\frac{\left(\sigma^2\tau+n\eta^2\right)}{\sqrt{\left(\sigma^2\tau+n\eta^2\right)}},$$

$$= \frac{\ln\left(\frac{S(t)}{X}\right) + \left(\left(r - \lambda\bar{k} + \frac{n\gamma}{\tau}\right) + \frac{1}{2}\left(\sigma^2 + \frac{n\eta^2}{\tau}\right)\right)\tau}{\sqrt{\left(\sigma^2\tau + n\eta^2\right)}}. \tag{271}$$

Given the above, expression (270) reduces to

$$e^{-r\tau} S(t) e^{\left(r - \lambda\bar{k} + \frac{n\gamma}{\tau}\right)\tau} N(d_1). \tag{272}$$

Now replacing the integrals in expression (250), with expressions (261) and (271) yields

$$C(S,t \mid n) = e^{-r\tau} S(t) e^{\left(r - \lambda\bar{k} + \frac{n\gamma}{\tau}\right)\tau} N(d_1) - e^{-r\tau} X N(d_2), \tag{273}$$

where

$$d_1 = \frac{\ln\left(\frac{S(t)}{X}\right) + \left(\left(r - \lambda\bar{k} + \frac{n\gamma}{\tau}\right) + \frac{1}{2}\left(\sigma^2 + \frac{n\eta^2}{\tau}\right)\right)\tau}{\sqrt{\left(\sigma^2\tau + n\eta^2\right)}}, \tag{274}$$

and

$$d_2 = d_1 - \sqrt{\left(\sigma^2\tau + n\eta^2\right)}. \tag{275}$$

Substituting the conditioned Black-Scholes solution into Merton's expression yields a closed form solution. This is

$$C(SY,t) = \sum_{n=0}^{\infty} \frac{e^{-\lambda\tau}(\lambda\tau)^n}{n!}\left[e^{-r\tau}\left(S(t)\exp\left(\left(r - \lambda\bar{k} + \frac{n\gamma}{\tau}\right)\tau\right) N(d_1) - XN(d_2)\right)\right]. \tag{276}$$

Expression (276) is a closed form solution for an option contract written on a stock whose price follows a mixed jump-diffusion process. Furthermore, it is assumed that the jump size follows a log-normal distribution. To obtain this solution Merton follows the argument originally presented in Black and Scholes (1973) seminal work. With complications arising from discontinuous breaks in the spot price dynamics, Merton must make additional assumptions. First, he states the capital asset pricing model holds. Secondly, he treats jump occurrences as idiosyncratic risk that is not priced by the CAPM model. With these assumptions in place, Merton is capable of allowing the option price to trend at the riskless rate of return. This in turn, yields a general solution for the option contract. If we presume, jumps sizes are log-normally distributed, we obtain the closed form expression given by equation (276).

Merton's analysis leads to a unique and insightful result. By incorporating discrete random jumps in a price process, he gives greater flexibility and reality to the model. We have already seen where the Black-Scholes analysis is capable of pricing options on futures contracts, where the futures price is found to follow a geometric Brownian motion. Casual observations of commodity spot and futures prices, however, indicate this is not a realistic assumption. These prices often exhibit price spikes and would therefore not follow a geometric Brownian motion all the time. In light of this evidence, we need a model that does incorporate jumps in the underlying commodity price. Merton's model is a possible candidate for our analysis. The only problem with Merton's analysis is his treatment of jumps. That is, he presumes they are idiosyncratic and therefore diversifiable. It is unrealistic to believe that a commodity's price spikes are anything but systematic and therefore does not fit the assumptions of Merton's model. As a result, we need an option pricing model which allows for systematic jumps in commodity prices.

4.6 Bates' Model

In Merton's work, he assumes the jump risk in a stock's price is diversifiable, which permits returns on the otherwise risk-free replicating portfolio to be equal to the risk-free rate. The assumption is important in that it allows Merton a means of obtaining a fundamental partial differential equation for pricing options written on the underlying stock. The question remains whether Merton's assumption of diversifiable jump risk is reasonable. Recently, some literature has developed to consider this phenomenon. In particular, Bates (1988, 1991) notes that the stock market crash in October of 1987 is evidence that asset prices are prone to exhibit significant random discrete jumps. Furthermore, these jumps manifest themselves in the S&P 500 index. It is hard to accept that jumps in the S&P 500 index can be seen as idiosyncratic. Therefore, in consideration of asset price behavior, Bates constructs a model for pricing contingent claims, where asset prices incorporate systematic jumps in the diffusion processes.

Bates notes that introducing systematic jumps to asset prices introduces new forms of risk, namely jump risk. The challenge to modeling this behavior when pricing options is that the jump risk is embodied in option prices, but this risk is not directly priced by any instrument currently traded in financial markets. In the Black-Scholes analysis, the only sort of risk that is considered is the regular price vibrations of the underlying security. Since the stock price is a tradable asset, the underlying risk premium is said to be

$$\lambda = \frac{\mu - r}{\sigma},$$

where μ is the instantaneous expected return, r is the riskless rate of return, and σ is the diffusion coefficient. Under the proposed model of Bates, to price options written on stock prices one will need a risk premium for normal price vibrations and a risk premium for large discrete price vibrations as well. The result is that the Black-Scholes arbitrage-based methodology cannot be used. To model this phenomenon, one must impose restrictions on preferences and technologies in order to price those forms of risk, and consequently to price options. Bates extends the Cox, Ingersoll, Ross (1985a) general equilibrium framework, by including a jump process, to examine how options can be priced under jump risk.

The general framework for Bates' model is as follows. There are a large number of infinitely lived consumers, with identical preferences, endowments, and information sets. Each consumer seeks to maximize a lifetime expected utility function of the form

$$E_0 \int_0^{\infty} e^{-\rho t} U(C_t, Y_t) dt \tag{277}$$

subject to initial wealth W_0 and initial underlying state of the economy Y_0. $U(C_t, \cdot)$ is assumed strictly concave in the consumption flow C.

There are N investment opportunities in real assets available to every investor. The model is a pure capital growth model where it is assumed that labor is unnecessary in production. The return on each investment follows a state-dependent jump-diffusion process

$$\frac{dP_i}{P_i} = [\alpha_i(Y) - \lambda E_Y(k_i)]dt + g_i(Y)dZ + k_i(Y)dq, \tag{278}$$

where $\alpha_i(Y)$ is the state-dependent instantaneous expected return on the process. Z is an (N+K)x1 vector of independent standard Weiner processes. N is the number of investment opportunities and K is the number of state variables, in the Kx1 vector Y. $g_i(Y)$ is a 1x(N+K) state-dependent vector reflecting the sensitivity of returns to the various shocks. q is a Poisson counter with intensity λ, and $k_i(Y)$ is the random percentage jump size conditional on the Poisson-distributed event occurring. The state dependent distribution of $k_i(Y)$ has a mean of $E_Y(k_i)$ conditional on state Y. The movement of the K-dimensional vector of state variables Y is determined by a system of stochastic differential equations of the form

$$dY = [\mu(Y) - \lambda E_Y(\Delta Y)]dt + \sigma_Y(Y)dZ + \Delta Y dq \qquad (279)$$

where $\mu(Y)$ is the state-dependent drift in Y, $\sigma_Y(Y)$ is a Kx(N+K) matrix giving the state-dependent sensitivities of the underlying state variables, and ΔY is a Kx1 vector of random increments to the state variables conditional on the Poisson-distributed event occurring. The Poisson shock is assumed to affect all investment opportunities and underlying state variables simultaneously, and is therefore systematic and non-diversifiable risk.

At each instant, each consumer chooses a consumption flow and an investment strategy to maximize expected utility over the consumer's remaining lifetime given the state of the economy. Solving this maximization problem yields a general capital asset pricing model for jump diffusions. Excess returns on any investment are generated by the security's content of various forms of systematic risk. These are 1) market risk conditional on no jumps, 2) technological risks (shifts in the investment opportunity set) conditional on no jumps, and 3) jump risk, which includes both market and technological risks. In theorem three of his analysis, Bates shows the asset pricing model implies that options on non-dividend paying stocks are priced as if investors are risk-neutral. The system of stochastic differential equations for the asset price, wealth, and the state variables are

$$\frac{dS}{S} = [r - \lambda^* E_{W,Y}(k_s^*)]dt + \sigma_s dZ^* + k_s^* dq^* \qquad (280)$$

$$\frac{dW}{W} = \left[r - \lambda^* E_{W,Y}\left(k_W^*\right) - \frac{C(W,Y)}{W} \right] dt + \sigma_W dZ^* + k_W^* dq^* \tag{281}$$

$$dY = \left[\mu(Y) - \lambda E_Y(\Delta Y) - \Phi_Y(W,Y)\right] dt + \sigma_Y(Y) dZ^* + \Delta Y^* dq^* \tag{282}$$

where

$$\Phi_Y(W,Y) = -Cov\left(dJ_W / J_W, dY'\right) \tag{283}$$

and jumps occur with frequency $\lambda^*(W,Y) = \frac{\lambda E_{W,Y}\left(J_W(W + k_W W, Y + \Delta Y)\right)}{J_W}$. dS and dW are Nx1 vectors and dY is a Kx1 vector. E_Y denotes the expectation operator conditional on state Y relative to the true joint probability density function of $\left[k_s, k_W, \Delta Y\right]$ denoted $f(k_s, k_W, \Delta Y \mid Y)$. $E_{W,Y}^*$ denotes the expectations operator conditional on state$[W,Y]$ relative to the marginal weighted joint probability distribution function of $\left[k_s, k_W, \Delta Y\right]$

$$f(k_s^*, k_W^*, \Delta Y^* \mid Y) = \frac{J_W(W + k_W W, Y + \Delta Y)}{E_{W,Y}\left(J_W(W + k_W W, Y + \Delta Y)\right)} f(k_s, k_W, \Delta Y \mid Y). \tag{284}$$

Note that the beginning analysis defines P as the price of any asset available for investment. The analysis now uses S for the asset price, because we are interested in pricing options written on a stock price.

Bates' asset pricing model is capable of pricing options under varying assumptions. The Black-Scholes solution is attainable by imposing three restrictions on the model. In particular if there are no jumps, the underlying asset's price volatility is not a function of other underlying state variables Y, and the risk-free rate is nonstochastic. The Black-Scholes model reflects the fact that the assumptions ensure that the option contains only one form of systematic risk, namely the underlying asset. Bates, however, has a model that involves forms of risk that are not directly price by the market; jump risk. Consequently, Bates shows it is necessary to derive prices for these forms of risk from restrictions of preferences and on distributions. For jump risk models, the price of risk is reflected in the modified jump frequency parameter $\lambda^*(W,Y)$ and the modified jump size distribution $f(k_s^*, k_W^*, \Delta Y^* \mid Y)$. To find an option price with systematic jump risk we need to make the following assumptions: 1) the representative agent has a log utility, 2) the underlying asset price follows a jump-diffusion of the type state above, with constant volatility and random state-independent percentage jump amplitude k_s, 3) wealth follows a similar jump-diffusion process, with state-independent random percentage jump amplitude k_W but with possible state independent volatility, 4) one plus the percentage jump amplitudes in the underlying asset price and wealth have a joint log-normal distribution, and 5) the instantaneous risk-free rate is nonstochastic. With the above assumptions, the option prices depend only on the underlying asset price and time, and are evaluated using a risk-neutral jump-diffusion with log-normal random jumps

$$\frac{dS}{S} = \left[r - \lambda^* E^*\left(k_s^*\right)\right]dt + \sigma_s dZ^* + k_s^* dq^*,$$

where the terms are defined above and

$$\ln(1 + k_s^*) \sim N\left[\gamma_s - R\delta_{sW} - \frac{1}{2}\delta_s^2, \delta_s^2\right] \equiv N\left[\gamma_s^* - \frac{1}{2}\delta_s^2, \delta_s^2\right],$$

where γ^* is the risk adjusted mean jump size. Bates (1991) notes for options on futures contracts the jump diffusion process becomes

$$\frac{dF}{F} = -\lambda^* E^*\left(k_F^*\right)dt + \sigma dZ^* + k_F^* dq^*, \qquad (285)$$

where σ is the instantaneous variance conditional on no jumps; dZ^* is the increment of a standard Brownian motion under a risk-adjusted probability measure; k_F^* is the risk-adjusted random percentage jump conditional on a Poisson distributed event occurring; $(1+k_F^*)$ is log-normally distributed: $\ln(1+k_F^*) \sim N\left(\gamma^* - \frac{1}{2}\omega^2, \omega^2\right)$; λ^* is the modified frequency of Poisson events; and q^* is a Poisson counter with intensity λ^*: $\Pr(dq^*(t)=1)=\lambda^* dt$, $\Pr(dq^*(t)=0)=1-\lambda^* dt$. Define $\bar{k}_F^* \equiv E^*(k_F^*)$. From Press (1967) and Merton (1976) (expression (276)) the option price written on a futures contract that follows a jump diffusion is

$$C(FY,t) = e^{-r\tau} \sum_{n=0}^{\infty} \frac{e^{-\lambda^*\tau}(\lambda^*\tau)^n}{n!} \left[F(t)e^{b(n)\tau} N(d_1) - XN(d_2)\right], \tag{286}$$

where $b(n) = \left(-\lambda^* \bar{k}_F^* + \frac{n\gamma^*}{\tau}\right)$,

$$d_1 = \frac{\ln\left(\frac{F(t)}{X}\right) + \left(\left(-\lambda^* \bar{k}_F^* + \frac{n\gamma^*}{\tau}\right)\tau + \frac{1}{2}(\sigma^2\tau + n\omega^2)\right)}{\sqrt{\sigma^2\tau + n\omega^2}}, \tag{287}$$

and

$$d_2 = d_1 - \sqrt{\sigma^2\tau + n\omega^2}\,. \tag{288}$$

This result is shown below for the process in expression (285).

The process in equation (285) resembles a geometric Brownian motion most of the time, but on average λ^* times per year the price jumps discretely by a random amount. Working on the futures price, let $H(t) = \ln F(t)$. The transformation, and Ito's lemma yields the following process for $H(t)$

$$dH = H_F dF_{dq=0} + \frac{1}{2} H_{FF} \left[dF_{dq=0}\right]^2 + H(FY,t) - H(F,t),$$

$$dH = \frac{1}{F}(F)\left(-\lambda^* \bar{k}_F^* dt + \sigma dZ^*(t)\right) - \frac{1}{2}\frac{1}{F^2}\left(F^2\sigma^2 dt\right) + \ln FY - \ln F,$$

$$dH = \left(-\lambda^* \bar{k}_F^* dt + \sigma dZ^*(t)\right) - \frac{1}{2}\left(\sigma^2 dt\right) + \ln\left(\frac{FY}{F}\right),$$

$$dH = \left(-\lambda^* \bar{k}_F^* - \frac{1}{2}\sigma^2\right)dt + \sigma dZ^*(t) + \ln(Y)$$

$$\int_t^T dH = \int_t^T \left(-\lambda^* \bar{k}_F^* - \frac{1}{2}\sigma^2\right)ds + \sigma \int_t^T dZ^*(s) + \sum_{i=1}^n \ln(Y_i),$$

$$H(T) - H(t) = \int_t^T \left(-\lambda^* \bar{k}_F^* - \frac{1}{2}\sigma^2\right)ds + \sigma \int_t^T dZ^*(s) + \sum_{i=1}^n \ln(Y_i),$$

$$H(T) - H(t) = \left(-\lambda^* \bar{k}_F^* - \frac{1}{2}\sigma^2\right)(T-t) + \sigma \int_t^T dZ^*(s) + \sum_{i=1}^n \ln(Y_i),$$

$$H(T) = H(t) + \left(-\lambda^* \bar{k}_F^* - \frac{1}{2}\sigma^2\right)\tau + \sigma \int_t^T dZ^*(s) + \sum_{i=1}^n \ln(Y_i),$$

$$F(T) = F(t)\exp[B(T)]Y(n), \tag{289}$$

where $B(T) = \left(-\lambda^* \bar{k}_F^* - \frac{1}{2}\sigma^2\right)\tau + \sigma \int_t^T dZ^*(s)$, and $Y(n) = \prod_{i=1}^n Y_i$. Expression (289) is an alternative representation for the futures price process. Here the futures price is given for a particular time conditional on the instantaneous trend, movements in the Weiner process and n jumps. $F(T)$ is a random variable. The uncertainty underlying $F(T)$ is introduced by the movements in the Weiner process and the size of the jumps. While $F(T)$ may be random, we do, however, have some expectation about the futures price at time T. This expectation is given as[14]

$$E_t^*[F(T)] = \sum_{n=0}^{\infty} \frac{e^{-\lambda^*\tau}(\lambda^*\tau)^n}{n!}\left[E_t^*(F(T) \mid n\ jumps)\right], \quad (290)$$

where $E(F(T) \mid n\ jumps)$ is the conditional expectation of the continuous movements of $F(T)$ over the interval $(T\text{-}t)$. Expression (290) is simply the Poisson weighted sum of n expected values of $F(T)$. The form of $E(F(T) \mid n\ jumps)$ will depend on the distribution of the continuous components of $F(T)$.

To determine the distribution of the futures price at time T, we look at expression (289). The two sources of randomness are the Brownian motion and the jump size. The Brownian motion is by definition normally distributed, and the distribution for the jump size is presumed to be log-normal. The continuous futures price vibrations are log-normal. Therefore, the conditional expected value on the right-hand side of expression (290) is found as

$$E_t^*[F(T) \mid n\ jumps] = \exp\left[E_t^*[H(T) \mid n\ jumps] + \frac{1}{2}V_t^*[H(T) \mid n\ jumps]\right]. \quad (291)$$

The conditional mean of $H(T)$ is

$$\begin{aligned} E_t^*[H(T) \mid n\ jumps] &= H(t) + \left(-\lambda^*\bar{k}_F^* - \frac{1}{2}\sigma^2\right)\tau + \sum_{i=1}^{n} E_n^*[\ln(Y_i)], \\ &= H(t) + \left(-\lambda^*\bar{k}_F^* - \frac{1}{2}\sigma^2\right)\tau + \sum_{i=1}^{n}\left(\gamma^* - \frac{1}{2}\omega^2\right), \\ &= H(t) + \left(-\lambda^*\bar{k}_F^* - \frac{1}{2}\sigma^2\right)\tau + n\left(\gamma^* - \frac{1}{2}\omega^2\right). \quad (292) \end{aligned}$$

The conditional variance of $H(T)$ is

$$V_t^*[H(T) \mid n\ jumps] = V_t^*\left[\sigma\int_t^T dZ^*(s)\right] + \sum_{i=1}^{n} V_n^*[\ln(Y_i)].$$

Recall the jump term and the Wiener process are uncorrelated.

$$V_t[H(T) \mid n\ jumps] = \sigma^2\tau + \sum_{i=1}^{n}\omega^2\ ,$$

$$= \sigma^2\tau + n\omega^2\ . \tag{293}$$

Note that the expressions for the conditional mean and variance of *H(T)* are identical in form for Bates and Merton models. The two models, however, are fundamentally different. One, Merton presumes idiosyncratic jumps, while Bates does not. Secondly, the parameters for the jump process in the Bates model are risk-adjusted, whereas, Merton uses the actual parameter values. Continuing, we substitute expressions (292) and (293) into (291) to obtain

$$E_t^*[F(T) \mid n\ jumps] = \exp\left[H(t) + \left(-\lambda^*\bar{k}_F^* - \frac{1}{2}\sigma^2\right)\tau + n\left(\gamma^* - \frac{1}{2}\omega^2\right) + \frac{1}{2}\left(\sigma^2\tau + n\omega^2\right)\right],$$

$$= \exp\left[H(t) - \lambda^*\bar{k}_F^*\tau - \frac{1}{2}\sigma^2\tau + n\gamma^* - \frac{1}{2}n\omega^2 + \frac{1}{2}\sigma^2\tau + \frac{1}{2}n\omega^2\right],$$

$$= \exp\left[H(t) - \lambda^*\bar{k}_F^*\tau + n\gamma^*\right],$$

$$= \exp\left[H(t) + \left(-\lambda^*\bar{k}_F^* + \frac{n\gamma^*}{\tau}\right)\tau\right],$$

$$= F(t)\exp\left[\left(-\lambda^*\bar{k}_F^* + \frac{n\gamma^*}{\tau}\right)\tau\right]. \tag{294}$$

Expression (290) is the expected futures price at time *T* for the continuous movements conditioned on *n* jumps occurring over the interval *(T-t)*. Rewriting expression (294) in a more simplified manner yields

$$E_t^*[F(T) \mid n\ jumps] = F(t)e^{b(n)\tau} \tag{295}$$

where $b(n) = -\lambda^* \bar{k}^* + \frac{n\gamma^*}{\tau}$. Taking the result from equation (295) and substituting it into expression (290) yields

$$E_t^*[F(T)] = \sum_{n=0}^{\infty} \frac{e^{-\lambda^* \tau} (\lambda^* \tau)^n}{n!} \left[F(t) e^{b(n)\tau}\right], \tag{296}$$

which is the risk-adjusted expected futures price at time T contingent on n jumps occurring over the investment horizon. With the characteristics of the futures price in place, we are ready to begin pricing call options written on the futures contract.

Expression (285) is the return dynamics for futures prices. This expression is very similar to Merton's price dynamics. As noted above, there are subtle differences between the two expressions. One, the instantaneous drift in Bates' expression is zero and two, the Bates' formula is a risk-neutral process allowing for systematic price and jump risk. The form of Bates' option pricing formula in expression (286) is identical to the form of Merton's (1976) formula (expression (276)) with the difference being the treatment of the jump parameters. Merton's formula incorporates the actual parameters for the idiosyncratic jump distribution, whereas, Bates has the modified systematic jump parameters. To find a closed form solution to Bates, we need the Black-Scholes solution conditioned on n jumps occurring over the investment horizon.

In a risk neutral world, the price of a European call option written on the risk-adjusted futures contract can be expressed as

$$C(F,t \mid n) = e^{-r(T-t)} E_t^*\left[C(F,T \mid n\ jumps)\right], \tag{297}$$

where $C(F,t \mid n)$ is the price of the call option today, and $C(F,T \mid n)$ is the risk adjusted call option at a terminal date T. Substituting in for the risk-adjusted terminal call price we get

$$C(F,t \mid n) = e^{-r\tau} E_t^*\left[\max\left(F(T) - X, 0\right) \mid n\ jumps\right], \tag{298}$$

From expression (289), we have the futures price at time T written as

$$F(T) = F(t)\exp\left[B(T)\right]Y(n),$$

where $B(T) = \left(-\lambda^* \bar{k}_F^* - \frac{1}{2}\sigma^2\right)\tau + \sigma\int_t^T dZ^*(s)$, and $Y(n) = \prod_{i=1}^{n} Y_i$. We now define $G(T)$ as $G(T) = \left(-\lambda^* \bar{k}_F^* - \frac{1}{2}\sigma^2\right)\tau + \sigma\int_t^T dZ^*(s) + \sum_{i=1}^{n}\ln(Y_i)$, where $\ln(Y_i)$ is normally distributed with mean $\gamma^* - \frac{1}{2}\omega^2$ and variance of ω^2. That is, we move the jump component back into the exponential of the futures price in expression (289). The max function in expression (298) can be rewritten as

$$C(F,t \mid n) = e^{-r\tau}\left\{E_t^*\left[F(t)e^{G(T)} - X \mid F(T) \geq X, n\ jumps\right] + E_t^*\left[0 \mid F(T) < X, n\ jumps\right]\right\},$$

$$= e^{-r\tau}\left\{E_t^*\left[F(t)e^{G(T)} - X \mid F(t)e^{G(T)} \geq X, n\ jumps\right] + E_t^*\left[0 \mid F(t)e^{G(T)} < X, n\ jumps\right]\right\},$$

$$= e^{-r\tau}\left\{E_t^*\left[F(t)e^{G(T)} - X \mid e^{G(T)} \geq \frac{X}{F(t)}, n\ jumps\right] + E_t^*\left[0 \mid e^{G(T)} < \frac{X}{F(t)}, n\ jumps\right]\right\},$$

$$= e^{-r\tau}\left\{E_t^*\left[F(t)e^{G(T)} - X \mid G(T) \geq \ln\left(\frac{X}{F(t)}\right), n\ jumps\right]\right.$$

$$\left. + E_t^*\left[0 \mid G(T) < \ln\left(\frac{X}{F(t)}\right), n\ jumps\right]\right\} \tag{299}$$

From expression (299), we see that the price of a call today is simply the linear combination of two partial expectations conditioned on n jumps occurring. The first partial expectation considers the option when it is in the money and the second is for the option when it is out of the money. Conditioning on n jumps occurring, we may formally write the probability of the option expiring in the money. This is

$$P^{*}\{X \leq F(T) \leq \infty\},$$

where P^{*} is an equivalent martingale probability. Transforming the variable then constitutes the following

$$P^{*}\left\{\frac{X}{F(t)} \leq \frac{F(T)}{F(t)} \leq \infty\right\},$$

$$P^{*}\left\{\ln\left(\frac{X}{F(t)}\right) \leq \ln\left(\frac{F(T)}{F(t)}\right) \leq \infty\right\}.$$

Since $\ln\left(\frac{F(T)}{F(t)}\right) = G(T)$, the above becomes

$$P^{*}\left\{\ln\left(\frac{X}{F(t)}\right) \leq G(T) \leq \infty\right\}.$$

The second partial expectation operator in expression (299) considers the probability that the spot price is less than the exercise price at maturity. Formally, this is

$$P^{*}\{-\infty \leq F(T) \leq X\},$$

where P^{*} is an equivalent martingale probability. Again, transforming the variable then constitutes the following

$$P^{*}\left\{-\infty \leq \frac{F(T)}{F(t)} \leq \frac{X}{F(t)}\right\},$$

$$P^{*}\left\{-\infty \leq \ln\left(\frac{F(T)}{F(t)}\right) \leq \ln\left(\frac{X}{F(t)}\right)\right\},$$

$$P^{*}\left\{-\infty \leq G(T) \leq \ln\left(\frac{X}{F(t)}\right)\right\}.$$

Given the probability statements above, the linear combination of the partial expectations in expression (299) may be written as

$$C(F,t\mid n)=e^{-r\tau}\left[\int_{-\infty}^{\ln\left(\frac{X}{F(t)}\right)}(0)dP^{*}+\int_{\ln\left(\frac{X}{F(t)}\right)}^{\infty}\left(F(t)e^{G(T)}-X\right)dP^{*}\right], \tag{300}$$

where

$$dP^{*}=\frac{1}{\sqrt{2\pi\left(\sigma^{2}\tau+n\omega^{2}\right)}}\exp\left[-\frac{\left(g(T)-\left(b(n)\tau-\frac{1}{2}(\sigma^{2}\tau+n\omega^{2})\right)\right)^{2}}{2\left(\sigma^{2}\tau+n\omega^{2}\right)}\right]dg(T). \tag{301}$$

Note that Bates' model is a risk neutral pricing model. In light of this observation the expectation operator in expression (300), is taken with respect to the equivalent probability measure, dP^{*}, and not the actual probability distribution. This notion is discussed in section 4.1 expression (31). Combining (300) and (301) we have

$$C(F,t\mid n)=e^{-r\tau}\left[\int_{-\infty}^{\ln\left(\frac{X}{F(t)}\right)}\frac{0}{\sqrt{2\pi\left(\sigma^{2}\tau+n\omega^{2}\right)}}\exp\left[-\frac{\left(g(T)-\left(b(n)\tau-\frac{1}{2}\left(\sigma^{2}\tau+n\omega^{2}\right)\right)\right)^{2}}{2\left(\sigma^{2}\tau+n\omega^{2}\right)}\right]dg(T)\right.$$

$$\left.+\int_{\ln\left(\frac{X}{F(t)}\right)}^{\infty}\frac{\left(F(t)e^{g(T)}-X\right)}{\sqrt{2\pi\left(\sigma^{2}\tau+n\omega^{2}\right)}}\exp\left[-\frac{\left(g(T)-\left(b(n)\tau-\frac{1}{2}\left(\sigma^{2}\tau+n\omega^{2}\right)\right)\right)^{2}}{2\left(\sigma^{2}\tau+n\omega^{2}\right)}\right]dg(T)\right] \tag{302}$$

The first term in expression (302) is equal to zero leaving us with

$$C(F,t\,|\,n) = e^{-r\tau} \int_{\ln\left(\frac{X}{F(t)}\right)}^{\infty} \frac{\left(F(t)e^{g(T)} - X\right)}{\sqrt{2\pi\left(\sigma^2\tau + n\omega^2\right)}} \exp\left[-\frac{\left(g(T) - \left(b(n)\tau - \frac{1}{2}\left(\sigma^2\tau + n\omega^2\right)\right)\right)^2}{2\left(\sigma^2\tau + n\omega^2\right)} \right] dg(T)$$

(303)

Expression (303) can be split up into two different integrals. For ease of exposition, the brackets of the exponential in expression (303) are acknowledged with a dot.

$$\int_{\ln\left(\frac{X}{F(t)}\right)}^{\infty} e^{-r\tau} \frac{F(t)e^{g(T)}}{\sqrt{2\pi\left(\sigma^2\tau + n\omega^2\right)}} \exp[\cdot]dg(T) - \int_{\ln\left(\frac{X}{F(t)}\right)}^{\infty} e^{-r\tau} \frac{X}{\sqrt{2\pi\left(\sigma^2\tau + n\omega^2\right)}} \exp[\cdot]dg(T)$$

(304)

We start with the second integral in expression (304). Consider the transformation of *G(T)*

$$Z = \frac{G(T) - \left(b(n)\tau - \frac{1}{2}\left(\sigma^2\tau + n\omega^2\right)\right)}{\sqrt{\left(\sigma^2\tau + n\omega^2\right)}}. \qquad (305)$$

If we wish to write the integrals in (304) as a function of *z*, we must adjust the limits. Currently, we are considering the probability that the spot price is greater than the exercise price at maturity. Formally, this is

$$P^*\left\{\ln\left(\frac{X}{F(t)}\right) \le G(T) \le \infty\right\}. \qquad (306)$$

Recall *G(T)* has a mean of $b(n)\tau - \frac{1}{2}\left(\sigma^2\tau + n\omega^2\right)$ and a variance of $\left(\sigma^2\tau + n\omega^2\right)$. If we subtract the mean of *G(T)* and then divide by the standard deviation, this yields

$$P^{*}\left\{\frac{\ln\left(\frac{X}{F(t)}\right)-\left(b(n)\tau-\frac{1}{2}\left(\sigma^{2}\tau+n\omega^{2}\right)\right)}{\sqrt{\left(\sigma^{2}\tau+n\omega^{2}\right)}}\leq\frac{G(T)-\left(b(n)\tau-\frac{1}{2}\left(\sigma^{2}\tau+n\omega^{2}\right)\right)}{\sqrt{\left(\sigma^{2}\tau+n\omega^{2}\right)}}\leq\infty\right\}.$$

From equation (305) we have

$$P^{*}\left\{\frac{\ln\left(\frac{X}{F(t)}\right)-\left(b(n)\tau-\frac{1}{2}\left(\sigma^{2}\tau+n\omega^{2}\right)\right)}{\sqrt{\left(\sigma^{2}\tau+n\omega^{2}\right)}}\leq Z\leq\infty\right\}. \tag{307}$$

Thus, if we change the variable of integration in the above to Z, then the limits of integration for this problem become $\frac{\ln\left(\frac{X}{F(t)}\right)-\left(b(n)\tau-\frac{1}{2}\left(\sigma^{2}\tau+n\omega^{2}\right)\right)}{\sqrt{\left(\sigma^{2}\tau+n\omega^{2}\right)}}$ and ∞, making the integral

$$-e^{-r\tau}X\int_{\frac{\ln\left(\frac{X}{F(t)}\right)-\left(b(n)\tau-\frac{1}{2}\left(\sigma^{2}\tau+n\omega^{2}\right)\right)}{\sqrt{\left(\sigma^{2}\tau+n\omega^{2}\right)}}}^{\infty}\frac{1}{\sqrt{2\pi}}\exp\left[-\frac{1}{2}z^{2}\right]dz. \tag{308}$$

Recall the symmetry property of the normal distribution

$$\int_{L}^{\infty}f(z)dz=\int_{-\infty}^{-L}f(z)dz. \tag{309}$$

Using the property above, we may rewrite expression (308) as

$$-e^{-r\tau}X\int_{-\infty}^{\frac{-\left(\ln\left(\frac{X}{F(t)}\right)-\left(b(n)\tau-\frac{1}{2}\left(\sigma^{2}\tau+n\omega^{2}\right)\right)\right)}{\sqrt{\left(\sigma^{2}\tau+n\omega^{2}\right)}}}\frac{1}{\sqrt{2\pi}}\exp\left[-\frac{1}{2}z^{2}\right]dz. \tag{310}$$

The upper limit in (310) can be rewritten. This is,

$$-\frac{\ln\left(\frac{X}{F(t)}\right)-\left(b(n)\tau-\frac{1}{2}\left(\sigma^2\tau+n\omega^2\right)\right)}{\sqrt{\left(\sigma^2\tau+n\omega^2\right)}}$$

$$\frac{\ln\left(\frac{F(t)}{X}\right)+\left(b(n)\tau-\frac{1}{2}\left(\sigma^2\tau+n\omega^2\right)\right)}{\sqrt{\left(\sigma^2\tau+n\omega^2\right)}}.$$

We define the upper limit as

$$d_2=\frac{\ln\left(\frac{F(t)}{X}\right)+\left(b(n)\tau-\frac{1}{2}\left(\sigma^2\tau+n\omega^2\right)\right)}{\sqrt{\left(\sigma^2\tau+n\omega^2\right)}} \tag{311}$$

Using expression (311) in (310) yields

$$-e^{-r\tau}X\int_{-\infty}^{d_2}\frac{1}{\sqrt{2\pi}}\exp\left[-\frac{1}{2}z^2\right]dz,$$

which equals

$$-e^{-r\tau}XN(d_2), \tag{312}$$

Now working on the first integral in expression (304), we start with

$$\int_{\ln\left(\frac{X}{F(t)}\right)}^{\infty} e^{-r\tau}F(t)e^{g(T)}\frac{1}{\sqrt{2\pi\left(\sigma^2\tau+n\omega^2\right)}}\exp[\cdot]dg(T).$$

Changing the variable of integration for the integral above to *dz* we must change the limits. This is

$$\int_{\frac{\ln\left(\frac{X}{F(t)}\right)-\left(b(n)\tau-\frac{1}{2}\left(\sigma^2\tau+n\omega^2\right)\right)}{\sqrt{\left(\sigma^2\tau+n\omega^2\right)}}}^{\infty} e^{-r\tau}F(t)e^{g(T)}\frac{1}{\sqrt{2\pi}}\exp\left[-\frac{1}{2}z^2\right]dz$$

The limits of integration are for the variable *Z*, but we still have *g(T)* inside the integral. We need to transform this variable. First, the lower limit can be rewritten. This is

$$\int_{-d_2}^{\infty} e^{-r\tau} F(t) e^{g(T)} \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{1}{2}z^2\right] dz.$$

Working on the variable $g(T)$ we multiply the above expression by 1. This yields

$$\exp\left(b(n)\tau - \frac{1}{2}\left(\sigma^2\tau + n\omega^2\right)\right)\exp\left(-\left[b(n)\tau - \frac{1}{2}\left(\sigma^2\tau + n\omega^2\right)\right]\right). \tag{313}$$

Further simplification yields

$$e^{-r\tau} e^{\left(b(n)\tau - \frac{1}{2}\left(\sigma^2\tau + n\omega^2\right)\right)} F(t) \int_{-d_2}^{\infty} e^{g(T)} e^{\left(-\left(b(n)\tau - \frac{1}{2}\left(\sigma^2\tau + n\omega^2\right)\right)\right)} \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{1}{2}z^2\right] dz.$$

This is equal to

$$e^{-r\tau} e^{\left(b(n)\tau - \frac{1}{2}\left(\sigma^2\tau + n\omega^2\right)\right)} F(t) \int_{-d_2}^{\infty} e^{\left(g(T) - \left(b(n)\tau - \frac{1}{2}\left(\sigma^2\tau + n\omega^2\right)\right)\right)} \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{1}{2}z^2\right] dz. \tag{314}$$

We know

$$Z = \frac{G(T) - \left(b(n)\tau - \frac{1}{2}\left(\sigma^2\tau + n\omega^2\right)\right)}{\sqrt{\sigma^2\tau + n\omega^2}},$$

which is equal to

$$Z\left(\sqrt{\sigma^2\tau + n\omega^2}\right) = G(T) - \left(b(n)\tau - \frac{1}{2}\left(\sigma^2\tau + n\omega^2\right)\right). \tag{315}$$

Substituting (315) into (314) yields

$$e^{-r\tau} e^{\left(b(n)\tau - \frac{1}{2}\left(\sigma^2\tau + n\omega^2\right)\right)} F(t) \int_{-d_2}^{\infty} e^{\left(\sqrt{\sigma^2\tau + n\omega^2}\right)z} \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{1}{2}z^2\right] dz,$$

$$e^{-r\tau} e^{\left(b(n)\tau - \frac{1}{2}\left(\sigma^2\tau + n\omega^2\right)\right)} F(t) \int_{-d_2}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(z^2 - 2z\sqrt{\sigma^2\tau + n\omega^2}\right)\right] dz. \tag{316}$$

The exponential term inside the integral in expression (316) can be completed into a square. To do so we need to multiply (316) by

$$\exp\left[\frac{\sigma^2\tau + n\omega^2}{2}\right]\exp\left[-\frac{\sigma^2\tau + n\omega^2}{2}\right]. \tag{317}$$

This yields

$$e^{-r\tau}F(t)e^{\frac{\sigma^2\tau+n\omega^2}{2}+\left(b(n)\tau-\frac{1}{2}(\sigma^2\tau+n\omega^2)\right)}\int_{-\infty}^{d_2}\frac{1}{\sqrt{2\pi}}\exp\left[-\frac{1}{2}\left(z^2 - 2z\sqrt{\sigma^2\tau + n\omega^2} - \left(\sigma^2\tau + n\omega^2\right)\right)\right]dz$$

or

$$e^{-r\tau}F(t)e^{b(n)\tau}\int_{-\infty}^{d_2}\frac{1}{\sqrt{2\pi}}\exp\left[-\frac{1}{2}\left(z - \sqrt{\sigma^2\tau + n\omega^2}\right)^2\right]dz. \tag{318}$$

The integral in expression (318) is no longer for a standard normal variable. Its mean has been moved to the left by $\sqrt{\sigma^2\tau + n\omega^2}$ units. We can restate the integral in terms of a standard normal variable. If we let $w = z - \sqrt{\sigma^2\tau + n\omega^2}$ then we see that $\frac{dw}{dz} = 1$. When $z = \infty$ then $w = \infty$, and when

$$z = \frac{\ln\left(\frac{X}{F(t)}\right) - \left(b(n)\tau - \frac{1}{2}\left(\sigma^2\tau + n\omega^2\right)\right)}{\sqrt{\sigma^2\tau + n\omega^2}}$$

then

$$w = \frac{\ln\left(\frac{X}{F(t)}\right) - \left(b(n)\tau - \frac{1}{2}\left(\sigma^2\tau + n\omega^2\right)\right)}{\sqrt{\sigma^2\tau + n\omega^2}} - \sqrt{\left(\sigma^2\tau + n\omega^2\right)}.$$

The first term in the expression above is $-d_2$, thus we have

$$w = -d_2 - \sqrt{\left(\sigma^2\tau + n\omega^2\right)}.$$

Rewriting expression (318) in terms of W reduces to

$$e^{-r\tau}F(t)e^{b(n)\tau}\int_{-d_2-\sqrt{\sigma^2\tau+n\omega^2}}^{\infty}\frac{1}{\sqrt{2\pi}}\exp\left[-\frac{1}{2}w^2\right]dw. \tag{319}$$

Using the symmetry property of the normal distribution expression (319) can be rewritten as

$$e^{-r\tau}F(t)e^{b(n)\tau}\int_{-\infty}^{d_2+\sqrt{\sigma^2\tau+n\omega^2}}\frac{1}{\sqrt{2\pi}}\exp\left[-\frac{1}{2}w^2\right]dw,$$

$$e^{-r\tau}F(t)e^{b(n)\tau}\int_{-\infty}^{d_1}\frac{1}{\sqrt{2\pi}}\exp\left[-\frac{1}{2}w^2\right]dw. \tag{320}$$

The upper limit is equal to

$$d_1 = d_2 + \sqrt{\sigma^2\tau + n\omega^2},$$

$$= \frac{\ln\left(\frac{F(t)}{X}\right) + \left(b(n)\tau - \frac{1}{2}(\sigma^2\tau + n\omega^2)\right)}{\sqrt{\sigma^2\tau + n\omega^2}} + \sqrt{\sigma^2\tau + n\omega^2},$$

$$= \frac{\ln\left(\frac{F(t)}{X}\right) + \left(b(n)\tau - \frac{1}{2}(\sigma^2\tau + n\omega^2)\right)}{\sqrt{\sigma^2\tau + n\omega^2}} + \frac{\sigma^2\tau + n\omega^2}{\sqrt{\sigma^2\tau + n\omega^2}},$$

$$= \frac{\ln\left(\frac{F(t)}{X}\right) + \left(b(n)\tau + \frac{1}{2}(\sigma^2\tau + n\omega^2)\right)}{\sqrt{\sigma^2\tau + n\omega^2}}. \tag{321}$$

Given the above, expression (320) reduces to

$$e^{-r\tau}F(t)e^{b(n)\tau}N(d_1), \tag{322}$$

Now replacing the integrals in expression (304) with expressions (312) and (322) yields

$$C(F,t\mid n) = e^{-r\tau}F(t)e^{b(n)\tau}N(d_1) - e^{-r\tau}XN(d_2), \tag{323}$$

where $b(n) = \left(-\lambda^{*}\bar{k}^{*} + \frac{n\gamma^{*}}{\tau}\right)$, (324)

$$d_1 = \frac{\ln\left(\frac{F(t)}{X}\right) + \left(b(n)\tau + \frac{1}{2}(\sigma^2\tau + n\omega^2)\right)}{\sqrt{\sigma^2\tau + n\omega^2}}, \tag{325}$$

and

$$d_2 = d_1 - \sqrt{\sigma^2 \tau + n\omega^2}\,. \tag{326}$$

Substituting expression (323) into Merton's solution, expression (276), results in Bates' solution for an option contract written on an asset that follows a jump diffusion process. The expression is

$$C(FY,t) = \sum_{n=0}^{\infty} \frac{e^{-\lambda^* \tau}(\lambda^* \tau)^n}{n!}\left[e^{-r\tau} F(t) e^{b(n)\tau} N(d_1) - e^{-r\tau} X N(d_2)\right].$$

Further, simplifying the above yields

$$C(FY,t) = e^{-r\tau} \sum_{n=0}^{\infty} \frac{e^{-\lambda^* \tau}(\lambda^* \tau)^n}{n!}\left[F(t) e^{b(n)\tau} N(d_1) - X N(d_2)\right]. \tag{327}$$

Expression (327) is a closed form solution for an option contract written on futures price which is influenced by systematic jumps. To obtain the result, Bates has to make restrictions on preferences, technology, and distributions. This pricing formula is similar to Merton (1976) with the difference being the treatment of the jumps. Merton's formula presumes the jumps are idiosyncratic and has the actual parameters for the jump distribution in his equation. Alternatively, Bates allows for systematic jumps and has the risk-adjusted parameters in his formula. The Bates model is a more realistic depiction of commodity markets, and thereby, is a plausible model for pricing options written on commodity futures.

4.7 Hilliard and Reis (1998)

Hilliard and Reis (1998) investigates the pricing of options on commodity futures under stochastic convenience yields, interest rates and jumps in the spot price. The development of Hilliard and Reis' model draws on Bates' (1988, 1991) to provide a medium for pricing commodity futures options. Offering support for their choice, they state that commodity prices occasionally exhibit large discrete changes due to weather and other significant events. On average, one may expect that these jumps would have a non-zero mean. The implication is that the jumps systematically impact commodity prices and therefore cannot be ignored. The systematic nature of these discrete random jumps rules out Merton's (1976) model for pricing options written on commodity prices, since Merton presumes that the jumps are idiosyncratic. The assumption implyies the mean jump is zero and the risk is therefore ignored. Alternatively, Bates'(1988, 1991) model deals specifically with systematic jumps in the asset's price, thereby, making it a natural candidate for pricing options on commodity futures.

The option model given by Hilliard and Reis, is an extension to their three-factor futures model. The system of stochastic differential equations used by the authors is given below

$$\frac{dS(t)}{S(t)} = (r(t) - \delta(t) - \lambda^* \bar{k}^*)dt + \sigma_s dZ_s^*(t) + k^* dq^*, \tag{328}$$

$$d\delta(t) = \left(k_c(\alpha - \delta(t)) - \lambda\sigma_c\right)dt + \sigma_c dZ_c^*(t), \tag{329}$$

$$dr(t) = \left(f_t(s,t) + k_r f(s,t) + \frac{\sigma_r^2}{2k_r}\left(1 - e^{-2k_r(t-s)}\right) - k_r r(t) \right)dt + \sigma_r dZ_r^*(t). \tag{330}$$

where λ^*, k^*, and dq^* are described by Bates (1988,1991).The authors replace the spot price process in there three factor model

$$dS(t) = (r(t) - \delta(t))S(t)dt + \sigma_s S(t) dZ_s^*(t), \tag{331}$$

with expression (328) There is a problem with this system. In expression (328), the authors are using the results from Bates' model to begin their analysis. The parameters for Bates are the outcome of a general equilibrium model. That is, these parameters are determined endogenously. The problem here is that Hilliard and Reis wish to use these parameters exogenously. Furthermore, the authors are using Bates' parameters in a model that allows for a stochastic convenience yield and random interest rates. Bates' model never considered the impact of a convenience yield and he held the interest rate constant. The construction of Hilliard and Reis' model is not valid in terms of Bates' model.

Over looking the observation above, the authors posit that the futures price is a function of the above joint stochastic process (equations (331), (329) and (330)). Using Ito's lemma the authors express the increment for the futures price as

$$dF(t) = \frac{1}{2}F_{ss}[dS(t)]^2 + \frac{1}{2}F_{\delta\delta}[d\delta(t)]^2 \frac{1}{2}F_{rr}[dr(t)]^2 + F_s dS(t) + F_\delta d\delta(t) + F_r dr(t)$$

$$F_t dt + F_{s\delta} dS d\delta + F_{sr} dS dr + F_{r\delta} dr d\delta . \tag{332}$$

Substituting the differentials into (332) yields

$$\begin{aligned} dF(t) = {} & \frac{1}{2}F_{ss}\sigma_s^2[S(t)]^2 dt + \frac{1}{2}F_{\delta\delta}\sigma_c^2 dt + \frac{1}{2}F_{rr}\sigma_r^2 dt \\ & + F_s\left((r(t) - \delta(t) - \lambda^* \bar{k}^*)S(t)dt + \sigma_s S(t) dZ_s^*(t)\right) \\ & + F_\delta\left((k_c(\alpha - \delta(t)) - \lambda\sigma_c)dt + \sigma_c dZ_c^*(t)\right) \\ & + F_r\left(\left(f_t(s,t) + k_r f(s,t) + \frac{\sigma_r^2}{2k_r}\left(1 - e^{-2k_r(t-s)}\right) - k_r r(t)\right)dt + \sigma_r dZ_r^*(t)\right) \\ & + F_{s\delta}S(t)\sigma_s\sigma_c\rho_{sc}dt + F_{sr}S(t)\sigma_s\sigma_r\rho_{sr}dt + F_{r\delta}\sigma_r\sigma_c\rho_{rc}dt \\ & + F_t dt + F_s S(t) k^* dq^* . \end{aligned}$$

Rearranging the above yields

$$dF(t) = \left[\frac{1}{2}F_{ss}\sigma_s^2[S(t)]^2 + \frac{1}{2}F_{\delta\delta}\sigma_\delta^2 + \frac{1}{2}F_{rr}\sigma_r^2 + F_s\left((r(t)-\delta(t))S(t)\right)\right.$$

$$+ F_\delta\left(k_c(\alpha-\delta(t)) - \lambda\sigma_c\right) + F_r\left(f_t(s,t) + k_r f(s,t) + \frac{\sigma_r^2}{2k_r}\left(1-e^{-2k_r(t-s)}\right) - k_r r(t)\right)$$

$$\left. + F_{s\delta}S(t)\sigma_s\sigma_c\rho_{sc} + F_{sr}S(t)\sigma_s\sigma_r\rho_{sr} + F_{r\delta}\sigma_r\sigma_c\rho_{rc} + F_t\right]dt$$

$$+ \sigma_s S(t)dZ_s^*(t) + \sigma_c dZ_c^*(t) + \sigma_r dZ_r^*(t) - F_s S(t)\lambda^*\bar{k}^* dt + F_s S(t)k^* dq^*. \quad (333)$$

Under a risk neutral measure the bracket term in (333) should equal zero, thereby, reducing expression (333) to

$$dF(t) = F_s\sigma_s S(t)dZ_s^*(t) + F_\delta\sigma_c dZ_c^*(t) + F_r\sigma_r dZ_r^*(t) - F_s S(t)\lambda^*\bar{k}^* dt + F_s S(t)k^* dq^*. \quad (334)$$

Expression (334) is the stochastic differential reported by Hilliard and Reis. This price dynamic is incorrect. There are two faults with the above and it begins with the authors' use of Ito's lemma. Hilliard and Reis' treatment of Ito's lemma for the futures price is consistent if and only if the futures price is a function of continuous state variables. For the current case this is not true. Given the last two terms in expressions (333) and (334), we see that the authors do include the jump in the differential for the futures price. The correct version of Ito's lemma for their problem is

$$dF = \frac{1}{2}F_{SS}\left[dS_{dq=0}\right]^2 dt + \frac{1}{2}F_{\delta\delta}[d\delta(t)]^2 \frac{1}{2}F_{rr}[dr(t)]^2 + F_s dS_{dq=0} + F_\delta d\delta(t) + F_r dr(t)$$

$$+ F_t dt + F_{s\delta}dS_{dq=0}d\delta + F_{sr}dS_{dq=0}dr + F_{r\delta}drd\delta + F(SY,t) - F(S,t).^{15} \quad (335)$$

The author's, however, ignore this rule and simply introduce the jump into the futures price dynamic through the stochastic differential for the spot price. Again, this is shown by the last two terms in expression (333). Next, Hilliard and Reis set the risk adjusted drift of the futures price equal to zero. First, this operation is true only when no jumps are present. Secondly, the authors pull the term $F_s S(t)\lambda^*\bar{k}^* dt$ out of the drift before setting the drift to zero. This is mathematically inconsistent with theory of stochastic calculus.

The importance of the model is predicated on the futures price being influenced by random discrete jumps in the commodity's spot price. The authors should have developed the stochastic differential in expression (335) and not expression (332). The difference between these two differentials is the treatment of the jump. The last two terms in expression (335) indicate the impact in the futures price given a jump in the spot price. This change is a discrete change in the futures price. Expression (332) omits this term. To introduce the jump into the futures stochastic differential, the authors include it with the diffusion of the spot price when they make their substitutions into (332). Notice that this term is multiplied by the infinitesimal change in the futures price. This is completely inconsistent with how the futures price is changing given a jump in the spot price. That is, anytime a jump occurs there will be a discrete change in the futures price and not an infinitesimal change. The operation in expression (332) is simply an invalid operation.

After developing their stochastic differential for the futures price, Hilliard and Reis state that the jump in the futures price is exactly the same as the jump in the spot price. By making such a statement, the authors wish to take advantage of the results derived by Bates. Hilliard and Reis' statement, however, is incorrect. By definition the futures price is a contingent claim, whereby, it derives it value from an underlying spot commodity. In chapter three, we show that the futures price is a function of the spot price. We cannot believe that the actual jump in the spot commodity is the same in the futures price, because the jump should be introduced through some functional relationship. Indeed, inspection of the last term in expression (335) shows this to be correct.

Continuing with Hilliard and Reis, the authors further simplify expression (334) by substituting in for the partial derivatives of their three factor futures model *F(t)*. To do this the authors claim that futures prices are unchanged when a jump component is added to the spot price diffusion. That is, Hilliard and Reis argue that the futures price is simply the risk adjusted expected spot price and this value is the same whether the jump diffusion is added to the analysis or not. The authors verify this by taking the expectation of

$$\frac{dS(t)}{S(t)} = (r(t) - \delta(t) - \lambda^* \bar{k}^*)dt + \sigma dZ_s^*(t) + k^* dq^*,$$

and

$$\frac{dS(t)}{S(t)} = (r(t) - \delta(t))dt + \sigma dZ_s^*(t).$$

This is

$$\begin{aligned} E_t^*\left[\frac{dS(t)}{S(t)}\right] &= E_t^*\left[(r(t) - \delta(t) - \lambda^* \bar{k}^*)dt + \sigma dZ_s^*(t) + k^* dq^*\right], \\ &= (r(t) - \delta(t) - \lambda^* \bar{k}^*)dt + E_t^*\left[\sigma dZ_s^*(t)\right] + E_t^*\left[k^* dq^*\right], \\ &= (r(t) - \delta(t) - \lambda^* \bar{k}^*)dt + \lambda^* \bar{k}^* dt \\ &= (r(t) - \delta(t))dt. \end{aligned} \tag{336}$$

$$\begin{aligned} E_t^*\left[\frac{dS(t)}{S(t)}\right] &= E_t^*\left[(r(t) - \delta(t))dt + \sigma dZ_s^*(t)\right], \\ &= (r(t) - \delta(t))dt + E_t^*\left[\sigma dZ_s^*(t)\right], \\ &= (r(t) - \delta(t))dt. \end{aligned} \tag{337}$$

The authors state that since the expectation in expression (336) is the same as the expression in (337), the jump does not alter the risk-adjusted expected spot price. This is simply not correct. Recall from Merton's (1976) model that he presumes the jumps were diversifiable. This implies that the jumps are idiosyncratic and do not impact the market portfolio. Consequently, this allows him to ignore the risk associated with the jumps when it comes to pricing an option, but the jumps did impact the price of the option. Press (1967) shows that the expected value for the spot price at time T given the dynamics in expression (331) is

$$E_t\left[S(T)\right]=\sum_{n=0}^{\infty}\frac{e^{-\lambda^*\tau}(\lambda^*\tau)^n}{n!}\left[\int_{-\infty}^{\infty}s(T)f\left(s(T)\right)ds\mid n\ jumps\right].$$

This expectation above is not the same as the risk adjusted expected value for the terminal three-factor futures price given in chapter three. Here, the jump component does alter the expected value, and in fact the only way for this value to equal the expected value of the three-factor model is if $\lambda=0$.

If we presume Hilliard and Reis' assumptions about the futures price are correct, then taking the partial derivatives of equation (151) with respect to the spot price, convenience yield and interest rate yields

$$F_s=A(\tau)D_1(\tau)D_2(\tau)D_3(\tau)\exp\left(-H_c(\tau)\delta(t)\right)\frac{1}{P(t,T)}, \tag{338}$$

$$F_\delta=-H_c(\tau)S(t)A(\tau)D_1(\tau)D_2(\tau)D_3(\tau)\exp\left(-H_c(\tau)\delta(t)\right)\frac{1}{P(t,T)}, \tag{339}$$

$$F_r=H_r(\tau)S(t)A(\tau)D_1(\tau)D_2(\tau)D_3(\tau)\exp\left(-H_c(\tau)\delta(t)\right)\frac{1}{P(t,T)}. \tag{340}$$

Substituting (338)-(340) into (335) yields

$$dF(t) = \sigma_s S(t) dZ_s^*(t) A(\tau) D_1(\tau) D_2(\tau) D_3(\tau) \exp(-H_c(\tau)\delta(t)) \frac{1}{P(t,T)}$$

$$+ \sigma_c dZ_c^*(t) \left(-H_c(\tau) S(t) A(\tau) D_1(\tau) D_2(\tau) D_3(\tau) \exp(-H_c(\tau)\delta(t)) \frac{1}{P(t,T)} \right)$$

$$+ \sigma_r dZ_r^*(t) \left(H_r(\tau) S(t) A(\tau) D_1(\tau) D_2(\tau) D_3(\tau) \exp(-H_c(\tau)\delta(t)) \frac{1}{P(t,T)} \right)$$

$$- S(t) A(\tau) D_1(\tau) D_2(\tau) D_3(\tau) \exp(-H_c(\tau)\delta(t)) \frac{1}{P(t,T)} \lambda^* \bar{k}^* dt$$

$$+ S(t) A(\tau) D_1(\tau) D_2(\tau) D_3(\tau) \exp(-H_c(\tau)\delta(t)) \frac{1}{P(t,T)} k^* dq^*. \quad (341)$$

Simplifying

$$dF(t) = F(t)\sigma_s dZ_s^*(t) - F(t) H_c(\tau) \sigma_c dZ_c^*(t) + F(t) H_r(\tau) \sigma_r dZ_r^*(t)^*$$

$$- F(t) \lambda^* \bar{k}^* dt + F(t) k^* dq^*. \quad (342)$$

Define $\sigma_F(\tau) dZ_F^* = \sigma_s dZ_s^*(t) - H_c(\tau)\sigma_c dZ_c^*(t) + H_r(\tau)\sigma_r dZ_r^*(t)$. The above becomes

$$\frac{dF(t)}{F(t)} = \sigma_F(\tau) dZ_F^*(t) - \lambda^* \bar{k}^* dt + k^* dq^*$$

$$\frac{dF(t)}{F(t)} = -\lambda^* \bar{k}^* dt + \sigma_F(\tau) dZ_F^*(t) + k^* dq^*. \quad (343)$$

Hilliard and Reis (1998) claim to have found a process for the futures prices which is consistent with Bates (1991). Provided this observation, the authors use Bates' solution to price options written on the futures contract. As we discussed above, there are a couple of problems with Hilliard and Reis' assertions about the futures price process in equation (343). First, the percentage jump in the futures price will not be the same as the spot price. The only time this statement would hold is if there is no time left to maturity. Then, for this case by definition the futures price would equal the spot price. Secondly, it is true the futures price is linear in the spot price, but this does not translate to a perfect correlation in movements. The futures price is a function of the spot as well as other variables (time to maturity, convenience yield and interest rates). The movement in the spot price is one component of the futures price, and there is no reason to expect the futures price to behave identically with the spot price. In fact, Samuelson (1965) asserts the futures price volatility should be less than the spot. To conclude, the authors derive a result based on incorrect assumptions and mathematic operations. Any results derived from this analysis are tenuous at best.

4.8 One-factor jump-diffusion model

We see the analysis of Hilliard and Reis (1998) is flawed. If, however, one overlooks the problems with the methodology posited by these authors, we may use their insight to derive pricing models for the different futures prices derived in chapter three. Let us start with a one-factor model, where the futures price is influence only by the spot price. According to Hilliard and Reis, the spot price process is

$$\frac{dS}{S} = \left[k\left(\alpha^{*} - \ln S\right) - \lambda^{*}\bar{J}^{*}\right]dt + \sigma dZ_{s}^{*} + J^{*}dq^{*}, \tag{344}$$

where J^{*} now represents the percentage jump in the spot price. k represents the speed of adjustment for the spot price around its mean. The futures price is a twice continuously function of the spot price and time. Using Ito's lemma for the continuous part and analogous lemma for the jump part, the increment of the futures price may be expressed as

$$dF = \frac{1}{2}F_{ss}\left[dS_{dq=0}\right]^{2}dt + F_{s}dS_{dq=0} + F_{t}dt + F(SY,t) - F(S,t). \tag{345}$$

Substituting the expressions for the increments of the spot price into (345) and rearranging yields

$$dF = \frac{1}{2}F_{ss}\sigma^2 S^2 dt + F_s\left[\left(k\left(\alpha^* - \ln S\right) - \lambda^* \bar{J}^*\right)Sdt + \sigma S dZ_s^*\right] + F_t dt + F(SY,t) - F(S,t),$$

$$dF = \left[\frac{1}{2}F_{ss}\sigma^2 S^2 + F_s\left(k\left(\alpha^* - \ln S\right) - \lambda^* \bar{J}^*\right)S + F_t\right]dt + F_s\sigma S dZ_s^*$$

$$+ F(SY,t) - F(S,t) - \lambda E\left[F(SY,t) - F(S,t)\right]dt + \lambda E\left[F(SY,t) - F(S,t)\right]dt,$$

$$dF = \left[\frac{1}{2}F_{ss}\sigma^2 S^2 + F_s\left(k\left(\alpha^* - \ln S\right) - \lambda^* \bar{J}^*\right)S + F_t + \lambda E\left[F(SY,t) - F(S,t)\right]\right]dt + F_s\sigma S dZ_s^*$$

$$+ F(SY,t) - F(S,t) - \lambda E\left[F(SY,t) - F(S,t)\right]dt.$$

The expression above is the risk adjusted dynamic for the futures price. In a risk neutral world the futures price is a martingale process. Again, this implies that the drift term must equal zero, thus the expression above reduces to

$$dF = F_s\sigma S dZ_s^* + F(SY,t) - F(S,t) - \lambda E\left[F(SY,t) - F(S,t)\right]dt. \qquad (346)$$

Under the assumption that the jump does not alter the pricing formula for a futures contract, the partial of the futures price with respect to the spot price is

$$F_s = e^{-k\tau} S^{e^{-k\tau}-1} \exp\left[\alpha^*\left(1 - e^{-k\tau}\right) + \frac{\sigma^2}{4k}\left[1 - e^{-2k\tau}\right]\right].$$

This is given by expression (72) on page 143. Substituting the partial derivative into equation (346) yields

$$dF = e^{-k\tau} S S^{e^{-k\tau}-1} \exp\left[\alpha^*\left(1 - e^{-k\tau}\right) + \frac{\sigma^2}{4k}\left[1 - e^{-2k\tau}\right]\right]\sigma dZ_s^*$$

$$+ F(SY,t) - F(S,t) - \lambda E\left[F(SY,t) - F(S,t)\right]dt,$$

$$dF = e^{-k\tau} S^{e^{-k\tau}} \exp\left[\alpha^*\left(1-e^{-k\tau}\right)+\frac{\sigma^2}{4k}\left[1-e^{-2k\tau}\right]\right]\sigma dZ_s$$

$$+ F(SY,t) - F(S,t) - \lambda E[F(SY,t)-F(S,t)]dt ,$$

$$dF = Fe^{-k\tau}\sigma dZ_s + F(SY,t) - F(S,t) - \lambda E[F(SY,t)-F(S,t)]dt . \tag{347}$$

Dividing the left-hand side and right-hand side of expression (347) yields

$$\frac{dF}{F} = e^{-k\tau}\sigma dZ_s + \frac{F(SY,t)-F(S,t)}{F} - \frac{\lambda E[F(SY,t)-F(S,t)]}{F}dt ,$$

$$\frac{dF}{F} = e^{-k\tau}\sigma dZ_s^* + J_F^* dq^* - \lambda^* \bar{J}_F^* dt ,$$

$$\frac{dF}{F} = -\lambda^* \bar{J}_F^* dt + e^{-k\tau}\sigma dZ_s^* + J_F^* dq^* . \tag{348}$$

Expression (348) is equivalent to Bates (1988, 1991). Therefore, we may use Bates' option pricing model to price options written on commodity futures.

The option price according to Bates (1988, 1991) is

$$C(FY,t) = e^{-r\tau} \sum_{n=0}^{\infty} \frac{e^{-\lambda^*\tau}(\lambda^*\tau)^n}{n!}\left[F(t)e^{b(n)\tau}N(d_1) - XN(d_2)\right], \tag{349}$$

where $b(n) = \left(-\lambda^*\bar{J}_F^* + \frac{n\gamma^*}{\tau}\right)$, (350)

$$d_1 = \frac{\ln\left(\frac{F(t)}{X}\right) + \left(\left(-\lambda^*\bar{J}_F^* + \frac{n\gamma^*}{\tau}\right)\tau + \frac{1}{2}(\sigma^2\tau + n\omega^2)\right)}{\sqrt{\sigma^2\tau + n\omega^2}}, \tag{351}$$

and

$$d_2 = d_1 - \sqrt{\sigma^2\tau + n\omega^2} . \tag{352}$$

We need to tailor this expression according to the price dynamic given in equation (348). The bracket term in expression (349) is the Black-Scholes option pricing formula. Once this value is evaluated, we then have a solution for the option.

To price the option we first need to determine the terminal futures price. Recall the terminal futures price we are interested in is the futures price when the option expires and not when the futures expires. Let $H(t) = \ln F(t)$. The transformation, and Ito's lemma for a jump diffusion yields the following process for $H(t)$

$$dH = H_F dF_{dq=o} + \frac{1}{2} H_{FF} \left[dF_{dq=0}\right]^2 + H(FY,t) - H(F,t),$$

$$dH = \frac{1}{F}(F)\left(-\lambda^* \bar{J}_F^* dt + e^{-k\tau} \sigma dZ^*(t)\right) - \frac{1}{2}\frac{1}{F^2}\left(F^2 e^{-2k\tau}\sigma^2 dt\right) + \ln FY - \ln F,$$

$$dH = \left(-\lambda^* \bar{J}_F^* dt + e^{-k\tau}\sigma dZ^*(t)\right) - \frac{1}{2}\left(e^{-2k\tau}\sigma^2 dt\right) + \ln\left(\frac{FY}{F}\right),$$

$$dH = \left(-\lambda^* \bar{J}_F^* - \frac{1}{2} e^{-2k\tau}\sigma^2\right)dt + \sigma dZ^*(t) + \ln(Y)$$

$$\int_t^{T_1} dH = \int_t^{T_1}\left(-\lambda^* \bar{J}_F^* - \frac{1}{2} e^{-2k(T-s)}\sigma^2\right)ds + \sigma\int_t^{T_1} e^{-k(T-s)} dZ^*(s) + \sum_{i=1}^{n}\ln(Y_i),$$

$$H(T_1) - H(t) = \int_t^{T_1}\left(-\lambda^* \bar{J}_F^* - \frac{1}{2} e^{-2k(T-s)}\sigma^2\right)ds + \sigma\int_t^{T_1} e^{-k(T-s)} dZ^*(s) + \sum_{i=1}^{n}\ln(Y_i),$$

$$H(T_1) - H(t) = -\lambda^* \bar{J}_F^*(T_1 - t) - \frac{1}{2}\int_t^{T_1} e^{-2k(T-s)}\sigma^2 ds + \int_t^{T_1} e^{-k(T-s)}\sigma dZ^*(s) + \sum_{i=1}^{n}\ln(Y_i).$$

Recall from section 4.2 that $v^2 = \int_t^{T_1} e^{-2k(T-s)}\sigma^2 ds$. Therefore the above reduces to

$$H(T_1) = H(t) + -\lambda^* \bar{J}_F^* \tau_1 - \frac{1}{2}v^2 + \int_t^{T_1} e^{-k(T-s)}\sigma dZ^*(s) + \sum_{i=1}^{n}\ln(Y_i),$$

$$F(T_1) = F(t)\exp\left[-\lambda^* \bar{J}_F^* \tau_1 - \frac{1}{2}v^2 + \int_t^{T_1} e^{-k(T-s)}\sigma dZ^*(s)\right]Y(n),$$

$$F(T_1) = F(t)\exp[\alpha]Y(n), \qquad (353)$$

where $\alpha = \left(-\lambda^* \bar{J}_F^* \tau_1 - \frac{1}{2}v^2 + \int_t^{T_1} e^{-k(T-s)}\sigma dZ^*(s) \right)$, and $Y(n) = \prod_{i=1}^{n} Y_i$. Again, the terminal futures price, given in equation (353) is still a stochastic process. We may obtain a solution for (353), by finding the expected terminal futures price.

The expected terminal futures price is

$$E_t^*[F(T_1) \mid n\ jumps] = \exp\left[E_t^*[H(T_1) \mid n\ jumps] + \frac{1}{2}V_t^*[H(T_1) \mid n\ jumps] \right]. \quad (354)$$

The conditional mean of $H(T_1)$ is

$$E_t^*[H(T_1) \mid n\ jumps] = H(t) + -\lambda^* \bar{J}_F^* \tau_1 - \frac{1}{2}v^2 + \sum_{i=1}^{n} E_n^*[\ln(Y_i)].$$

Recall on page 215 that $\ln(1 + k_F^*) \sim N\left(\gamma^* - \frac{1}{2}\omega^2, \omega^2 \right)$, and that $Y = 1 + k_F^*$. Therefore the expectation operator $E_n^*[\ ]$ yields

$$= H(t) - \lambda^* \bar{J}_F^* \tau_1 - \frac{1}{2}v^2 + \sum_{i=1}^{n}\left(\gamma^* - \frac{1}{2}\omega^2 \right),$$

$$= H(t) - \lambda^* \bar{J}_F^* \tau_1 - \frac{1}{2}v^2 + n\left(\gamma^* - \frac{1}{2}\omega^2 \right). \qquad (355)$$

The conditional variance of $H(T_1)$ is

$$V_t^*[H(T_1) \mid n\ jumps] = V_t^*\left[\int_t^{T_1} e^{-k(T-s)}\sigma dZ^*(s) \right] + \sum_{i=1}^{n} V_n^*[\ln(Y_i)].$$

Recall the jump term and the Wiener process are uncorrelated. Continuing,

$$V_t^*[H(T_1) \mid n\ jumps] = \int_t^{T_1} e^{-2k(T-s)}\sigma^2 ds + \sum_{i=1}^{n} \omega^2,$$

$$= v^2 + \sum_{i=1}^{n} \omega^2 ,$$

$$= v^2 + n\omega^2 . \quad (356)$$

Substituting expressions (355) and (356) into (354) yields

$$E_t^*[F(T_1) \mid n\ jumps] = \exp\left[H(t) - \lambda^* \bar{J}_F^* \tau_1 - \frac{1}{2}v^2 + n\left(\gamma^* - \frac{1}{2}\omega^2\right) + \frac{1}{2}\left(v^2 + n\omega^2\right)\right],$$

$$= \exp\left[H(t) - \lambda^* \bar{J}_F^* \tau_1 + n\gamma^* - \frac{1}{2}v^2 + \frac{1}{2}v^2 - \frac{1}{2}n\omega^2 + \frac{1}{2}n\omega^2\right],$$

$$= \exp\left[H(t) - \lambda^* \bar{J}_F^* \tau_1 + n\gamma^*\right],$$

$$= \exp\left[H(t) + \left(-\lambda^* \bar{J}_F^* + \frac{n\gamma^*}{\tau_1}\right)\tau_1\right],$$

$$= F(t)\exp\left[\left(-\lambda^* \bar{J}_F^* + \frac{n\gamma^*}{\tau_1}\right)\tau_1\right],$$

$$= F(t)\exp\left[b(n)\tau_1\right], \quad (357)$$

where $b(n) = -\lambda^* \bar{J}_F^* + \frac{n\gamma^*}{\tau_1}$. Expression (357) is the expected futures price for the instantaneous movements in the futures price given n jumps have occurred. Recall, the instantaneous volatility for the one-factor model in section 4.2 is given by expression (82). This is

$$v^2 = \frac{\sigma^2}{2k}\left(e^{-2k(T-T_1)} - e^{-2k(T-t)}\right).$$

Substituting this expression into expression (356) yields

$$V_t^*[H(T) \mid n\ jumps] = \frac{\sigma^2}{2k}\left(e^{-2k(T-T_1)} - e^{-2k(T-t)}\right) + n\omega^2 .$$

$$= \frac{\sigma^2}{2k}\left[e^{-2k\tau_1} - e^{-2k\tau}\right] + n\omega^2. \tag{358}$$

Expression (358) is the volatility of the futures price over the life of the option contract. Expressions (357) and (358) give us the characteristics of the terminal futures price at the time of expiration of the option contract. Using these expressions in Bates' option pricing formula yields

$$C(FY,t) = P(t,T)\sum_{n=0}^{\infty}\frac{e^{-\lambda^*\tau_1}(\lambda^*\tau_1)^n}{n!}\left[F(t)e^{b(n)\tau_1}N(d_1) - XN(d_2)\right], \tag{359}$$

where $b(n) = \left(-\lambda^*\bar{J}_F^* + \frac{n\gamma^*}{\tau_1}\right)$, (360)

$$d_1 = \frac{\ln\left(\frac{F(t)}{X}\right) + \left(\left(-\lambda^*\bar{J}_F^* + \frac{n\gamma^*}{\tau_1}\right)\tau_1 + \frac{1}{2}\left(v^2 + n\omega^2\right)\right)}{\sqrt{v^2 + n\omega^2}}, \tag{361}$$

and

$$d_2 = d_1 - \sqrt{v^2 + n\omega^2}. \tag{362}$$

Expression (359) is the option pricing formula for an option written on a futures contract, whose price follows the jump-diffusion in expression (348).

4.9 Two-factor jump-diffusion model

In chapter three section 3.2 on page 38, we show that if the futures price is influenced by more than one state variable (the spot price), then a different solution exists for the futures price. In particular, we investigate the effect a stochastic convenience yield has on the futures price. In section 4.3 of this chapter, we price options written on a two-factor futures price. Now, we investigate an option pricing model that considers the two-factor futures price, where the futures price is influenced by a stochastic convenience yield and a jump-diffusion spot price. Again, we assume that the assumptions from Hilliard and Reis are applicable. The system of stochastic differential equations is

$$\frac{dS(t)}{S(t)} = \left(r - \delta(t) - \lambda^* \bar{J}^*\right)dt + \sigma_s dZ_s^*(t) + J^* dq^* \tag{363}$$

$$d\delta(t) = k(\alpha - \delta(t))dt + \sigma_c dZ_c^*(t). \tag{364}$$

If the futures price is function of the spot price, convenience yield, and time, namely $F(SY, \delta, t)$, then using Ito's lemma, we may express the futures price dynamic as

$$\begin{aligned} dF = &\frac{1}{2} F_{ss}\sigma_s^2 S^2 dt + F_s\left[\left(r - \delta(t) - \lambda^* \bar{J}^*\right)Sdt + \sigma_s S dZ_s^*\right] + \frac{1}{2} F_{\delta\delta}\sigma_c^2 dt \\ &+ F_\delta\left[k(\alpha - \delta(t))dt + \sigma_c dZ_c^*(t)\right] + F_{s\delta}\rho_{sc}\sigma_s\sigma_c dt + F_t dt \\ &+ F(SY, t) - F(S, t). \end{aligned} \tag{365}$$

We may rewrite expression (365) by adding and subtracting $\lambda E[F(SY,t) - F(S,t)]dt$. This yields

$$\begin{aligned} dF = &\frac{1}{2} F_{ss}\sigma_s^2 S^2 dt + F_s\left[\left(r - \delta(t) - \lambda^* \bar{J}^*\right)Sdt + \sigma_s S dZ_s^*\right] + \frac{1}{2} F_{\delta\delta}\sigma_c^2 dt \\ &+ F_\delta\left[k(\alpha - \delta(t))dt + \sigma_c dZ_c^*(t)\right] + F_{s\delta}\rho_{sc}\sigma_s\sigma_c dt + F_t dt \\ &+ F(SY,t) - F(S,t) - \lambda E[F(SY,t) - F(S,t)]dt + \lambda E[F(SY,t) - F(S,t)]dt. \end{aligned}$$

Substituting in for the increments of the spot price and convenience yield and rearranging the above yields

$$\begin{aligned} dF = &\frac{1}{2} F_{ss}\sigma_s^2 S^2 dt + \frac{1}{2} F_{\delta\delta}\sigma_c^2 dt + F_s\left(r - \delta(t) - \lambda^* \bar{J}^*\right)Sdt + F_\delta\left[k(\alpha - \delta(t))dt\right] \\ &+ F_{s\delta}\rho_{sc}\sigma_s\sigma_c dt + F_t dt + \lambda E[F(SY,t) - F(S,t)]dt + F_s\sigma_s S dZ_s^* + F_d\sigma_c dZ_c^*(t) \\ &+ F(SY,t) - F(S,t) - \lambda E[F(SY,t) - F(S,t)]dt, \end{aligned}$$

$$dF = \left[\frac{1}{2}F_{ss}\sigma_s^2 S^2 + \frac{1}{2}F_{\delta\delta}\sigma_c^2 + F_s\left(r - \delta(t) - \lambda^* \bar{J}^*\right)S + F_\delta k(\alpha - \delta(t))\right.$$

$$\left. + F_{s\delta}\rho_{sc}\sigma_s\sigma_c + F_t + \lambda E[F(SY,t) - F(S,t)]\right]dt + F_s\sigma_s S dZ_s^* + F_d\sigma_c dZ_c^*(t)$$

$$+ F(SY,t) - F(S,t) - \lambda E[F(SY,t) - F(S,t)]dt. \qquad (366)$$

Expression (366) is the risk adjusted dynamic for the futures price. In a risk neutral world the futures price is a martingale process. Again, this implies that the drift term must equal zero, thereby, making the above equal to

$$dF = F_s\sigma_s S dZ_s^* + F_\delta\sigma_c dZ_c^*(t) + F(SY,t) - F(S,t) - \lambda E[F(SY,t) - F(S,t)]dt. \qquad (367)$$

The above expression may be further reduced, by substituting in for the partial derivatives of the two-factor futures price reported by expression (94) in chapter three. These partial derivatives are given by expressions (127) and (128) above. The partial with respect to the spot price is

$$F_s = \exp\left\{-\left[\frac{1}{2}\sigma_s^2 + \alpha - \frac{1}{k}\sigma_c\lambda\right](T-t) - \frac{1}{k}\left(\delta(t) - \alpha + \frac{1}{k}\sigma_c\lambda\right)(1-\theta)\right.$$
$$+\frac{1}{2}\left[\left(\sigma_s^2 - 2\frac{1}{k}\sigma_c\sigma_s\rho_{cs} + \left(\frac{1}{k}\right)^2\sigma_c^2\right)(T-t)\right.$$
$$\left.\left. + 2\left(\left(\frac{1}{k}\right)^2\sigma_c\sigma_s\rho_{cs} - \left(\frac{1}{k}\right)^3\sigma_c^2\right)(1-\theta) + \left(\frac{1}{k}\right)^2\frac{\sigma_c^2}{2k}\left(1-\theta^2\right)\right]\right\}.$$

(368)

The partial derivative of the futures price with respect to the convenience yield is

$$F_\delta = -\left(\frac{1}{k}\right)(1-\theta)F = -H_c(\tau)F, \qquad (369)$$

where $H_c(\tau) = -\frac{\left(1-e^{-k\tau}\right)}{k}$. Substituting expressions (368) and (369) into (367) yields

$$dF = \left[S(t)\exp\left\{ -\left[\frac{1}{2}\sigma_s^2 + \alpha - \frac{1}{k}\sigma_c\lambda\right](T-t) - \frac{1}{k}\left(\delta(t) - \alpha + \frac{1}{k}\sigma_c\lambda\right)(1-\theta) \right.\right.$$

$$+\frac{1}{2}\left[\left(\sigma_s^2 - 2\frac{1}{k}\sigma_c\sigma_s\rho_{cs} + \left(\frac{1}{k}\right)^2\sigma_c^2\right)(T-t)\right.$$

$$\left.\left.\left. +2\left(\left(\frac{1}{k}\right)^2\sigma_c\sigma_s\rho_{cs} - \left(\frac{1}{k}\right)^3\sigma_c^2\right)(1-\theta) + \left(\frac{1}{k}\right)^2\frac{\sigma_c^2}{2k}\left(1-\theta^2\right)\right]\right\}\right]\sigma_s dZ_s^*$$

$$-H_c(\tau)F\sigma_c dZ_c^*(t) + F(SY,t) - F(S,t) - \lambda E\left[F(SY,t) - F(S,t)\right]dt,$$

$$dF = F\sigma_s dZ_s^*(t) - H_c(\tau)F\sigma_c dZ_c^*(t) + F(SY,t) - F(S,t) - \lambda E\left[F(SY,t) - F(S,t)\right]dt. \tag{370}$$

Dividing the left-hand side and right-hand side of expression (370) by F yields

$$\frac{dF}{F(S,t)} = \sigma_s dZ_s^*(t) - H_c(\tau)\sigma_c dZ_c^*(t) + \frac{F(SY,t) - F(S,t)}{F(S,t)} - \frac{\lambda E\left[F(SY,t) - F(S,t)\right]}{F(S,t)}dt,$$

$$\frac{dF}{F(S,t)} = \sigma_s dZ_s^*(t) - H_c(\tau)\sigma_c dZ_c^*(t) + J_F^* dq - \lambda^* \bar{J}_F^* dt,$$

$$\frac{dF}{F(S,t)} = -\lambda^* \bar{J}_F^* dt + \sigma_s dZ_s^*(t) - H_c(\tau)\sigma_c dZ_c^*(t) + J_F^* dq, \tag{371}$$

Now if we let

$$\sigma_F dZ_F^*(t) \equiv \sigma_s dZ_s^*(t) - H_c(\tau)\sigma_c dZ_c^*(t), \tag{372}$$

then, expression (371) becomes

$$\frac{dF}{F(S,t)} = -\lambda^* \bar{J}_F^* dt + \sigma_F dZ_F^* + J_F^* dq^*. \tag{373}$$

Equation (373) is the price dynamic for a futures price, which is influenced by a stochastic convenience yield and a jump-diffusion spot price process.

Given the dynamic, expression in (373), we may find the terminal futures price. With this, Let $H = \ln F$. The transformation and Ito's lemma gives

$$dH = H_F dF_{dq=o} + \frac{1}{2} H_{FF} \left[dF_{dq=0} \right]^2 + H(FY,t) - H(F,t),$$

$$dH = H_F \left[-\lambda^* \bar{J}_F^* F dt + \sigma_F F dZ_F^* \right] + \frac{1}{2} H_{FF} \sigma_F^2 F^2 dt + \ln FY - \ln F,$$

$$dH = \frac{1}{F} \left[-\lambda^* \bar{J}_F^* F dt + \sigma_F F dZ_F^* \right] + \frac{1}{2} \left(-\frac{1}{F^2} \right) \sigma_F^2 F^2 dt + \ln \frac{FY}{F},$$

$$dH = -\lambda^* \bar{J}_F^* dt + \sigma_F dZ_F^* - \frac{1}{2} \sigma_F^2 dt + \ln Y,$$

$$dH = \left(-\lambda^* \bar{J}_F^* - \frac{1}{2} \sigma_F^2 \right) dt + \sigma_F dZ_F^* + \ln Y,$$

$$\int_t^{T_1} dH = \int_t^{T_1} \left(-\lambda^* \bar{J}_F^* - \frac{1}{2} \sigma_F^2 \right) ds + \int_t^{T_1} \sigma_F dZ_F^* + \sum_{i=1}^{n} \ln Y_i,$$

$$H(T_1) - H(t) = -\lambda^* \bar{J}_F^* (T_1 - t) - \int_t^{T_1} \frac{1}{2} \sigma_F^2 ds + \int_t^{T_1} \sigma_F dZ_F^* + \sum_{i=1}^{n} \ln Y_i$$

$$H(T_1) = H(t) - \lambda^* \bar{J}_F^* (T_1 - t) - \int_t^{T_1} \frac{1}{2} \sigma_F^2 ds + \int_t^{T_1} \sigma_F dZ_F^* + \sum_{i=1}^{n} \ln Y_i,$$

$$F(T_1) = F(t) \exp\left[-\lambda^* \bar{J}_F^* (T_1 - t) - \int_t^{T_1} \frac{1}{2} \sigma_F^2 ds + \int_t^{T_1} \sigma_F dZ_F^* \right] Y(n),$$

$$F(T_1) = F(t) \exp[\alpha] Y(n), \tag{374}$$

where $\alpha = -\lambda^* \bar{J}_F^* (T_1 - t) - \int_t^{T_1} \frac{1}{2} \sigma_F^2 ds + \int_t^{T_1} \sigma_F dZ_F^*$ and $Y(n) = \prod_{i=1}^{n} Y_i$. Under a risk neutral measure the expected futures price is

$$E_t^* [F(T_1) \mid n\ jumps] = \exp\left[E_t^* [H(T_1) \mid n\ jumps] + \frac{1}{2} V_t^* [H(T_1) \mid n\ jumps] \right]. \tag{375}$$

The expected value of $H(T_1)$ is

$$E_t^*\left[H(T_1)\,|\,n\ jumps\right] = H(t) - \lambda^* \bar{J}_F^*(T_1 - t) - \int_t^{T_1} \frac{1}{2}\sigma_F^2 ds + \sum_{i=1}^{n} E_n^*[\ln Y_i]$$

$$= H(t) - \lambda^* \bar{J}_F^*(T_1 - t) - \int_t^{T_1} \frac{1}{2}\sigma_F^2 ds + \sum_{i=1}^{n}\left(\gamma^* - \frac{1}{2}\omega^2\right),$$

$$= H(t) - \lambda^* \bar{J}_F^*(T_1 - t) - \int_t^{T_1} \frac{1}{2}\sigma_F^2 ds + n\left(\gamma^* - \frac{1}{2}\omega^2\right). \quad (376)$$

The conditional variance of $H(T_1)$ is

$$V_t^*\left[H(T_1)\,|\,n\ jumps\right] = V_t^*\left[\int_t^{T_1} \sigma_F dZ^*(s)\right] + \sum_{i=1}^{n} V_n^*[\ln(Y_i)].$$

Recall the jump term and the Wiener process are uncorrelated.

$$V_t^*\left[H(T_1)\,|\,n\ jumps\right] = V_t^*\left[\int_t^{T_1}\left(\sigma_s dZ_s(s) - H_c(\tau)\sigma_c dZ_c(s)\right)\right] + \sum_{i=1}^{n} V_n^*[\ln(Y_i)]$$

$$= \left[\int_t^{T_1} \sigma_s^2 ds + \int_t^{T_1} \sigma_c^2 H_c^2(\tau) ds - \int_t^{T_1} 2\rho_{sc}\sigma_s\sigma_c H_c(\tau) ds\right] + \sum_{i=1}^{n} \omega^2 .$$

Substituting the results from equations (138)-(140) into the above yields

$$V_t^*\left[H(T_1)\,|\,n\ jumps\right] = \sigma_s^2(T_1 - t)$$

$$+ \frac{\sigma_c^2}{k_c^2}\left[(T_1 - t) + \frac{1}{2k_c}\left(e^{-2k_c(T-T_1)} - e^{-2k_c(T-t)}\right) - \frac{2}{k_c}\left(e^{-k_c(T-T_1)} - e^{-k_c(T-t)}\right)\right]$$

$$- \frac{2\rho_{sc}\sigma_s\sigma_c}{k_c}\left[(T_1 - t) - \frac{1}{k_c}\left(e^{-k_c(T-T_1)} - e^{-k_c(T-t)}\right)\right] + \sum_{i=1}^{n}\omega^2 . \quad (377)$$

Recall from expression (141) on page 164 of this chapter that

$$v^2 = \sigma_s^2(T_1 - t) + \frac{\sigma_c^2}{k_c^2}\left[(T_1 - t) + \frac{1}{2k_c}\left(e^{-2k_c(T-T_1)} - e^{-2k_c(T-t)}\right) - \frac{2}{k_c}\left(e^{-k_c(T-T_1)} - e^{-k_c(T-t)}\right)\right]$$

$$-\frac{2\rho_{sc}\sigma_s\sigma_c}{k_c}\left[(T_1-t)-\frac{1}{k_c}\left(e^{-k_c(T-T_1)}-e^{-k_c(T-t)}\right)\right].$$

Therefore, we may reduce (377) to

$$V_t^*[H(T_1)\mid n\ jumps]=v^2+\sum_{i=1}^{n}\omega^2,$$

$$=v^2+n\omega^2. \tag{378}$$

Substituting expressions (376) and (378) into (375) gives us a expected futures price as

$$E_t^*[F(T_1)\mid n\ jumps]=\exp\left[H(t)-\lambda^*\bar{J}_F^*(T_1-t)-v^2+n\gamma^*+\frac{1}{2}\left(v^2+n\omega^2\right)\right],$$

$$=\exp\left[H(t)-\lambda^*\bar{J}_F^*(T_1-t)+n\gamma^*-\frac{1}{2}v^2+\frac{1}{2}v^2-n\frac{1}{2}\omega^2+\frac{1}{2}n\omega^2\right],$$

$$=\exp\left[H(t)-\lambda^*\bar{J}_F^*(T_1-t)+n\gamma^*\right],$$

$$=F(t)\exp\left[-\lambda^*\bar{J}_F^*\tau_1+n\gamma^*\right],$$

$$=F(t)\exp\left[\left(-\lambda^*\bar{J}_F^*+\frac{n\gamma^*}{\tau_1}\right)\tau_1\right],$$

$$=F(t)\exp\left[b(n)\tau_1\right], \tag{379}$$

where $b(n)=-\lambda^*\bar{J}_F^*+\frac{n\gamma^*}{\tau_1}$. Equation (379) is the expected futures price at time T_1 given that n jumps occurred. Substituting expression (379) into Bates' formula yields

$$C(FY,t)=e^{-r\tau_1}\sum_{n=0}^{\infty}\frac{e^{-\lambda^*\tau_1}(\lambda^*\tau_1)^n}{n!}\left[F(t)e^{b(n)\tau_1}N(d_1)-XN(d_2)\right], \tag{380}$$

where $b(n)=\left(-\lambda^*\bar{J}_F^*+\frac{n\gamma^*}{\tau_1}\right),$ (381)

$$d_1 = \frac{\ln\left(\frac{F(t)}{X}\right) + \left(\left(-\lambda^* \bar{J}_F^* + \frac{n\gamma^*}{\tau_1}\right)\tau_1 + \frac{1}{2}\left(v^2 + n\omega^2\right)\right)}{\sqrt{v^2 + n\omega^2}}, \tag{382}$$

and

$$d_2 = d_1 - \sqrt{v^2 + n\omega^2}. \tag{383}$$

Expression (380) is the option pricing formula for an option written on a futures contract, whose price follows the jump-diffusion in expression (373).

4.10 Three-factor jump-diffusion model

The last model we consider is a three-factor jump diffusion model presented in Hilliard and Reis (1998). Here the futures price is a function of the following system of stochastic differential equations (these were stated earlier on page 231)

$$\frac{dS(t)}{S(t)} = (r(t) - \delta(t) - \lambda^* \bar{J}^*)dt + \sigma_s dZ_s^*(t) + J^* dq^*, \tag{384}$$

$$d\delta(t) = \left(k_c(\alpha - \delta(t)) - \lambda_c \sigma_c\right)dt + \sigma_c dZ_c^*(t), \tag{385}$$

$$dr(t) = \left(f_t(s,t) + k_r f(s,t) + \frac{\sigma_r^2}{2k_r}\left(1 - e^{-2k_r(t-s)}\right) - k_r r(t)\right)dt + \sigma_r dZ_r^*(t). \tag{386}$$

The increment of the futures price can be expressed using Ito's lemma. This is

$$dF(t) = \frac{1}{2}F_{ss}\left[dS(t)_{dq=o}\right]^2 + \frac{1}{2}F_{\delta\delta}\left[d\delta(t)\right]^2 + \frac{1}{2}F_{rr}\left[dr(t)\right]^2 + F_s dS(t)_{dq=o} + F_\delta d\delta(t) + F_r dr(t)$$

$$F_t dt + F_{s\delta} dS d\delta + F_{sr} dS dr + F_{r\delta} dr d\delta + F(SY,t) - F(S,t). \tag{387}$$

Substituting the stochastic differentials in expression (384)-(386) into (387) yields

$$dF(t) = \frac{1}{2}F_{ss}\sigma_s^2\left[S(t)_{dq=0}\right]^2 dt + \frac{1}{2}F_{\delta\delta}\sigma_c^2 dt + \frac{1}{2}F_{rr}\sigma_r^2 dt$$

$$+ F_s\left((r(t) - \delta(t) - \lambda^* \bar{k}^*)S(t)dt + \sigma_s S(t) dZ_s^*(t)\right)$$

$$+F_\delta\left(\left(k_c(\alpha-\delta(t))-\lambda_c\sigma_c\right)dt+\sigma_c dZ_c^*(t)\right)$$

$$+F_r\left(\left(f_t(s,t)+k_r f(s,t)+\frac{\sigma_r^2}{2k_r}\left(1-e^{-2k_r(t-s)}\right)-k_r r(t)\right)dt+\sigma_r dZ_r^*(t)\right)$$

$$+F_{s\delta}S(t)\sigma_s\sigma_c\rho_{sc}dt+F_{sr}S(t)\sigma_s\sigma_r\rho_{sr}dt+F_{r\delta}\sigma_r\sigma_c\rho_{rc}dt$$

$$+F_t dt+F(SY,t)-F(S,t).$$

Rearranging the above yields

$$dF(t)=\left[\frac{1}{2}F_{ss}\sigma_s^2[S(t)]^2+\frac{1}{2}F_{\delta\delta}\sigma_c^2+\frac{1}{2}F_{rr}\sigma_r^2+F_s\left((r(t)-\delta(t)-\lambda^*\bar{k}^*)S(t)\right)\right.$$

$$+F_\delta\left(k_c(\alpha-\delta(t))-\lambda_c\sigma_c\right)+F_r\left(f_t(s,t)+k_r f(s,t)+\frac{\sigma_r^2}{2k_r}\left(1-e^{-2k_r(t-s)}\right)-k_r r(t)\right)$$

$$\left.+F_{s\delta}S(t)\sigma_s\sigma_c\rho_{sc}+F_{sr}S(t)\sigma_s\sigma_r\rho_{sr}+F_{r\delta}\sigma_r\sigma_c\rho_{rc}+F_t\right]dt$$

$$+F_s\sigma_s S(t)dZ_s^*(t)+F_\delta\sigma_c dZ_c^*(t)+F_r\sigma_r dZ_r^*(t)+F(SY,t)-F(S,t)$$

$$+\lambda E[F(SY,t)-F(S,t)]dt-\lambda E[F(SY,t)-F(S,t)]dt,$$

$$dF(t)=\left[\frac{1}{2}F_{ss}\sigma_s^2[S(t)]^2+\frac{1}{2}F_{\delta\delta}\sigma_\delta^2+\frac{1}{2}F_{rr}\sigma_r^2+F_s\left((r(t)-\delta(t)-\lambda^*\bar{k}^*)S(t)\right)\right.$$

$$+F_\delta\left(k_c(\alpha-\delta(t))-\lambda_c\sigma_c\right)+F_r\left(f_t(s,t)+k_r f(s,t)+\frac{\sigma_r^2}{2k_r}\left(1-e^{-2k_r(t-s)}\right)-k_r r(t)\right)$$

$$\left.+F_{s\delta}S(t)\sigma_s\sigma_c\rho_{sc}+F_{sr}S(t)\sigma_s\sigma_r\rho_{sr}+F_{r\delta}\sigma_r\sigma_c\rho_{rc}+F_t+\lambda E[F(SY,t)-F(S,t)]dt\right]dt$$

$$+F_s\sigma_s S(t)dZ_s^*(t)+F_\delta\sigma_c dZ_c^*(t)+F_r\sigma_r dZ_r^*(t)+F(SY,t)-F(S,t)$$

$$-\lambda E[F(SY,t)-F(S,t)]dt. \tag{388}$$

Under risk neutrality, the futures price is a martingale process implying the drift term in (388) is equal to zero. Expression (388) reduces to

$$dF(t)=F_s\sigma_s S(t)dZ_s^*(t)+F_\delta\sigma_c dZ_c^*(t)+F_r\sigma_r dZ_r^*(t).$$

$$+F(SY,t)-F(S,t)-\lambda E[F(SY,t)-F(S,t)]dt\,, \qquad (389)$$

Using the partial derivatives in equations (338)-(340) in (389) yields

$$dF(t)=\sigma_s S(t)dZ_s^*(t)A(\tau)D_1(\tau)D_2(\tau)D_3(\tau)\exp(-H_c(\tau)\delta(t))\frac{1}{P(t,T)}$$

$$+\sigma_c dZ_c^*(t)\left(-H_c(\tau)S(t)A(\tau)D_1(\tau)D_2(\tau)D_3(\tau)\exp(-H_c(\tau)\delta(t))\frac{1}{P(t,T)}\right)$$

$$+\sigma_r dZ_r^*(t)\left(H_r(\tau)S(t)A(\tau)D_1(\tau)D_2(\tau)D_3(\tau)\exp(-H_r(\tau)\delta(t))\frac{1}{P(t,T)}\right)$$

$$+F(SY,t)-F(S,t)-\lambda E[F(SY,t)-F(S,t)]dt\,.$$

Rearranging the above, we obtain

$$dF(t)=F(S,t)\sigma_s dZ_s^*(t)+F(S,t)H_c(\tau)\sigma_c dZ_c^*(t)+F(S,t)H_r(\tau)\sigma_r dZ_r^*(t)$$

$$+F(SY,t)-F(S,t)-\lambda E[F(SY,t)-F(S,t)]dt$$

Dividing the expression above by F yields

$$\frac{dF(t)}{F(S,t)}=\sigma dZ_s^*(t)+H_c(\tau)\sigma_c dZ_c^*(t)+H_r(\tau)\sigma_r dZ_r^*(t)$$

$$+\frac{F(SY,t)-F(S,t)}{F(S,t)}-\frac{\lambda E[F(SY,t)-F(S,t)]dt}{F(S,t)},$$

$$\frac{dF(t)}{F(S,t)}=\sigma dZ_s^*(t)+H_c(\tau)\sigma_c dZ_c^*(t)+H_r(\tau)\sigma_r dZ_r^*(t)+J_F^* dq^*-\lambda\bar{J}_F^* dt\,. \qquad (390)$$

If we let $\sigma_F dZ_F^*(t)=\sigma dZ_s^*(t)+H_c(\tau)\sigma_c dZ_c^*(t)+H_r(\tau)\sigma_r dZ_r^*(t)$, then expression (390) reduces to

$$\frac{dF(t)}{F(S,t)}=-\lambda\bar{J}_F^* dt+\sigma_F dZ_F^*(t)+J_F^* dq^*\,. \qquad (391)$$

The dynamics of the futures price is similar form to the one- and two-factor models. Thus, we know the expected futures price is of the form

$$E_t^*[F(T_1) \mid n\ jumps] = F(t)\exp[b(n)\tau_1],$$

where $b(n) = -\lambda^* \bar{J}_F^* + \frac{n\gamma^*}{\tau_1}$. The volatility for the log returns, $\ln\left(\frac{F(T_1)}{F(t)}\right)$, is equal to $\nu^2 + n\omega^2$, where expression (141) on page164 of this chapters states

$$\nu^2 = \int_t^{T_1} \sigma_F^2 ds,$$

$$\nu^2 = \int_t^{T_1} \left(\sigma dZ_s^*(t) + H_c(\tau)\sigma_c dZ_c^*(t) + H_r(\tau)\sigma_r dZ_r^*(t)\right)^2 ds,$$

$$\nu^2 = \sigma_s^2(T_1 - t) + \frac{\sigma_c^2}{k_c^2}\left[(T_1 - t) + \frac{1}{2k_c}\left(e^{-2k_c(T-T_1)} - e^{-2k_c(T-t)}\right) - \frac{2}{k_c}\left(e^{-k_c(T-T_1)} - e^{-k_c(T-t)}\right)\right]$$

$$+ \frac{\sigma_r^2}{k_r^2}\left[(T_1 - t) + \frac{1}{2k_r}\left(e^{-2k_r(T-T_1)} - e^{-2k_r(T-t)}\right) - \frac{2}{k_r}\left(e^{-k_r(T-T_1)} - e^{-k_r(T-t)}\right)\right]$$

$$- \frac{2\rho_{sc}\sigma_s\sigma_c}{k_c}\left[(T_1 - t) - \frac{1}{k_c}\left(e^{-k_c(T-T_1)} - e^{-k_c(T-t)}\right)\right]$$

$$+ \frac{2\rho_{sr}\sigma_s\sigma_r}{k_r}\left[(T_1 - t) - \frac{1}{k_r}\left(e^{-k_r(T-T_1)} - e^{-k_r(T-t)}\right)\right]$$

$$- \frac{\sigma_c\sigma_r\rho_{rc}}{k_c k_r}\left[(T_1 - t) - \frac{\left(e^{-k_r(T-T_1)} - e^{-k_r(T-t)}\right)}{k_r} - \frac{\left(e^{-k_c(T-T_1)} - e^{-k_c(T-t)}\right)}{k_c}\right.$$

$$\left. + \frac{1}{k_r + k_c}\left(e^{-k_r(T-T_1)}e^{-k_c(T-T_1)} - e^{-k_r(T-t)}e^{-k_c(T-t)}\right)\right]. \qquad (392)$$

In the one- and two-factor models, once we found the expected futures price, we could begin to price the option. In our present case, we may not follow the arguments presented in the one- and two-factor models. Currently, the model allows for stochastic movements in interest rates. From section 4.4, we know that no closed form solution exists for the Black-Scholes formula when interest rates are random. The reason is due to the correlation between the interest rate and the futures price. The expectation cannot be evaluated. Since Bates' model relies on the Black-Scholes solution, no closed form solution exists for this model either.

In contrast to the observation noted in the three factor model, Hilliard and Reis report a pricing formula for the three-factor jump-diffusion model. They state the option pricing formula is as follows

$$C(FY,t) = \sum_{n=0}^{\infty} \text{prob}^*(\text{n jumps}) E_t^* \left[e^{\int_t^{\tau_1} r(v)dv} \max[F(\tau_1,\tau) - X, 0] \mid \text{n jumps} \right],$$

$$C(FY,t) = \sum_{n=0}^{\infty} \frac{e^{-\lambda\tau_1}(\lambda\tau_1)^n}{n!} E_t^* \left[e^{\int_t^{\tau_1} r(v)dv} F(\tau_1,\tau) e^{b(n)\tau_1} N(d_1) - XP(t,T_1) N(d_2) \right], \tag{393}$$

where $b(n) = \left(-\lambda^* \bar{J}_F^* + \frac{n\gamma^*}{\tau_1} \right)$, (394)

$$d_1 = \frac{\ln\left(\frac{E_t^* \left[e^{\int_t^{\tau_1} r(v)dv} F(\tau_1,\tau) \right]}{XP(t,T_1)} \right) + \left(b(n)\tau_1 + \frac{1}{2}(v^2 + n\omega^2) \right)}{\sqrt{v^2 + n\omega^2}}, \tag{395}$$

and

$$d_2 = d_1 - \sqrt{v^2 + n\omega^2}. \tag{396}$$

Note, the authors essentially value the option by allowing only the futures price to be influenced by the stochastic interest rate. To arrive to a closed form solution, the authors claim they need to evaluate $E_t^*\left[e^{\int_t^{T_1} r(v)dv} F(\tau_1,\tau)\right]$. They state that since futures prices are not influenced by jumps, the solution of $E_t^*\left[e^{\int_t^{T_1} r(v)dv} F(\tau_1,\tau)\right]$ is also not affected by jumps in the futures price. Thus, this expectation can be taken with respect to the interest rate process and the futures price process without the jumps. That is, instead of using expression (391), the authors use the following diffusion for futures

$$\frac{dF(t)}{F(t)} = \sigma_F(\tau)dZ_F^*(t). \tag{397}$$

Now, finding a solution for $E_t^*\left[e^{\int_t^{T_1} r(v)dv} F(\tau_1,\tau)\right]$ let $G(t) = \ln F(\tau_1,\tau) - \int_t^{T_1} r(v)dv$. With the transformation, Ito's lemma and the risk-adjusted diffusion for the futures price, the diffusion for $G(t)$ is

$$\begin{aligned} dG(t) &= G_t dt + G_F dF + \frac{1}{2}G_{FF}dF^2 - d\left[\int_t^{T_1} r(v)dv\right], \\ &= \frac{1}{F(t)}\sigma_F(\tau)F(t)dZ_F^*(t) - \frac{1}{2}\left(\frac{1}{[F(t)]^2}\right)[F(t)]^2\sigma_F^2(\tau)dt - d\left[\int_t^{T_1} r(v)dv\right], \\ &= -\frac{1}{2}\sigma_F^2(\tau)dt + \sigma_F(\tau)dZ_F^*(t) - d\left[\int_t^{T_1} r(v)dv\right]. \end{aligned} \tag{398}$$

Now integrating over (398) we get

$$\int_t^{T_1} dG(v) = -\frac{1}{2}\int_t^{T_1}\sigma_F^2(v)dv + \int_t^{T_1}\sigma_F(v)dZ_F^*(v) - \int_t^{T_1} d\left[\int_t^{T_1} r(v)dv\right],$$

$$G(T_1) - G(t) = -\frac{1}{2}\int_t^{T_1}\sigma_F^2(v)dv + \int_t^{T_1}\sigma_F(v)dZ_F^*(v) - \int_t^{T_1}r(v)dv,$$

$$G(T_1) = G(t) - \frac{1}{2}\int_t^{T_1}\sigma_F^2(v)dv + \int_t^{T_1}\sigma_F(v)dZ_F^*(v) - \int_t^{T_1}r(v)dv. \qquad (399)$$

Hilliard and Reis claim that the distribution of $G(T_1)$ is normal, which implies that the expected value of the futures price is given as

$$E_t^*[F(T_1)] = E_t^*\left[e^{G(T_1)}\right],$$

$$= e^{E_t^*[G(T_1)] + \frac{1}{2}V_t^*[G(T_1)]}. \qquad (400)$$

The expectation in expression (400) is not the same as the expectation

$$E_t^*\left[e^{\int_t^{T_1} r(v)dv} F(\tau_1, \tau)\right].$$

The expectation above is for the product of two dependent random variables. These variables are the interest rate and futures price and we know they are dependent since the futures price is a function of the interest rate. From section 4.4, we know this expectation cannot be evaluated. Now, looking at the expectation in expression (400), we see that this is the expectation of a single log-normal random variable. Hilliard and Reis are evaluating the expected value for the futures price independent of the interest rate. Later in their analysis they will claim that the right hand side of expression (400) is equal to

$$E_t^*\left[e^{\int_t^{T_1} r(v)dv} F(\tau_1, \tau)\right].$$

This outcome will be shown in expression (412) and it is incorrect.

Continuing with Hilliard and Reis, the expected value of $G(T_1)$ is

$$E_t^*[G(T_1)] = \ln F(t) - \frac{1}{2}\int_t^{T_1}\sigma_F^2(v)dv - \int_t^{T_1}E_t^*[r(v)]dv. \tag{401}$$

Now recall from chapter 3 section 3.32 on page 80 we show that

$$\int_t^{T_1}E_t^*[r(v)]dv = \int_t^{T_1}f(t,v)dv - (H_r(\tau_1)-\tau)\frac{\sigma_r^2}{2k_r^2} - \frac{\sigma_r^2H_r^2(\tau_1)}{4k_r}.$$

Substituting this result into (401) yields

$$E_t^*[G(T_1)] = \ln F(t) - \frac{1}{2}\int_t^{T_1}\sigma_F^2(v)dv - \int_t^{T_1}f(t,v)dv + (H_r(\tau_1)-\tau_1)\frac{\sigma_r^2}{2k_r^2} + \frac{\sigma_r^2H_r^2(\tau_1)}{4k_r}. \tag{402}$$

The variance of $G(T_1)$ is

$$V_t^*[G(T_1)] = V_t^*\left[\int_t^{T_1}\sigma_F(v)dZ_F^*(v)\right] - V_t^*\left[\int_t^{T_1}r(v)dv\right] - 2Cov_t^*\left[\int_t^{T_1}\sigma_F(v)dZ_F^*(v), \int_t^{T_1}r(v)dv\right]. \tag{403}$$

From equation (170) in section 3 of chapter three Hilliard and Reis have

$$V_t^*\left[\int_t^{T_1}r(v)dv\right] = -(H_r(\tau_1)-\tau_1)\frac{\sigma_r^2}{k_r^2} - \frac{\sigma_r^2H_r^2(\tau_1)}{2k_r}. \tag{404}$$

The variance of the log returns for the futures price is found by invoking Ito's isometry.

$$V_t^*\left[\int_t^{T_1}\sigma_F(v)dZ_F^*(v)\right] = (1)\int_t^{T_1}\sigma_F^2(v)dv. \tag{405}$$

The last thing the authors need to work on is the covariance term. Recall

$$\sigma_F(\tau)dZ_F^* = \sigma_s dZ_s^*(t) - H_c(\tau)\sigma_c dZ_c^*(t) + H_r(\tau)\sigma_r dZ_r^*(t).$$

Therefore

$$\sigma_F^2 = \sigma_s^2 + \sigma_c^2H_c^2(\tau) + \sigma_r^2H_r^2(\tau) - 2\sigma_c\sigma_s\rho_{sc}H_c(\tau) + 2\sigma_r\sigma_s\rho_{sr}H_r(\tau)$$

$$-2\sigma_r\sigma_c\rho_{cr}H_r(\tau)H_c(\tau).$$

Now the covariance term in (403) can be written as

$$Cov_t^*\left[\int_t^{T_1}\sigma_F(v)dZ_F^*(v), \int_t^{T_1}r(v)dv\right] = Cov_t^*\left[\int_t^{T_1}\sigma_s dZ_s^*(v), \int_t^{T_1}r(v)dv\right]$$

$$-Cov_t^*\left[\int_t^{T_1}H_c(\tau)\sigma_c dZ_c^*(v), \int_t^{T_1}r(v)dv\right]$$

$$+Cov_t^*\left[\int_t^{T_1}H_r(\tau)\sigma_r dZ_r^*(v), \int_t^{T_1}r(v)dv\right]. \quad (406)$$

We need expressions for the individual terms on the right-hand side of expression (406). Recall from equation (171) in chapter 3

$$Cov_t^*\left[\sigma_s\int_t^{T_1}dZ_s^*(v), \int_t^{T_1}r(v)dv\right] = \frac{\sigma_s\sigma_r\rho_{rs}}{k_r}(\tau - H_r(\tau)).$$

Therefore, the above is expressed as

$$Cov_t^*\left[\sigma_s\int_t^{T_1}dZ_s^*(v), \int_t^{T_1}r(v)dv\right] = \frac{\sigma_s\sigma_r\rho_{rs}}{k_r}(\tau_1 - H_r(\tau_1)). \quad (407)$$

Now working on the second term on the right-hand side of expression (406)

$$Cov_t^*\left[\int_t^{T_1}H_c(\tau)\sigma_c dZ_c^*(v), \int_t^{T_1}r(v)dv\right] = Cov_t^*\left[\int_t^{T_1}H_c(\tau)\sigma_c dZ_c^*(v), \int_t^{T_1}\sigma_r e^{-k_r v}\int_t^{v}e^{k_r s}dZ_r^*(s)dv\right].$$

Switching the order of integration yields

$$= Cov_t^*\left[\int_t^{T_1}H_c(\tau)\sigma_c dZ_c^*(v), \int_t^{T_1}e^{k_r s}\int_s^{v}\sigma_r e^{-k_r v}dv dZ_r^*(s)\right].$$

Integrating the second integral in the second term above yields

$$= Cov_t^*\left[\int_t^{T_1}H_c(\tau)\sigma_c dZ_c^*(v), \int_t^{T_1}e^{k_r s}\frac{-\sigma_r}{k_r}\left[e^{-k_r T_1} - e^{-k_r s}\right]dZ_r^*(s)\right].$$

Substituting into the above for $H_c(\tau_1)$ and factoring gives

$$= \frac{\sigma_r \sigma_c}{k_r k_c} Cov_t^* \left[\int_t^{T_1} (1 - e^{-k_c \tau}) dZ_c^*(v), \int_t^{T_1} \left[1 - e^{-k_r (T_1 - s)}\right] dZ_r^*(s) \right],$$

$$= \frac{\sigma_r \sigma_c}{k_r k_c} Cov_t^* \left[\int_t^{T_1} (1 - e^{-k_c (T-s)}) dZ_c^*(v), \int_t^{T_1} \left[1 - e^{-k_r (T_1 - s)}\right] dZ_r^*(s) \right].$$

Using Ito's isometry yields

$$= \frac{\sigma_r \sigma_c \rho_{rc}}{k_r k_c} \left[\int_t^{T_1} \left(1 - e^{-k_c (T-s)}\right)\left(1 - e^{-k_r (T_1 - s)}\right) ds \right],$$

$$= \frac{\sigma_r \sigma_c \rho_{rc}}{k_r k_c} \left[\int_t^{T_1} \left(1 - e^{-k_c (T-s)} - e^{-k_r (T_1 - s)} + e^{-k_c (T-s) - k_r (T_1 - s)}\right) ds \right],$$

$$= \frac{\sigma_r \sigma_c \rho_{rc}}{k_r k_c} \left[(T_1 - t) - \left(\frac{1}{k_c}\right)\left(e^{-k_c (T-T_1)} - e^{-k_c (T-t)}\right) - \left(\frac{1}{k_r}\right) e^{-k_r (T_1 - T_1)} - e^{-k_r (T_1 - t)} \right.$$

$$\left. + \int_t^{T_1} e^{(k_c + k_r) s - k_c T - k_r T_1} ds \right],$$

$$= \frac{\sigma_r \sigma_c \rho_{rc}}{k_r k_c} \left[\tau_1 - \left(\frac{1}{k_r}\right)\left(1 - e^{-k_r (T_1 - t)}\right) - \left(\frac{1}{k_c}\right)\left(e^{-k_c (T-T_1)} - e^{-k_c (T-t)}\right) + \int_t^{T_1} e^{(k_c + k_r) s - k_c T - k_r T_1} ds \right],$$

$$= \frac{\sigma_r \sigma_c \rho_{rc}}{k_r k_c} \left[\tau_1 - H_r(\tau_1) - \left(\frac{e^{-k_c (T-T_1)} - e^{-k_c \tau}}{k_c}\right) + \frac{e^{(k_c + k_r) T_1 - k_c T - k_r T_1} - e^{(k_c + k_r) t - k_c T - k_r T_1}}{(k_c + k_r)} \right],$$

$$= \frac{\sigma_r \sigma_c \rho_{rc}}{k_r k_c} \left[\tau_1 - H_r(\tau_1) - \left(\frac{e^{-k_c (T-T_1)} - e^{-k_c \tau}}{k_c}\right) + \frac{e^{-k_c (T-T_1)} - e^{(k_c + k_r) t - k_c T - k_r T_1}}{(k_c + k_r)} \right]. \quad (408)$$

Working on the third and final term on the left-hand side of equation (406) yields

$$Cov_t^* \left[\int_t^{T_1} H_r(\tau) \sigma_r dZ_r^*(v), \int_t^{T_1} r(v) dv \right] = Cov_t^* \left[\int_t^{T_1} H_r(\tau) \sigma_r dZ_r^*(v), \int_t^{T_1} \sigma_r e^{-k_r v} \int_t^{v} e^{k_r s} dZ_r^*(s) dv \right].$$

Switching the order of integration of the iterated integral in the expression above yields

$$= Cov_t^* \left[\int_t^{T_1} H_r(\tau) \sigma_r dZ_r^*(v), \int_t^{T_1} e^{k_r s} \int_s^{v} \sigma_r e^{-k_r v} dv dZ_r^*(s) \right].$$

Integrating the second integral in the above expression we obtain

$$= Cov_t^*\left[\int_t^{T_1} H_r(\tau)\sigma_r dZ_r^*(v), \int_t^{T_1} e^{k_r s}\frac{-\sigma_r}{k_r}\left[e^{-k_r T_1} - e^{-k_r s}\right]dZ_r^*(s)\right].$$

$$= Cov_t^*\left[\int_t^{T_1}\frac{1-e^{-k_r(T-s)}}{k_r}\sigma_r dZ_r^*(v), \int_t^{T_1}\frac{\sigma_r}{k_r}\left[1-e^{-k_r(T_1-s)}\right]dZ_r^*(s)\right],$$

$$= \frac{\sigma_r^2}{k_r^2}Cov_t^*\left[\int_t^{T_1} 1-e^{-k_r(T-s)}dZ_r^*(v), \int_t^{T_1}\left[1-e^{-k_r(T_1-s)}\right]dZ_r^*(s)\right].$$

Using Ito's isometry we get

$$= \frac{\sigma_r^2}{k_r^2}\left[\int_t^{T_1}\left(1-e^{-k_r(T-s)}\right)\left(1-e^{-k_r(T_1-s)}\right)ds\right],$$

$$= \frac{\sigma_r^2}{k_r^2}\left[\int_t^{T_1}\left(1-e^{-k_r(T_1-s)} - e^{-k_r(T-s)} + e^{-k_r(T-s)}e^{-k_r(T_1-s)}\right)ds\right],$$

$$= \frac{\sigma_r^2}{k_r^2}\left[(T_1-t) - \frac{1}{k_r}\left(e^{-k_r(T_1-T_1)} - e^{-k_r(T_1-t)}\right) - \frac{1}{k_r}\left(e^{-k_r(T-T_1)} - e^{-k_r(T-t)}\right) + \int_t^{T_1} e^{2k_r s - k_r T_1 - k_r T}ds\right],$$

$$= \frac{\sigma_r^2}{k_r^2}\left[\tau_1 - H_r(\tau_1) - \frac{\left(e^{-k_r(T-T_1)} - e^{-k_r\tau}\right)}{k_r} + \frac{1}{2k_r}\left[e^{2k_r T_1 - k_r T_1 - k_r T} - e^{2k_r t - k_r T_1 - k_r T}\right]\right],$$

$$= \frac{\sigma_r^2}{k_r^2}\left[\tau_1 - H_r(\tau_1) - \frac{\left(e^{-k_r(T-T_1)} - e^{-k_r\tau}\right)}{k_r} + \frac{\left[e^{-(k_r T - T_1)} - e^{2k_r t - k_r(T+T_1)}\right]}{2k_r}\right]. \qquad (409)$$

Substitute expressions (409)-(407) into (406) to obtain

$$Cov_t^*\left[\int_t^{T_1}\sigma_F(v)dZ_F^*(v), \int_t^{T_1} r(v)dv\right] = \frac{\sigma_s\sigma_r\rho_{rs}}{k_r}\left(\tau_1 - H_r(\tau_1)\right)$$

$$-\frac{\sigma_r\sigma_c\rho_{rc}}{k_r k_c}\left[\tau_1 - H_r(\tau_1) - \left(\frac{e^{-k_c(T-T_1)} - e^{-k_c\tau}}{k_c}\right) + \frac{e^{-k_c(T-T_1)} - e^{(k_c+k_r)t - k_c T - k_r T_1}}{(k_c+k_r)}\right]$$

$$+\frac{\sigma_r^2}{k_r^2}\left[\tau_1 - H_r(\tau_1) - \frac{\left(e^{-k_r(T-T_1)} - e^{-k_r\tau}\right)}{k_r} + \frac{\left[e^{-(k_r T - T_1)} - e^{2k_r t - k_r(T+T_1)}\right]}{2k_r}\right].$$

(410)

Now substitute expressions (404), (405) and (410) into (403) to obtain the variance

$$V_t^*[G(T_1)] = \int_t^{T_1}\sigma_F^2(v)dv - (H_r(\tau_1)-\tau_1)\frac{\sigma_r^2}{k_r^2} - \frac{\sigma_r^2 H_r^2(\tau_1)}{2k_r} - 2\left[\frac{\sigma_s\sigma_r\rho_{rs}}{k_r}(\tau_1 - H_r(\tau_1))\right.$$

$$-\frac{\sigma_r\sigma_c\rho_{rc}}{k_r k_c}\left[\tau_1 - H_r(\tau_1) - \left(\frac{e^{-k_c(T-T_1)} - e^{-k_c\tau}}{k_c}\right) + \frac{e^{-k_c(T-T_1)} - e^{(k_c+k_r)t-k_cT-k_rT_1}}{(k_c+k_r)}\right]$$

$$\left.+\frac{\sigma_r^2}{k_r^2}\left[\tau_1 - H_r(\tau_1) - \frac{\left(e^{-k_r(T-T_1)} - e^{-k_r\tau}\right)}{k_r} + \frac{\left[e^{-(k_rT-T_1)} - e^{2k_rt-k_r(T+T_1)}\right]}{2k_r}\right]\right].$$

(411)

Lastly, substitute (402) and (411) into (400) to obtain

$$E_t^*\left[e^{\int_t^{T_1} r(v)dv}F(\tau_1,\tau)\right] = F(t)\exp\left\{-\frac{1}{2}\int_t^{T_1}\sigma_F^2(v)dv - \int_t^{T_1} f(t,v)dv + (H_r(\tau_1)-\tau_1)\frac{\sigma_r^2}{2k_r^2} + \frac{\sigma_r^2 H_r^2(\tau_1)}{4k_r}\right.$$

$$+\frac{1}{2}\left[\int_t^{T_1}\sigma_F^2(v)dv - (H_r(\tau_1)-\tau_1)\frac{\sigma_r^2}{k_r^2} - \frac{\sigma_r^2 H_r^2(\tau_1)}{2k_r} - 2\left[\frac{\sigma_s\sigma_r\rho_{rs}}{k_r}(\tau_1 - H_r(\tau_1))\right.\right.$$

$$-\frac{\sigma_r\sigma_c\rho_{rc}}{k_r k_c}\left[\tau_1 - H_r(\tau_1) - \left(\frac{e^{-k_c(T-T_1)} - e^{-k_c\tau}}{k_c}\right) + \frac{e^{-k_c(T-T_1)} - e^{(k_c+k_r)t-k_cT-k_rT_1}}{(k_c+k_r)}\right]$$

$$\left.\left.\left.+\frac{\sigma_r^2}{k_r^2}\left[\tau_1 - H_r(\tau_1) - \frac{\left(e^{-k_r(T-T_1)} - e^{-k_r\tau}\right)}{k_r} + \frac{\left[e^{-(k_rT-T_1)} - e^{2k_rt-k_r(T+T_1)}\right]}{2k_r}\right]\right]\right]\right\}.$$

The above expression reduces to

$$E_t^*\left[e^{\int_t^{T_1} r(v)dv}F(\tau_1,\tau)\right] = F(t)\exp\left\{-\int_t^{T_1} f(t,v)dv - \frac{\sigma_s\sigma_r\rho_{rs}}{k_r}(\tau_1 - H_r(\tau_1))\right.$$

$$+\frac{\sigma_r\sigma_c\rho_{rc}}{k_r k_c}\left[\tau_1 - H_r(\tau_1) - \left(\frac{e^{-k_c(T-T_1)} - e^{-k_c\tau}}{k_c}\right) + \frac{e^{-k_c(T-T_1)} - e^{(k_c+k_r)t-k_cT-k_rT_1}}{(k_c+k_r)}\right]$$

$$-\frac{\sigma_r^2}{k_r^2}\left[\tau_1 - H_r(\tau_1) - \frac{\left(e^{-k_r(T-T_1)} - e^{-k_r\tau}\right)}{k_r} + \frac{\left[e^{-(k_rT-T_1)} - e^{2k_rt-k_r(T+T_1)}\right]}{2k_r}\right]\Bigg\},$$

which further reduces to

$$E_t^*\left[e^{\int_t^{T_1} r(v)dv} F(\tau_1,\tau)\right] = F(t)P(t,T)\exp\Bigg\{-\frac{\sigma_s\sigma_r\rho_{rs}}{k_r}(\tau_1 - H_r(\tau_1))$$

$$+\frac{\sigma_r\sigma_c\rho_{rc}}{k_r k_c}\left[\tau_1 - H_r(\tau_1) - \left(\frac{e^{-k_c(T-T_1)} - e^{-k_c\tau}}{k_c}\right) + \frac{e^{-k_c(T-T_1)} - e^{(k_c+k_r)t-k_cT-k_rT_1}}{(k_c+k_r)}\right]$$

$$-\frac{\sigma_r^2}{k_r^2}\left[\tau_1 - H_r(\tau_1) - \frac{\left(e^{-k_r(T-T_1)} - e^{-k_r\tau}\right)}{k_r} + \frac{\left[e^{-(k_rT-T_1)} - e^{2k_rt-k_r(T+T_1)}\right]}{2k_r}\right]\Bigg\}.$$

Simplifying the expression

$$E_t^*\left[e^{\int_t^{T_1} r(v)dv} F(\tau_1,\tau)\right] = F(t)P(t,T)Z(t,T), \tag{412}$$

where $Z(t,T) = \exp\Bigg\{-\frac{\sigma_s\sigma_r\rho_{rs}}{k_r}(\tau_1 - H_r(\tau_1))$

$$+\frac{\sigma_r\sigma_c\rho_{rc}}{k_r k_c}\left[\tau_1 - H_r(\tau_1) - \left(\frac{e^{-k_c(T-T_1)} - e^{-k_c\tau}}{k_c}\right) + \frac{e^{-k_c(T-T_1)} - e^{(k_c+k_r)t-k_cT-k_rT_1}}{(k_c+k_r)}\right]$$

$$-\frac{\sigma_r^2}{k_r^2}\left[\tau_1 - H_r(\tau_1) - \frac{\left(e^{-k_r(T-T_1)} - e^{-k_r\tau}\right)}{k_r} + \frac{\left[e^{-(k_rT-T_1)} - e^{2k_rt-k_r(T+T_1)}\right]}{2k_r}\right]\Bigg\}.$$

Substituting expression (412) into the solution for the option price yields

$$C(FY,t) = P(t,T)\sum_{n=0}^{\infty}\frac{e^{-\lambda\tau_1}(\lambda\tau_1)^n}{n!}\left[F(t)Z(t,T)e^{b(n)\tau_1}N(d_1) - XN(d_2)\right], \tag{413}$$

where $b(n) = \left(-\lambda^* \bar{J}_F^* + \frac{n\gamma^*}{\tau_1} \right)$ (414)

$$d_1 = \frac{\ln\left(\frac{F(t)Z(t,T)}{X} \right) + \left(b(n)\tau_1 + \frac{1}{2}(v^2 + n\omega^2) \right)}{\sqrt{v^2 + n\omega^2}}, \quad (415)$$

and

$$d_2 = d_1 - \sqrt{v^2 + n\omega^2}\,. \quad (416)$$

The solution presented in expression (413) is the three-factor jump-diffusion option formula given by Hilliard and Reis. The solution is based on the assumption that the futures price is influenced by the stochastic interest rate, while the option is not. The model is a contradiction of terms and the result is tenuous at best.

Endnotes

[1] The cautious reader should note while the spot price follows a geometric Brownian motion, the option price need not be Brownian motion. The option is a function of the spot price, and as such it too will be a stochastic process. But, this only means it will have a drift and diffusion term. We can characterize its return process by dividing the stochastic differential by its level. This leads to expression (2).

[2] The reader is directed to Appendix C for a review of equivalent martingale measures.

[3] This result is shown in chapter three on pages 26 to 28.

[4] The partial differential equation for the Black-Scholes analysis is

$$\frac{1}{2}F_{ss}\sigma^2[S(t)]^2 + F_s rS(t) + F_t - F(S(t),t)r = 0.$$

For the current model that considers a mean reverting spot price, the above partial differential equation would be

$$\frac{1}{2}F_{ss}\sigma^2[S(t)]^2 + F_s\left(k\left(\alpha^* - \ln S\right)\right)S(t) + F_t - F(S(t),t)r = 0.$$

[5] This result is consistent with Samuelson's (1965) assertion that changes in the futures price is equal to zero.

[6] This result is consistent with Samuelson's (1965) hypothesis that a commodity's futures price volatility is increasing as maturity decreases.

[7] It can be demonstrated that when the underlying spot price follows a geometric Brownian motion process, the futures price will too follow a geometric Brownian motion. Under this circumstance the volatility of the futures price will be the same as the spot price.

[8] The result of the expected growth rate in a futures price in a risk-neutral world is zero, is a very general one. It is true for all futures prices. It applies in the world where interest rates are stochastic as well as the world where they are constant. We see from expression (62), that since the expected growth rate in futures is zero

$$F(t) = E_t^*[F(T)],$$

where $F(T)$ is the futures price at the maturity of the contract, $F(t)$ is the futures price a time t, and E_t^* is the risk neutral expectation operator. At expiration, we know that $F(T) = S(T)$, where $S(T)$ is the spot price at time T. It follows that

$$F(t) = E_t^*[F(T)] = E_t^*[S(T)].$$

From equation (23), we know that

$$E_t^*[S(T)] = S(t)e^{r\tau}.$$

Substituting this into the above yields

$$F(t) = S(t)e^{r\tau}.$$

The above means that for all assets the futures price equals the expected future spot price in a risk neutral world.

[9] In the financial derivatives literature, economists use the Poisson process to model discontinuous jumps in asset prices. The Poisson process is well suited for modeling this phenomenon. To understand why, we develop a fundamental analysis of the Poisson distribution.

The Poisson distribution is the limiting form of the binomial distribution. The binomial distribution is the probability distribution associated with random variable that has only two possible outcomes: a success or failure. The probability of witnessing a success is measure by p and the likelihood of a failure is measure by $(1-p)$. Given a finite number of repeated trials, k, the probability of witnessing a certain number of successes, x, is

$$\Pr(x;k,p) = C\binom{k}{x}p^x(1-p)^{k-x}. \tag{1}$$

We shall investigate the limiting form of the binomial when $k \to \infty$, $p \to 0$, while kp remains constant. The reason for analyzing the limiting case, is due to the phenomenon we wish to model. We want a model that allows for discrete random jumps in asset prices. Over a small interval of time, we could witness several jumps, although, the likelihood of witnessing more than one jump period is negligible. The limiting form of the binomial distribution is one way to model this behavior. It allows for an infinite number of occurrences and the probability of witnessing these occurrences is small; all the while keeping the average rate of occurrence constant.

Continuing with the analysis, let $kp = \lambda$. This implies $p = \frac{\lambda}{k}$, and we can write

$$\Pr(x;k,p) = C\binom{k}{x}\left(\frac{\lambda}{k}\right)^x\left(1-\frac{\lambda}{k}\right)^{k-x}. \tag{2}$$

Expression (2) and (1) are equal. Expression two is just another may of writing the probability of the binomial. We can rewrite equation (2) to obtain

$$C\binom{k}{x}\left(\frac{\lambda}{k}\right)^x\left(1-\frac{\lambda}{k}\right)^{k-x} = \frac{k!}{x!(k-x)!}\left(\frac{\lambda}{k}\right)^x\left(1-\frac{\lambda}{k}\right)^{k-x},$$

$$= \frac{\lambda^x}{x!}\frac{k!}{(k-x)!k^x}\left(1-\frac{\lambda}{k}\right)^{k-x}.$$

This expression may be rewritten. First, note that

$$\frac{k!}{(k-x)!} = k(k-1)(k-2)\cdots(k-x+1). \qquad (3)$$

Thus, using equation (3) expression (2) becomes

$$\frac{\lambda^x}{x!}\frac{k!}{(k-x)!k^x}\left(1-\frac{\lambda}{k}\right)^{k-x} = \frac{\lambda^x}{x!}\frac{k(k-1)(k-2)\cdots(k-x+1)}{k^x}\left(1-\frac{\lambda}{k}\right)^{k-x}. \qquad (4)$$

Recognizing that

$$\left(1-\frac{\lambda}{k}\right)^{k-x} = \left(1-\frac{\lambda}{k}\right)^{k}\frac{1}{\left(1-\frac{\lambda}{k}\right)^{x}},$$

expression (4) can be expressed as

$$\frac{\lambda^x}{x!}\frac{k!}{(k-x)!k^x}\left(1-\frac{\lambda}{k}\right)^{k-x} = \frac{\lambda^x}{x!}\frac{k(k-1)(k-2)\cdots(k-x+1)}{k^x}\left(1-\frac{\lambda}{k}\right)^{k}\frac{1}{\left(1-\frac{\lambda}{k}\right)^{x}}. \qquad (5)$$

If we let $k \to \infty$, we obtain the pdf for the Poisson distribution. Taking the limit of (5) as $k \to \infty$ yields the following. The first term on the right-hand side of (5) is free of k and therefore remains unchanged. The second term, however, does change. Consider the ratio

$$\frac{k(k-1)(k-2)\cdots(k-x+1)}{k^x}.$$

This is also

$$\frac{k(k-1)(k-2)\cdots(k-x+1)}{k^x} = 1\left(1-\frac{1}{k}\right)\left(1-\frac{2}{k}\right)\cdots\left(1-\frac{(x-1)}{k}\right). \qquad (6)$$

Now, recall the property, $\lim(a_n, b_n) = \lim(a_n)\lim(b_n)$ and $\lim_{n\to\infty}\left(1-\frac{c}{n}\right) = 1$. For any fixed c

$$\lim_{n\to\infty} 1\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\cdots\left(1-\frac{(c-1)}{n}\right)=1.$$

Therefore, the limit of expression (6) as $k \to \infty$ is

$$\lim_{k\to\infty} 1\left(1-\frac{1}{k}\right)\left(1-\frac{2}{k}\right)\cdots\left(1-\frac{(x-1)}{k}\right)=1. \qquad (7)$$

We now to turn our focus to the third term on the right-hand side of (5). We want to know

$$\lim_{k\to\infty}\left(1-\frac{\lambda}{k}\right)^k = ?$$

Consider

$$X_n = \left(1+\frac{1}{n}\right)^n. \qquad (8)$$

Take the natural log of X_n to obtain

$$\ln(X_n) = n\ln\left(1+\frac{1}{n}\right).$$

Rearranging yields

$$= \frac{\ln\left(1+\frac{1}{n}\right)}{\frac{1}{n}}.$$

Now using l'hopital's rule, we get

$$= \frac{\frac{\partial \ln\left(1+\frac{1}{n}\right)}{\partial n}}{\frac{\partial \frac{1}{n}}{\partial n}},$$

$$=\frac{\frac{1}{1+\frac{1}{n}}\left(-\frac{1}{n^2}\right)}{-\frac{1}{n^2}},$$

$$=\frac{1}{1+\frac{1}{n}},$$

$$=\frac{n}{n+1}.$$

As $n \to \infty$, the above becomes

$$=\frac{n}{n+1} \to 1.$$

Therefore,

$$\lim_{n\to\infty} X_n = \lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n = e^1. \tag{9}$$

Now consider,

$$Y_n = \left(1-\frac{1}{n}\right)^n$$

Take the natural log of X_n to obtain

$$\ln(Y_n) = n\ln\left(1-\frac{1}{n}\right).$$

Rearranging yields

$$=\frac{\ln\left(1-\frac{1}{n}\right)}{\frac{1}{n}}.$$

Using l'hopital's rule, we get

$$= \frac{\dfrac{\partial \ln\left(1-\dfrac{1}{n}\right)}{\partial n}}{\dfrac{\partial \dfrac{1}{n}}{\partial n}},$$

$$= \frac{\dfrac{1}{1-\dfrac{1}{n}}\left(\dfrac{1}{n^2}\right)}{-\dfrac{1}{n^2}},$$

$$= -\frac{1}{1-\dfrac{1}{n}},$$

$$= -\frac{n}{n-1}.$$

As $n \to \infty$, the above becomes

$$= -\frac{n}{n-1} \to -1.$$

Therefore,

$$\lim_{n\to\infty} Y_n = \lim_{n\to\infty}\left(1-\frac{1}{n}\right)^n = e^{-1}. \qquad (10)$$

Thus, it follows that $\left(1-\frac{1}{n}\right)^n \to e^{-1}$. For our analysis, we want to know

$$\lim_{k\to\infty}\left(1-\frac{\lambda}{k}\right)^k.$$

Let $n' = \frac{k}{\lambda}$, and rewrite the above as

$$\lim_{k\to\infty}\left(1-\frac{1}{n'}\right)^{n'\lambda}.$$

We can use the result of equation (10) to write

$$\lim_{k\to\infty}\left(1-\frac{1}{n'}\right)^{n'\lambda} = e^{-\lambda} \qquad (11)$$

Equation (11) is the limit for the third term in expression (5). We are now left with finding the limit for the last term in expression (5). This is

$$\lim_{k\to\infty}\left(1-\frac{\lambda}{k}\right)^{-x} = 1. \qquad (12)$$

This follows immediately from $\lim(a_n, b_n) = \lim(a_n)\lim(b_n)$ and $\lim_{n\to\infty}\left(1-\frac{c}{n}\right) = 1$.

Substituting expressions (7), (11) and (12) into (5) gives the limiting function of the binomial pdf. It is

$$\Pr(x;\lambda) = \frac{\lambda^x}{x!}(1)e^{-\lambda}(1),$$

$$= \frac{e^{-\lambda}\lambda^x}{x!}. \qquad (13)$$

Equation (13) is the probability density function of the Poisson distribution. The mean and the variance of the Poisson distribution are given by

$$\mu = \lambda \qquad (14)$$

and

$$\sigma^2 = \lambda. \qquad (15)$$

We can show that the mean and variance for the Poisson is in fact λ. Working on the mean we take the expected value of X, which follows a Poisson distribution. That is,

$$E[X] = \sum_{x=0}^{\infty} xp(x) = \sum_{x=0}^{\infty} x \frac{\lambda^x e^{-\lambda}}{x!}, \qquad (16)$$

where $p(x)$ is the probability of realizing a value of X.. By definition, $\sum_{x=0}^{\infty} p(x) = 1$. Note that the first term in the sum of expression (16) will equal zero when $x = 0$ and hence

$$E[X] = \sum_{x=1}^{\infty} x \frac{\lambda^x e^{-\lambda}}{x!}. \qquad (17)$$

The expression in equation (17) can be further reduced to

$$E[X] = \sum_{x=1}^{\infty} \frac{\lambda^x e^{-\lambda}}{(x-1)!}. \qquad (18)$$

As it stands, the quantity above is not equal to the sum of a probability function, $p(x)$, over all values of x. We can, however, change it to the proper form by factoring a λ out of the expression and letting $z = x - 1$. Then the limits of summation become $z = 0$, when $x = 1$ and $z = \infty$ when $x = \infty$. Factoring λ out of expression (18) yields

$$E[X] = \lambda \sum_{x=1}^{\infty} \frac{\lambda^{x-1} e^{-\lambda}}{(x-1)!}. \qquad (19)$$

Now, changing the variables in equation (19) we get

$$E[X] = \lambda \sum_{z=0}^{\infty} \frac{\lambda^z e^{-\lambda}}{z!} = \lambda. \qquad (20)$$

Note, the summation in expression (19) is the probability function for a Poisson random variable, Z, and by definition sums to one. Therefore, the expected value of X, $E[X]$, equals the parameter λ, that appears in the expression for the Poisson probability function.

The variance of a Poisson distribution is by definition expressed

$$V[X]=\sum_{x=0}^{\infty}(x-E[X])^2 p(x)=\sum_{x=0}^{\infty}(x-E[X])^2 \frac{\lambda^x e^{-\lambda}}{x!}. \qquad (21)$$

We rewrite equation (21) as

$$V[X]=\sum_{x=0}^{\infty}\left(x^2 p(x)-2xE[X]p(x)+(E[X])^2 p(x)\right),$$

$$V[X]=\sum_{x=0}^{\infty}x^2 p(x)-\sum_{x=0}^{\infty}2xE[X]p(x)+\sum_{x=0}^{\infty}(E[X])^2 p(x),$$

$$V[X]=\sum_{x=0}^{\infty}x^2 p(x)-E[X]\sum_{x=0}^{\infty}2xp(x)+(E[X])^2\sum_{x=0}^{\infty}p(x),$$

$$V[X]=\sum_{x=0}^{\infty}x^2 p(x)-2E[X]E[X]+(E[X])^2(1),$$

$$V[X]=\sum_{x=0}^{\infty}x^2 p(x)-2(E[X])^2+(E[X])^2,$$

$$V[X]=\sum_{x=0}^{\infty}x^2 p(x)-(E[X])^2,$$

$$V[X]=E[X^2]-(E[X])^2. \qquad (22)$$

We know what the expected value of a Poisson random variable equals. The key to finding the variance is evaluating the term $E[X^2]$.

To evaluate $E[X^2]$ let us first consider the expected value of $X(X-1)$. This is

$$E[X(X-1)]=E[X^2-X]=E[X^2]-E[X],$$

which leads to

$$E[X^2]=E[X(X-1)]+E[X]. \qquad (23)$$

Following through, we have

$$E[X(X-1)] = \sum_{x=0}^{\infty} x(x-1)\frac{\lambda^x e^{-\lambda}}{x!}. \tag{24}$$

When x is equal to zero and one the first and second term are also zero, thus we have

$$E[X(X-1)] = \sum_{x=2}^{\infty} x(x-1)\frac{\lambda^x e^{-\lambda}}{x!}. \tag{25}$$

We can factor out the first two terms in expression (25). This leaves us with

$$E[X(X-1)] = \sum_{x=2}^{\infty} (x-1)\frac{\lambda^x e^{-\lambda}}{(x-1)!},$$

$$E[X(X-1)] = \sum_{x=2}^{\infty} \frac{\lambda^x e^{-\lambda}}{(x-2)!}. \tag{26}$$

Now factor two λ 's out of the sum in equation (26) and let $z = x - 2$. This yields

$$E[X(X-1)] = \lambda^2 \sum_{x=2}^{\infty} \frac{\lambda^{x-2} e^{-\lambda}}{(x-2)!},$$

$$E[X(X-1)] = \lambda^2 \sum_{z=2}^{\infty} \frac{\lambda^z e^{-\lambda}}{z!}. \tag{27}$$

The last term in expression (27) is the sum of a Poisson probability distribution and by definition equals one. Therefore, we have

$$E[X(X-1)] = \lambda^2. \tag{28}$$

Substituting the result in equation (28) into expression (23) yields

$$E[X^2] = E[X(X-1)] + E[X],$$

$$E[X^2] = \lambda^2 + E[X],$$

$$E[X^2] = \lambda^2 + \lambda. \tag{29}$$

We may now determine the variance of a Poisson random variable by substituting the result from equation (29) into expression (22). This is

$$V[X] = E[X^2] - (E[X])^2,$$

$$V[X] = \lambda^2 + \lambda - (E[X])^2,$$

$$V[X] = \lambda^2 + \lambda - (\lambda)^2,$$

$$V[X] = \lambda. \tag{30}$$

Thus, the variance for a Poisson random variable is the single parameter, λ, that appears in the expression for the Poisson probability distribution. The results given in expressions (20) and (30) agree with the earlier statement of equations (14) and (15).

Although the Poisson distribution has been derived as a limiting form of the binomial distribution, it has many applications which have no direct connection with the binomial. For example, the Poisson distribution can serve as a model for the number of successes that occur during a given time interval. This is true when (1) the number of successes occurring in non-overlapping time intervals is independent; (2) the probability of a single success occurring in a very short time interval is proportional to the length of the time interval; (3) the probability of more than one success occurring in such a short time interval is negligible. These properties are consistent with economists' assumptions for movements in market prices. Hence, the Poisson distribution seems to be a good candidate for modeling discontinuous jumps in a commodity's spot price over some short interval of time.

[10] Modeling jumps with a Poisson process we know that $f(q=1 \mid q=0)$ is not the same as $f(q=2 \mid q=1) = f(q=2 \mid q=0)$. That is, the probability of witnessing more than one jump over an interval *dt* dramatically decreasing. Some observers may suggest that the probability of witnessing further jumps could possibly be conditioned on the occurrence of a jump. That is, the

probability of witnessing successive jumps over the interval dt is conditional on prior jumps occurring during the interval. In this case, the Poisson distribution would not be a good model for jump behavior in asset pricing.

[11] To equate the options instantaneous drift equal to the risk-free rate of return Merton relies on the CAPM. Harrison and Kreps (1979) and Harrison and Pliska (1981) show under very general conditions, that in the absence of arbitrage opportunities, there exists a risk neutral probability measure. Rendleman and Carabini (1979) give empirical support that no arbitrage opportunities exist in the Treasury bill futures market. They indicate that when brokerage costs, bid-ask spreads and borrowing costs are taken into consideration, no pure arbitrage opportunities can be found.

[12] The characterization of Merton's solution comes from Press (1967). Press derives moments, pdf and cdf for a mixed jump diffusion process. The solution in Merton (1976) is just a stylized result of Press (1967).

[13] Refer to Chapter 3 pages 26 through 28 for a proof of this result.

[14] This result is shown in Press (1967)

[15] An excellent reference for this is found in Neftci (2000). For a more rigorous treatment of Ito's lemma, readers should consult Merton (1990) and Kushner (1995).

Chapter 5

Simulations

In chapters three and four, we developed closed form solutions for commodity futures and options on futures. These solutions greatly simplify the comparative statics and estimation of these contingent claims written on commodity spot contracts. The purpose of this chapter is to examine the pricing behavior of these models. Simulations for the futures prices and option prices are presented below.

4.1 One-factor Futures model

Using equation (56) from chapter 3,

$$F[S(T),T] = \exp\left[e^{-kT}\ln S(0) + \alpha^*\left(1-e^{-kT}\right) + \frac{\sigma^2}{4k}\left[1-e^{-2kT}\right]\right],$$

we calculate theoretical futures prices for various choices of spot price, speed of adjustment, and time to maturity. The results of these simulations are presented in Table 1, and there are a few causal observations worth noting.

In panel A of table 1, we see that when the initial level of the spot price is close to its long-run mean, the difference between the theoretical futures price and the long run mean spot price is small. This is expected. That is, we know the spot price is mean reverting, thus when the spot price reaches its mean, its tendency is to rest there. The futures price is consistent with this expectation.

Second, as the speed of adjustment k increases the difference between the theoretical futures price and the long run mean spot price is decreasing. This is simply illustrates that as the mean reversion becomes stronger there is a greater tendency for the spot price to revert to its mean, and the futures price should as well. In fact, if $k = \infty$ the theoretical futures will equal its risk-adjusted mean.

Additionally, as the maturity increases the difference between the theoretical futures price and the long run mean spot price is decreasing. Intuitively, as the investment horizon for the spot commodity increases, the spot price is allowed greater opportunity to revert to its mean. If this is the case, the futures price should reflect this behavior and indeed the behavior in the theoretical futures prices is consistent with our expectations.

The one-factor futures model shows that futures price is only influenced by an underlying spot price process. As a consequence, the pricing behavior of the futures contract should resemble the spot price's characteristics. Illustrations of the one-factor futures prices behavior is presented in Figures 2, 3, and 4 and we note that the futures price does behave to our expectations.

4.2 Two-factor Futures model

Theoretical futures prices, which are influenced by a stochastic spot price and convenience yield, are calculated using equation (109) in chapter 3. This expression is

$$F\big(S(T),\delta(T),t,T\big) = S(t)\exp\Bigg\{\left[-\alpha + \frac{1}{k}(\sigma_c\lambda - \sigma_c\sigma_s\rho_{cs}) + \frac{1}{2}\left(\frac{1}{k}\right)^2\sigma_c^2\right](T-t)$$

$$-\frac{1}{k}\left(\delta(t) - \alpha + \frac{1}{k}(\sigma_c\lambda - \sigma_c\sigma_s\rho_{cs}) + \left(\frac{1}{k}\right)^2\sigma_c^2\right)(1-\theta)$$

$$+\left(\frac{1}{k}\right)^2\frac{\sigma_c^2}{4k}\left(1-\theta^2\right)\Bigg\}.$$

Tables 2, 3 and 4 compares prices computed by the two-factor formula above with Black's well known cost of carry model

$$F(S(T),\delta(T),t,T)=S(t)\exp[(r-\delta)\tau].$$

Two points are noteworthy from Table 2 (the results are the same for Tables 3 and 4).

Differences arise between the two models above depending on the level of the convenience yield. When the initial level of convenience yield is close to its long-run mean or equal to it, theoretical futures price computed by Black's model and two-factor model are not significantly different. In addition, when δ_t moves away from its long-run mean (above or below), the difference between the two prices increases. This is expected since Black's model treats the convenience yield as a constant. That is, regardless of the current position of the convenience yield, Black's model will always use the long run mean to value the futures contract. The two-factor model, on the other hand, incorporates the movements of the convenience into its price. Thus, when the convenience yield deviates from its mean, we see differences in the futures prices generated by the two models.

The next observation worth noting is when the speed of adjustment k increases. In this case, the difference in prices between Black's model and the two-factor model decreases. In Black's model, the convenience yield is presumed constant and maintained at it long run level throughout the entire life of the contract. The two factor model, however, allows for random movements in the convenience yield and will therefore yield different prices for the futures contract when the convenience deviates from its mean. The difference between the two models is decreasing as the mean reversion becomes stronger in the convenience yield. This occurs because the convenience yield reverts to its mean quicker, and when the value of the convenience yield reaches its mean, the two-factor futures price equals the price computed by Black's model. Note that when $k = \infty$ the two-factor model becomes blacks cost of carry solution.

In summary, the difference in theoretical prices between the Black model and two-factor model arise due to the treatment of the convenience yield. Black's model hold's it constant and the two-factor allows for random movements. We see that price difference between the two models occur when deviation of the convenience yield from its mean increases. Furthermore, this difference decreases as the speed of adjustment increases. These differences are illustrated in Figures 5 through 13.

4.3 Three-factor Futures model

Using expression (157) from chapter three,

$$F(S(T),\delta(T),r(T),T) = S(t)A(\tau)D_1(\tau)D_2(\tau)D_3(\tau)e^{-H_c(\tau)\delta(t)}\frac{1}{P(t,T)},$$

where $$A(\tau) = \exp\left(\frac{(H_c(\tau)-\tau)\left(k_c^2\alpha - k_c\lambda\sigma_c - \frac{\sigma_c^2}{2} + \rho_{cs}\sigma_s\sigma_c k_c\right)}{k_c^2} - \frac{\sigma_c^2 H_c^2(\tau)}{4k_c}\right),$$

$$D_1(\tau) = \exp\left[\frac{\sigma_c\sigma_r\rho_{rc}}{k_r+k_c}\left[\frac{(H_r(\tau)-\tau)}{k_r} + \frac{(H_c(\tau)-\tau))}{k_c} + H_r(\tau)H_c(\tau)\right]\right],$$

$$D_2(\tau) = \exp\left[\left(\frac{\sigma_s\sigma_r\rho_{rs}}{k_r}(\tau - H_r(\tau))\right)\right],$$

$$D_3(\tau) = \exp\left[-(H_r(\tau)-\tau)\frac{\sigma_r^2}{k_r^2} - \frac{\sigma_r^2 H_r^2(\tau)}{2k_r}\right].$$

we calculate theoretical prices for the three-factor futures model. These prices are calculated given various parameter values for the spot price, convenience yield and interest rate. These values are reported in Tables 5 through 16.

The three factor model is a simple extension of the two-factor, whereby a stochastic interest rate is added. With that said, the two-factor model is nested in the three factor model. That is, if we hold the interest rate constant, the three-factor futures model reduces to the two-factor model. Furthermore, looking at the tables 5 through 16, we see that for a particular interest rate and speed of adjustment, the futures price behavior is consistent with the two factor model. For instance, in table 5 section one we see that as the speed of adjustment for the convenience yield increases the futures price tends to pull toward a central value. Moreover, as the mean reversion in the convenience becomes stronger the futures price tends revert much faster. This behavior is identical to the two-factor futures model.

Differences between the three-factor model and the two-factor model are seen by the interaction between the interest rate and both the convenience yield and spot price. This behavior is given by expressions (159) – (161). Expression (159), $D_1(\tau)$, indicates the premium or discount that is applied to pricing futures contracts due to the correlation between the interest rate and the convenience yield. When the correlation between the interest rate and convenience yield are zero this term is equal to one. That is no premium or discount is applied. When the value is positive (for our case we have the correlation equal to 0.5), this term is less than one, telling us that the futures price is discounted. Intuitively, since the interest rate goes up the convenience yield rises as well. The futures price is being discounted from increases in the convenience, in addition, the futures price is being further discounted by increases in the interest rate. This accounts for the time value of money, being lost by the recipient of the futures contract. That is, the individual who holds the futures contract not only loses the ability to take advantage of selling the commodity when the convenience yield increases, but also misses out on the time value of money. When the correlation between the convenience yield and interest rate is negative, we see the opposite is true.

Expression (160), $D_2(\tau)$, introduces the premium or discount applied to the futures price due to the correlation between the interest rate and the spot price. When the correlation is zero this value is equal to one. That is, no premium or discount is enacted since random movements in the interest rate do not seem to influence movements in the spot price. When the correlation is positive (for this example it is 0.5) the coefficient is greater than one, thus adding a premium to the futures price. When it is negative the coefficient is less than one, thereby, discounting the futures price.

Expression (160), $D_3(\tau)$, introduces the premium or discount applied to the futures price due to the volatility of the interest rate. The interest rate volatility is an exponentially dampened function. For very short time periods this expression is greater than one, thereby adding a premium to the futures price. As the time to maturity increases this expression is decreasing and in particular is less than one. The indication is that as time to maturity goes up greater discounts are applied to the futures price due to decreased volatility.

5.4 One-factor Option Model

Theoretical option prices are calculated using expressions (114) through (116) from chapter four. This is

$$C(F,t) = e^{-r\tau_1} F(t) N(d_1) - e^{-r\tau_1} X N(d_2),$$

where

$$d_1 = \frac{\ln\left(\frac{F(t)}{X}\right) + \frac{1}{2}v^2}{v},$$

and

$$d_2 = d_1 - v.$$

In addition, the prices generated by the one-factor model are compare with theoretical option values that are calculated using the Black-Scholes pricing formula, which is

$$C(F,t) = e^{-r\tau_1} F(t) N(d_1) - e^{-r\tau_1} X N(d_2),$$

where

$$d_1 = \frac{\ln\left(\frac{F(t)}{X}\right) + \left(r + \frac{1}{2}\sigma^2\tau_1\right)}{\sigma\sqrt{\tau_1}},$$

and

$$d_2 = d_1 - \sigma\sqrt{\tau_1}.$$

There are a few observations worth noting.

First, the option prices from the Black-Scholes model are always greater than the option prices from the one-factor model. We expect this since the spot price is said to mean revert. In chapter four, we show that the diffusion for the futures price dynamic is an exponentially dampened process, and the rate of depreciation is a function of time and the speed of adjustment. The prices reported in Tables 14, 15 and 16 (and illustrated in Figures 14 – 19), reflect this phenomenon. As maturity decreases the one-factor prices converge to the Black-Scholes prices. At maturity the two prices are equal, since at this point the futures price equals the spot price.

A second observation about the behavior of the one-factor option price is that as the speed of adjustment parameter increases, the price difference between the two models increases. This is expected since the futures price volatility is inversely related to the speed of adjustment. Additionally, we see this discrepancy is between the one-factor model and the Black-Scholes model is further magnified when the maturity increases.

Lastly, we see that when the instantaneous volatility for the spot price increases the price difference between the models increases. This price difference is more pronounced the longer the option has to mature. Since the futures price volatility is a function of the spot price volatility, we should expect the futures price volatility to increase as the spot volatility increases, and it does. The increase in the futures price volatility is not uniform across maturities though. The futures price volatility is exponentially dampened over time and with respect to the speed of adjustment. Therefore, as the time to maturity increases we see the volatility is exponentially decreasing.

5.5 Two-factor Option Model

Theoretical option prices are calculated for the two-factor model and reported in Tables 17, 18, and 19. The pricing formula is given by expression (142), in chapter four and is

$$C(F,t) = e^{-r\tau_1} F(t) N(d_1) - e^{-r\tau_1} X N(d_2),$$

where

$$d_1 = \frac{\ln\left(\frac{F(t)}{X}\right) + \frac{1}{2}v^2}{v},$$

and

$d_2 = d_1 - v$.

The one-factor model is not a special case of the two factor model, but there are similarities between the two models. Both pricing models show that when the speed of adjustment increases volatility decreases. Furthermore, when the time to maturity increase the volatility decreases. These similarities are due to the mean reverting behavior in commodity prices.

Differences between the one-factor and two-factor model occur due to the treatment of mean reversion in the model. The one-factor model captures mean reversion entirely from the spot price, where as, the two-factor model introduces mean reversion through a stochastic convenience yield. The volatility for the two factor model is decreasing with respect to time and speed of adjustment, but unlike the one-factor model, the volatility is converging to a particular value and not zero. In addition, the two-factor option volatility is influenced by the correlation between the spot price and the convenience yield. As the correlation increases the volatility for the option price decreases. This is expected since the convenience yield is mean reverting. That is, as the correlation increases the spot price must be exhibiting mean reverting tendencies itself, which should further dampen the price volatility. Observations of the option prices in Tables 17 through 19 illustrate this point. Furthermore, the difference between the two-factor option prices, where $\rho = 0$ and $\rho = 0.766$, is decreasing as the speed of adjustment for the convenience yield increases. That is, as the tendency for the convenience yield to revert back to its mean gets stronger the volatility to the futures price converges to the volatility of the spot price. As a consequence, both cases are converging to the same volatility term, and therefore, to the same option price.

Table 1

The values below list the futures prices for a one-factor model. Spot prices follow a mean reverting process $dS(t) = k(\mu - \ln S(t))S(t)dt + \sigma S(t)dZ(t)$. Parameter values are set for $\mu = \ln(20)$, $\sigma = 0.334$, and the futures pricing equation is

$$F[S,T] = \exp\left[e^{-kT}\ln S(0) + \alpha^*\left(1-e^{-kT}\right) + \frac{\sigma^2}{4k}\left[1-e^{-2kT}\right]\right].$$

A. Spot Price = 20

	Speed of Adjustment					
	0.5	5	7	10	12	15
time to maturity						
0.000	20.000	20.000	20.000	20.000	20.000	20.000
0.083	19.998	19.987	19.984	19.982	19.981	19.981
0.250	19.985	19.943	19.946	19.953	19.958	19.965
0.500	19.945	19.906	19.925	19.945	19.954	19.963
0.750	19.891	19.894	19.921	19.944	19.954	19.963
1.000	19.828	19.890	19.921	19.944	19.954	19.963
1.250	19.761	19.889	19.920	19.944	19.954	19.963

B. Spot Price = 15

	Speed of Adjustment					
	0.5	5	7	10	12	15
time to maturity						
0.000	15.000	15.000	15.000	15.000	15.000	15.000
0.083	15.176	16.534	17.020	17.619	17.643	18.400
0.250	15.504	18.365	18.973	19.439	19.504	19.830
0.500	15.942	19.442	19.753	19.807	19.925	19.960
0.750	16.323	19.760	19.891	19.792	19.960	19.963
1.000	16.653	19.852	19.915	19.746	19.963	19.963
1.250	16.940	19.878	19.920	19.697	19.963	19.963

C. Spot Price = 25

	Speed of Adjustment					
	0.500	5.000	7.000	10.000	12.000	15.000
time to maturity						
0.000	25.000	25.000	25.000	25.000	25.000	25.000
0.083	24.771	23.155	22.635	22.017	21.691	21.300
0.250	24.334	21.260	20.734	20.322	20.181	20.070
0.500	23.731	20.274	20.060	19.975	19.965	19.965
0.750	23.188	19.999	19.945	19.947	19.954	19.963
1.000	22.702	19.920	19.925	19.945	19.954	19.963
1.250	22.267	19.898	19.921	19.944	19.954	19.963

Figure 2

Theoretical Futures Prices for a One-Factor Model
When the Spot Price is Equal to the Long Run Mean

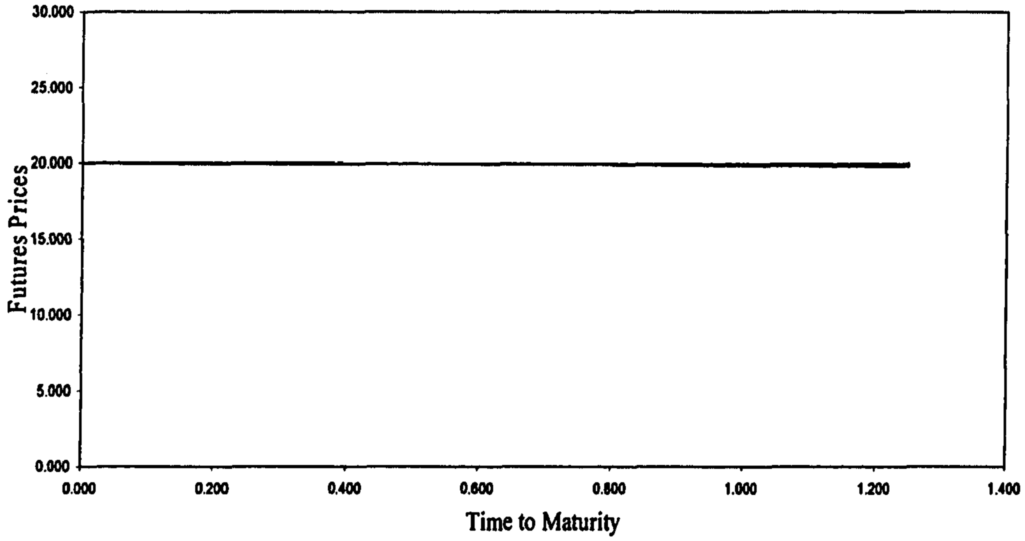


Figure 3

Futures Prices for a One-Factor Model
Spot Price Below Long Run Mean

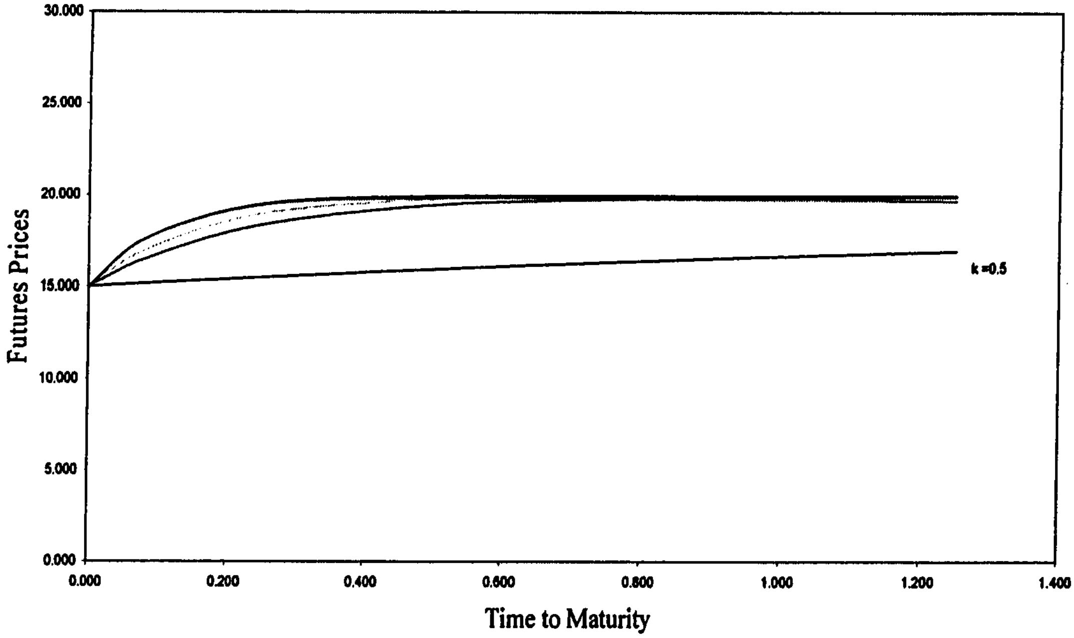


Figure 4

Futures Prices for a One-Factor Model
Spot Price Above Long Run Mean

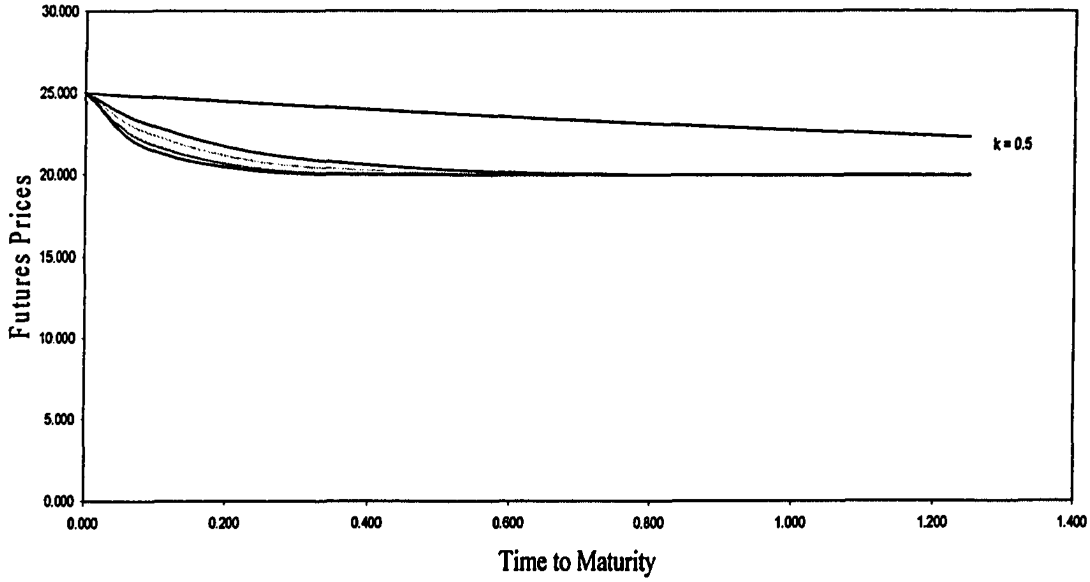


Table 2

This table illustrates the difference between Black's cost of carry model and a two-factor model with a stochastic spot price and mean reverting convenience yield. Black's cost of carry model is $F(S(T),T) = S(t)e^{(r-\alpha)t}$. Parameter values are: α = 0.1, where $\alpha = E(\delta(t))$, $\sigma_s = 0.393$, $\sigma_c = 0.1$, $\lambda = 0$, $\rho = 0$ $r = 0.15$, and $S(t) = 20$.

			Speed of Adjustment					
	C.C. Model	δ	0.5	1.876	5	7	10	15
	20.252	0.010	20.685	20.619	20.513	20.468	20.420	20.371
		0.030	20.588	20.536	20.455	20.420	20.383	20.345
		0.050	20.491	20.455	20.397	20.372	20.345	20.318
		0.070	20.395	20.373	20.338	20.324	20.308	20.292
		0.090	20.299	20.292	20.280	20.276	20.271	20.265
Three Month		0.100	20.251	20.252	20.252	20.252	20.252	20.252
		0.110	20.204	20.211	20.223	20.228	20.234	20.239
		0.130	20.109	20.131	20.165	20.180	20.196	20.213
		0.150	20.015	20.050	20.108	20.133	20.159	20.186
		0.170	19.921	19.971	20.050	20.085	20.122	20.160
		0.190	19.828	19.891	19.993	20.038	20.086	20.134

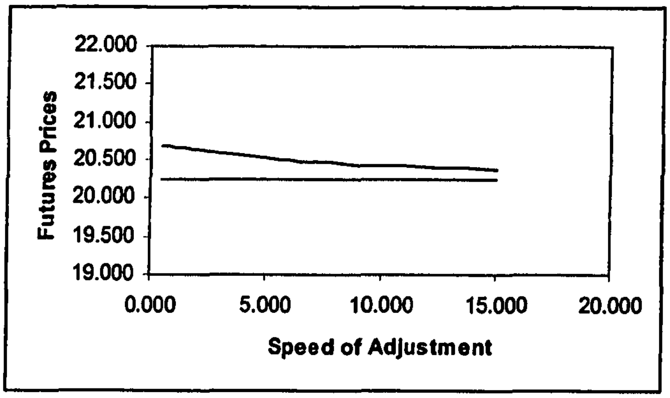
			Speed of Adjustment					
	C.C. Model	δ	0.5	1.876	5	7	10	15
	20.506	0.0100	21.339	21.114	20.848	20.764	20.690	20.630
		0.0300	21.151	20.977	20.771	20.706	20.649	20.602
		0.0500	20.965	20.842	20.695	20.649	20.608	20.575
		0.0700	20.780	20.707	20.619	20.592	20.567	20.547
		0.0900	20.597	20.573	20.544	20.535	20.527	20.520
Six Month		0.1000	20.506	20.506	20.506	20.506	20.506	20.506
		0.1100	20.416	20.440	20.469	20.478	20.486	20.493
		0.1300	20.236	20.308	20.394	20.421	20.445	20.465
		0.1500	20.058	20.176	20.319	20.365	20.405	20.438
		0.1700	19.881	20.046	20.244	20.308	20.364	20.411
		0.1900	19.706	19.916	20.170	20.252	20.324	20.384

			Speed of Adjustment					
	C.C. Model	δ	0.5	1.876	5	7	10	15
	20.764	0.010	21.967	21.530	21.132	21.031	20.952	20.889
		0.030	21.694	21.358	21.050	20.972	20.910	20.861
		0.050	21.424	21.186	20.968	20.912	20.868	20.834
		0.070	21.158	21.016	20.886	20.853	20.827	20.806
		0.090	20.895	20.848	20.805	20.794	20.785	20.778
Nine Month		0.100	20.763	20.765	20.764	20.764	20.764	20.764
		0.110	20.635	20.681	20.724	20.735	20.743	20.750
		0.130	20.379	20.515	20.643	20.676	20.702	20.723
		0.150	20.125	20.350	20.562	20.617	20.661	20.695
		0.170	19.875	20.187	20.482	20.559	20.619	20.668
		0.190	19.628	20.025	20.402	20.500	20.578	20.640

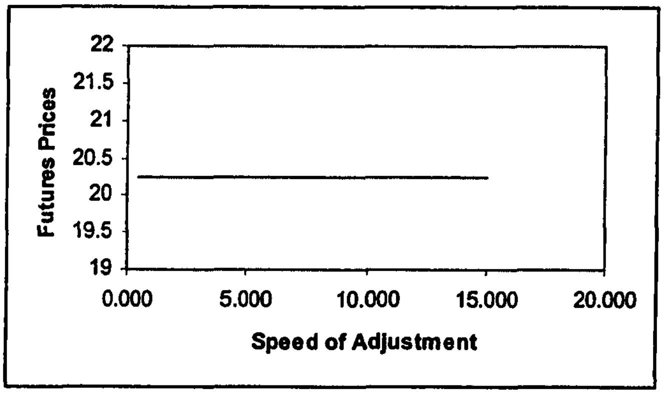
Figure 5

Three Month Theoretical Futures Prices from Table 2

A. Delta is below its long-run mean $\delta = 0.01$



B. Delta is equal to its long-run mean $\delta = 0.1$



C. Delta is below its long-run mean $\delta = 0.019$

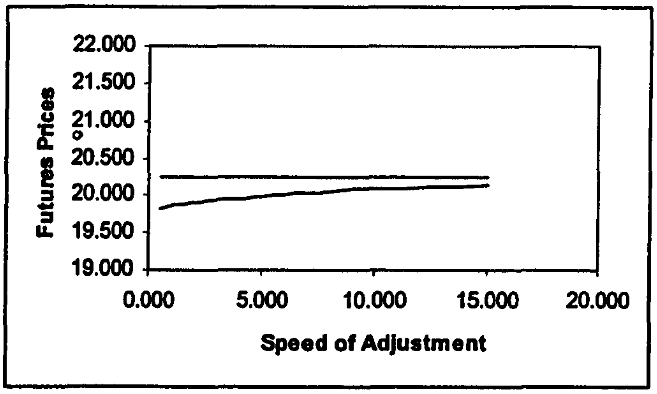
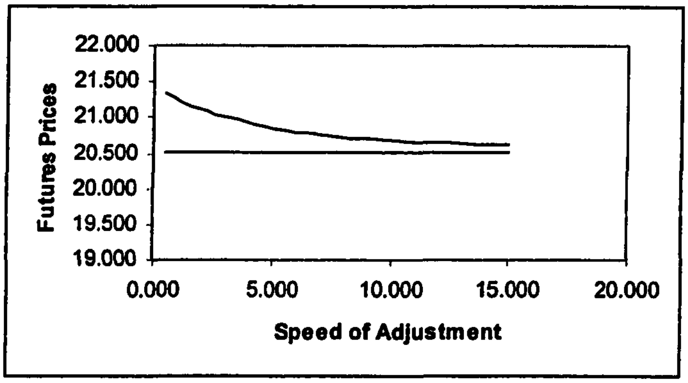


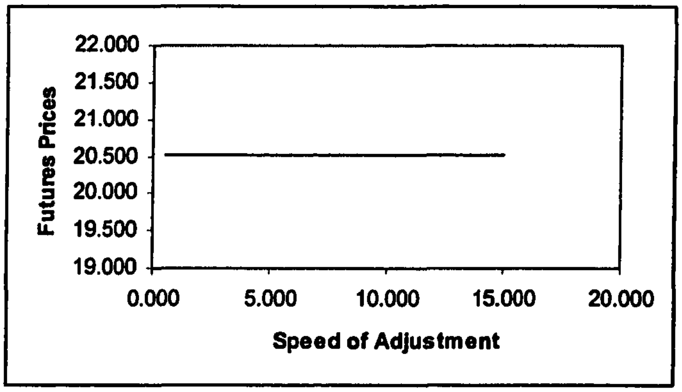
Figure 6

Six Month Theoretical Futures Prices from Table 2

A. Delta is below its long-run mean $\delta = 0.01$



B. Delta is equal to its long-run mean $\delta = 0.1$



C. Delta is below its long-run mean $\delta = 0.019$

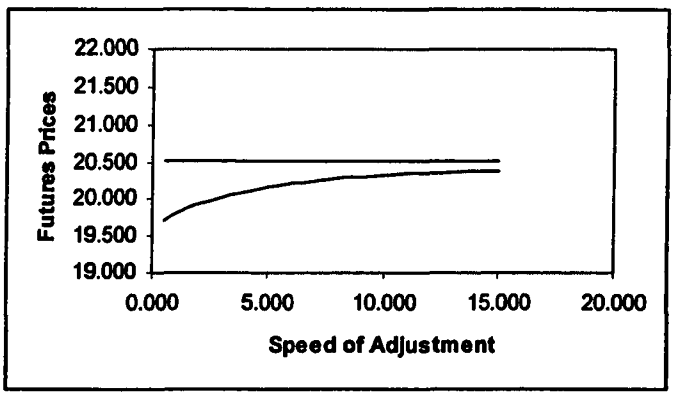
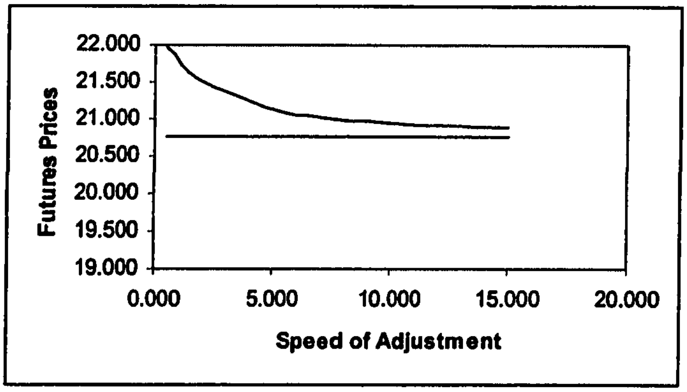


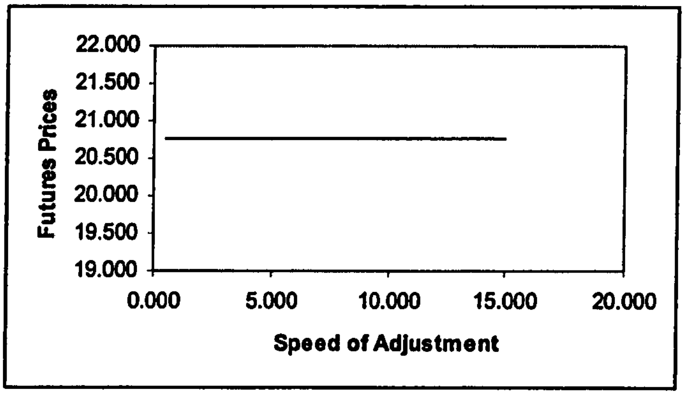
Figure 7

Nine Month Theoretical Futures Prices from Table 2

A. Delta is below its long-run mean $\delta = 0.01$



B. Delta is equal to its long-run mean $\delta = 0.1$



C. Delta is below its long-run mean $\delta = 0.019$

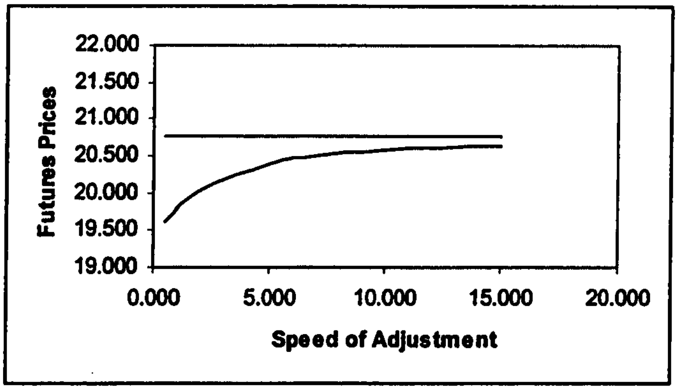


Table 3

This table illustrates the difference between Black's cost of carry model and a two-factor model with a stochastic spot price and mean reverting convenience yield. Black's cost of carry model is $F(S(T),T) = S(t)e^{(r-\alpha)t}$. Parameter values are: α = 0.1, where $\alpha = E(\delta(t))$, $\sigma_s = 0.393$, σ_c=0 .1, $\lambda = 0$, $\rho = 0.766$ $r = 0.15$, and $S(t) = 20$.

	C.C.Model	δ	Speed of Adjustment 0.5	1.876	5	7	10	15
	20.252	0.010	20.683	20.617	20.512	20.452	20.401	20.349
		0.030	20.586	20.535	20.454	20.404	20.363	20.322
		0.050	20.489	20.453	20.395	20.356	20.326	20.296
		0.070	20.393	20.372	20.337	20.308	20.289	20.269
		0.090	20.297	20.290	20.279	20.260	20.251	20.243
Three Month		0.100	20.192	20.250	20.250	20.250	20.236	20.233
		0.110	20.202	20.210	20.221	20.212	20.214	20.217
		0.130	20.108	20.129	20.164	20.164	20.177	20.190
		0.150	20.013	20.049	20.106	20.117	20.140	20.164
		0.170	19.919	19.969	20.049	20.069	20.103	20.138
		0.190	19.826	19.889	19.992	20.022	20.066	20.112

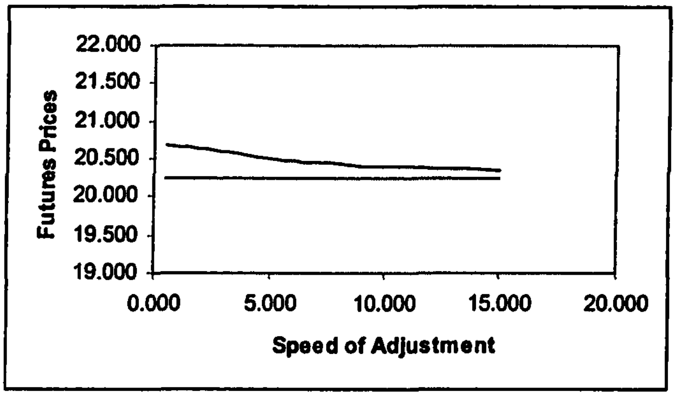
	C.C.Model	δ	Speed of Adjustment 0.5	1.876	5	7	10	15
	20.506	0.010	21.332	21.108	20.844	20.760	20.688	20.628
		0.030	21.144	20.971	20.767	20.703	20.647	20.600
		0.050	20.958	20.836	20.691	20.646	20.606	20.573
		0.070	20.773	20.701	20.616	20.588	20.565	20.546
		0.090	20.590	20.567	20.540	20.531	20.524	20.518
Six Month		0.100	20.389	20.499	20.500	20.502	20.503	20.504
		0.110	20.409	20.434	20.465	20.475	20.483	20.491
		0.130	20.229	20.302	20.390	20.418	20.443	20.464
		0.150	20.051	20.171	20.315	20.362	20.402	20.436
		0.170	19.874	20.040	20.241	20.305	20.362	20.409
		0.190	19.699	19.911	20.166	20.249	20.321	20.382

	C.C.Model	δ	Speed of Adjustment 0.5	1.876	5	7	10	15
	20.764	0.010	21.951	21.518	21.125	21.026	20.948	20.886
		0.030	21.678	21.346	21.043	20.966	20.906	20.858
		0.050	21.408	21.175	20.961	20.907	20.864	20.831
		0.070	21.142	21.005	20.879	20.847	20.822	20.803
		0.090	20.879	20.836	20.798	20.788	20.781	20.775
Nine Month		0.100	20.592	20.749	20.753	20.757	20.759	20.760
		0.110	20.620	20.669	20.717	20.729	20.739	20.748
		0.130	20.363	20.504	20.636	20.670	20.698	20.720
		0.150	20.110	20.339	20.556	20.612	20.657	20.692
		0.170	19.860	20.176	20.475	20.553	20.615	20.665
		0.190	19.613	20.014	20.396	20.495	20.574	20.637

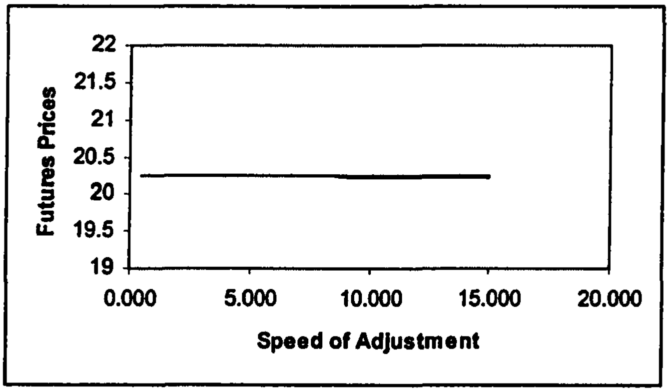
Figure 8

Three Month Theoretical Futures Prices from Table 3

A. Delta is below its long-run mean ($\delta = 0.01$)



B. Delta is equal to its long-run mean ($\delta = 0.1$)



C. Delta is above its long-run mean ($\delta = 0.19$)

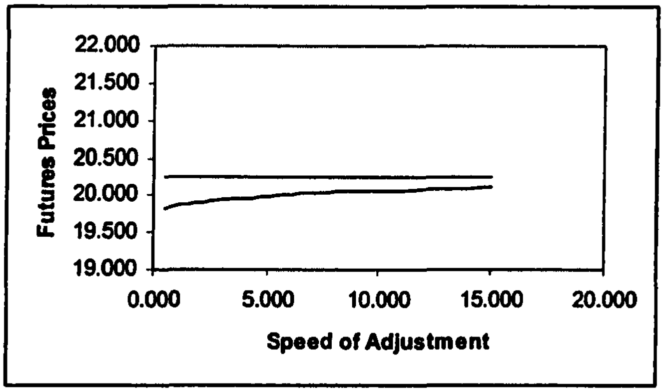
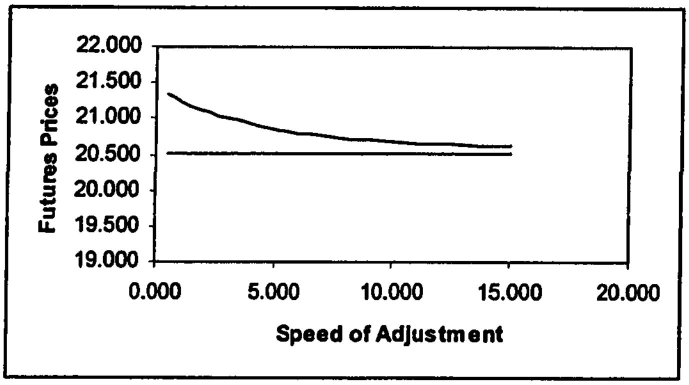


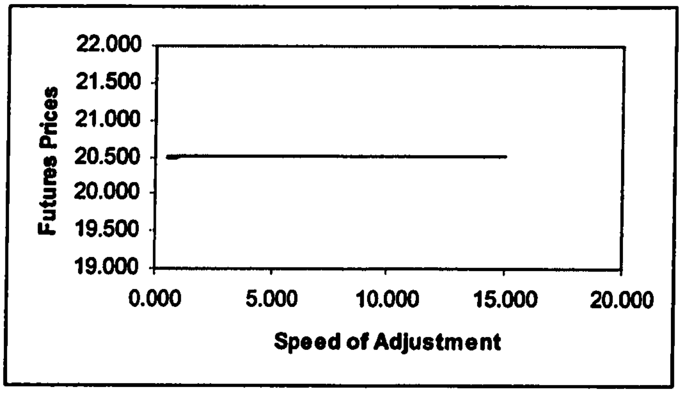
Figure 9

Six Month Theoretical Futures Prices from Table 3

A. Delta is below its long-run mean $\delta = 0.01$



B. Delta is equal to its long-run mean $\delta = 0.1$



C. Delta is below its long-run mean $\delta = 0.019$

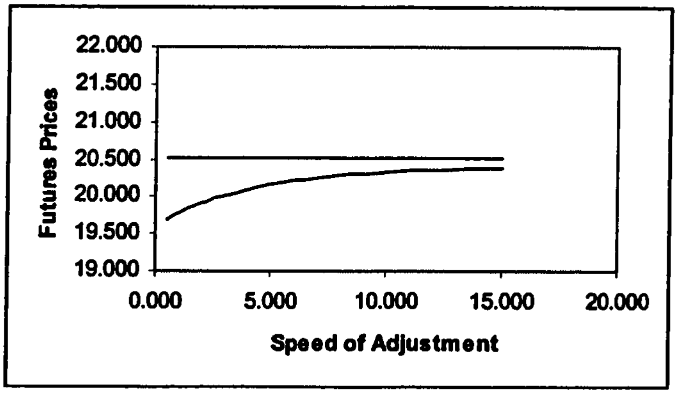
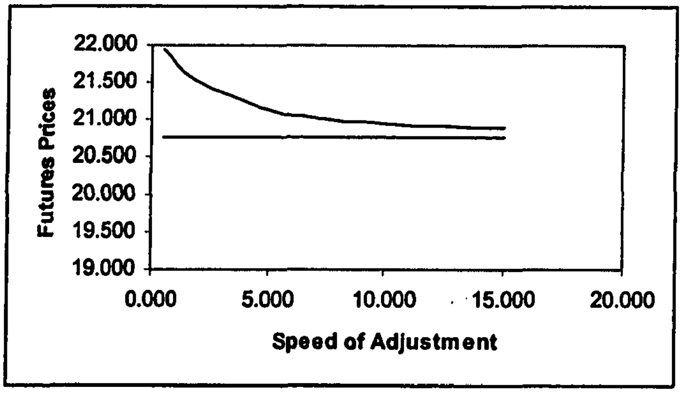


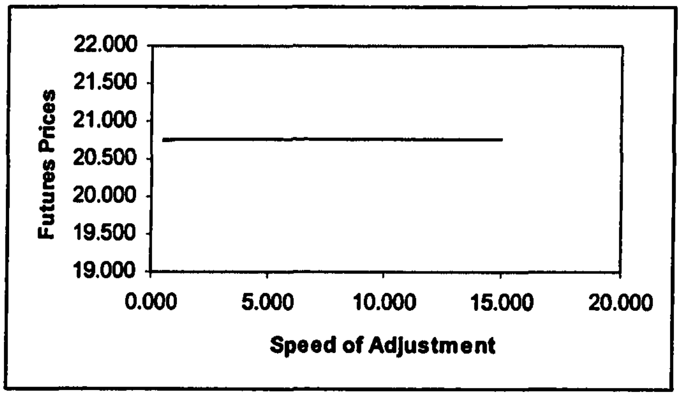
Figure 10

Nine Month Theoretical Futures Prices from Table 3

A. Delta is below its long-run mean $\delta = 0.01$



B. Delta is equal to its long-run mean $\delta = 0.1$



C. Delta is below its long-run mean $\delta = 0.019$

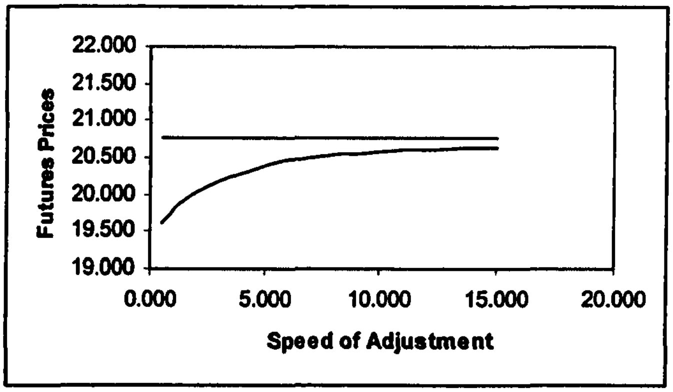


Table 4

This table illustrates the difference between Black's cost of carry model and a two-factor model with a stochastic spot price and mean reverting convenience yield. Black's cost of carry model is $F(S(T),T) = S(t)e^{(r-\alpha)t}$. Parameter values are: α = 0.1, where $\alpha = E(\delta(t))$, $\sigma_s = 0.393$, $\sigma_c = 0.1$, $\lambda = 0.198$, $\rho = 0.766$ $r = 0.15$, and $S(t) = 20$.

			Speed of Adjustment					
	C.C.Model	δ	0.5	1.876	5	7	10	15
	20.252	0.010	20.684	20.618	20.513	20.463	20.414	20.364
		0.030	20.587	20.536	20.454	20.415	20.376	20.337
		0.050	20.490	20.454	20.396	20.366	20.339	20.311
		0.070	20.394	20.373	20.338	20.318	20.301	20.284
		0.090	20.299	20.291	20.280	20.270	20.264	20.258
Three Month		0.100	20.153	20.251	20.251	20.251	20.247	20.246
		0.110	20.203	20.211	20.222	20.223	20.227	20.231
		0.130	20.109	20.130	20.165	20.175	20.190	20.205
		0.150	20.014	20.050	20.107	20.127	20.153	20.179
		0.170	19.921	19.970	20.050	20.080	20.116	20.153
		0.190	19.827	19.891	19.993	20.033	20.079	20.126

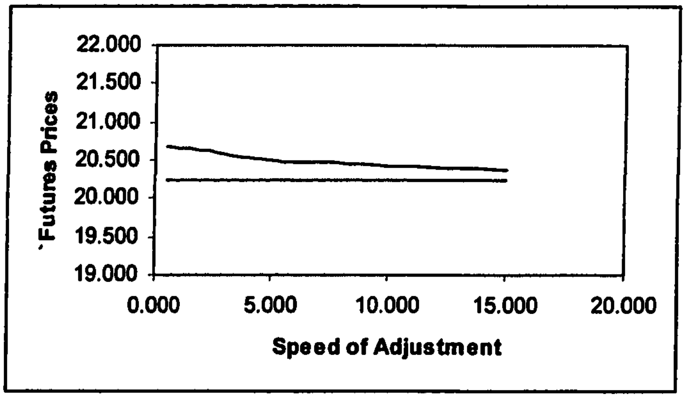
			Speed of Adjustment					
	C.C.Model	δ	0.5	1.876	5	7	10	15
	20.506	0.010	21.337	21.112	20.846	20.762	20.690	20.629
		0.030	21.149	20.975	20.770	20.705	20.648	20.602
		0.050	20.963	20.840	20.694	20.648	20.607	20.574
		0.070	20.778	20.705	20.618	20.591	20.567	20.547
		0.090	20.595	20.571	20.543	20.534	20.526	20.519
Six Month		0.100	20.313	20.504	20.504	20.505	20.505	20.505
		0.110	20.414	20.438	20.467	20.477	20.485	20.492
		0.130	20.234	20.306	20.392	20.420	20.444	20.465
		0.150	20.055	20.174	20.318	20.364	20.404	20.437
		0.170	19.879	20.044	20.243	20.307	20.363	20.410
		0.190	19.704	19.914	20.169	20.251	20.323	20.383

			Speed of Adjustment					
	C.C.Model	δ	0.5	1.876	5	7	10	15
	20.764	0.010	21.961	21.526	21.130	21.030	20.950	20.888
		0.030	21.396	21.688	21.353	21.048	20.970	20.909
		0.050	21.419	21.182	20.966	20.910	20.867	20.833
		0.070	21.153	21.012	20.884	20.851	20.825	20.805
		0.090	20.890	20.844	20.802	20.792	20.784	20.777
Nine Month		0.100	20.480	20.759	20.760	20.762	20.762	20.763
		0.110	20.630	20.677	20.721	20.733	20.742	20.749
		0.130	20.373	20.511	20.641	20.674	20.701	20.722
		0.150	20.120	20.347	20.560	20.615	20.659	20.694
		0.170	19.870	20.183	20.480	20.557	20.618	20.667
		0.190	19.623	20.022	20.400	20.498	20.577	20.639

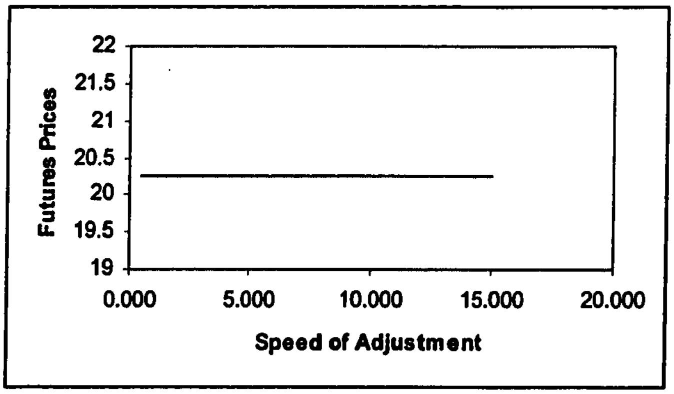
Figure 11

Three Month Theoretical Futures Prices from Table 4

A. Delta is below its long-run mean ($\delta = 0.01$)



B. Delta is equal to its long-run mean ($\delta = 0.1$)



C. Delta is above its long-run mean ($\delta = 0.19$)

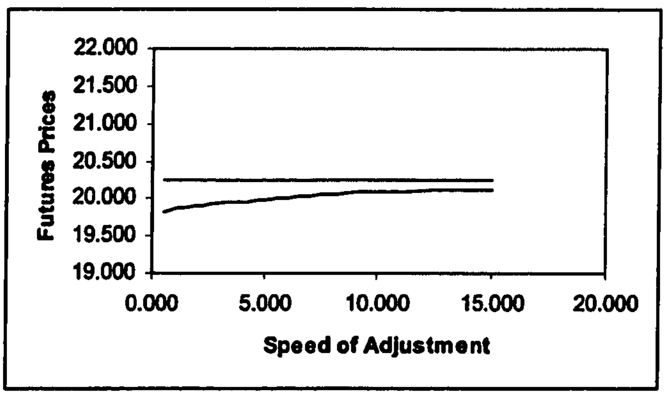
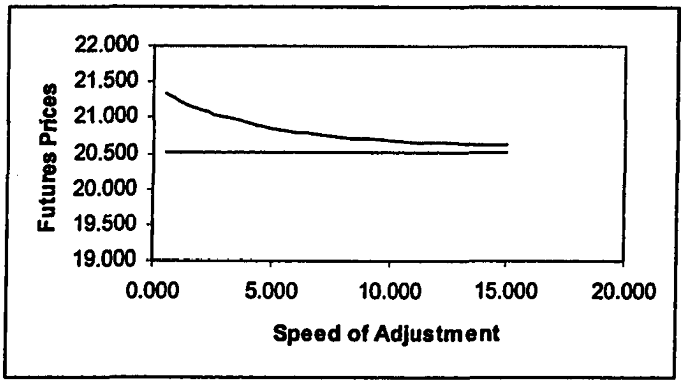


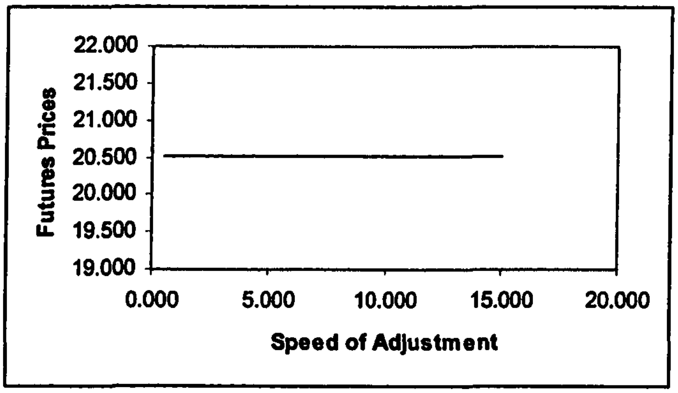
Figure 12

Six Month Theoretical Futures Prices from Table 4

A. Delta is below its long-run mean $\delta = 0.01$



B. Delta is equal to its long-run mean $\delta = 0.1$



C. Delta is below its long-run mean $\delta = 0.019$

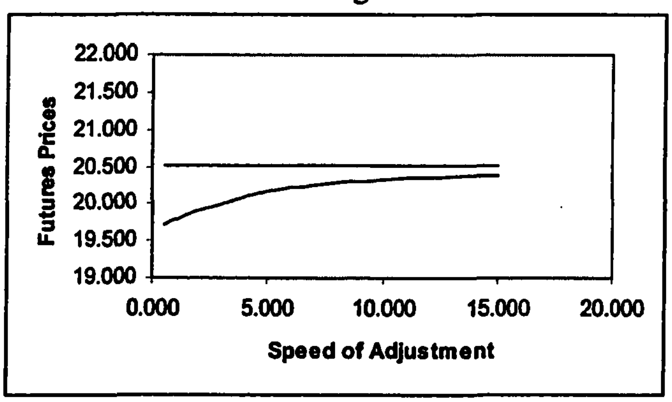
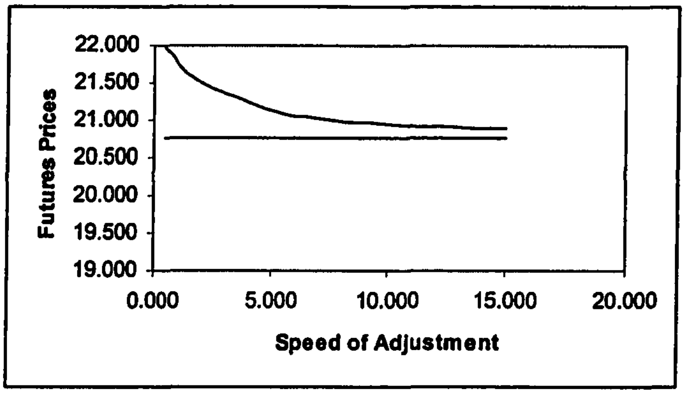


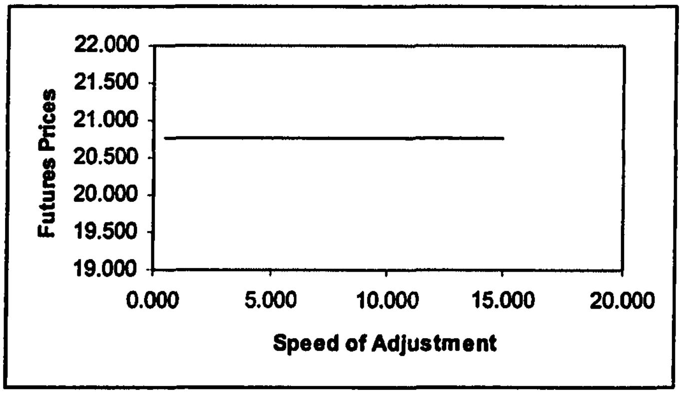
Figure 13

Nine Month Theoretical Futures Prices from Table 4

A. Delta is below its long-run mean $\delta = 0.01$



B. Delta is equal to its long-run mean $\delta = 0.1$



C. Delta is below its long-run mean $\delta = 0.019$

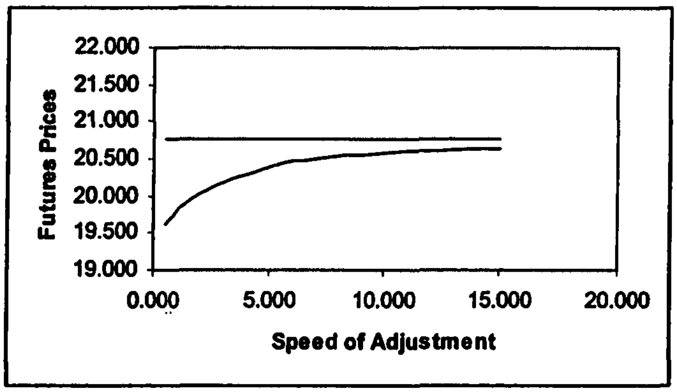


Table 5

This table shows theoretical values for the one month, three-factor futures model. Parameters for the model are $\rho_{sc} = \rho_{sr} = \rho_{cr} = 0$, $S(t)$ = 20, r = .15, $\sigma_s = .393$, $\sigma_c = \sigma_r = .1$ and $\lambda = 0.198$

	δ	Speed of Adjustment k_c 0.5	1.876	5	10
	0.03	20.202	20.194	20.181	20.164
	0.05	20.169	20.163	20.153	20.141
	0.07	20.136	20.132	20.126	20.118
	0.09	20.103	20.101	20.098	20.095
$k_r = 0.5$	0.1	20.086	20.085	20.085	20.084
	0.11	20.070	20.070	20.071	20.073
	0.13	20.037	20.039	20.044	20.050
	0.15	20.005	20.008	20.016	20.027
	0.17	19.972	19.977	19.989	20.005
	δ	0.5	1.876	5	10
	0.03	20.198	20.191	20.177	20.160
	0.05	20.165	20.160	20.150	20.138
	0.07	20.132	20.129	20.122	20.115
	0.09	20.100	20.098	20.095	20.092
$k_r = 5$	0.1	20.083	20.082	20.081	20.081
	0.11	20.067	20.067	20.068	20.069
	0.13	20.034	20.036	20.040	20.047
	0.15	20.001	20.005	20.013	20.024
	0.17	19.969	19.974	19.986	20.001
	δ	0.5	1.876	5	10
	0.03	20.197	20.190	20.176	20.159
	0.05	20.164	20.159	20.149	20.137
	0.07	20.131	20.128	20.122	20.114
	0.09	20.099	20.097	20.094	20.091
$k_r = 10$	0.1	20.082	20.081	20.080	20.080
	0.11	20.066	20.066	20.067	20.068
	0.13	20.033	20.035	20.039	20.046
	0.15	20.000	20.004	20.012	20.023
	0.17	19.968	19.973	19.985	20.001

Table 6

This table shows theoretical values for the six month, three-factor futures model. Parameters for the model are $\rho_{sc} = \rho_{sr} = \rho_{cr} = 0$, $S(t)$ = 20, r = .15, $\sigma_s = .393$, $\sigma_c = \sigma_r = .1$ and $\lambda = 0.198$

		Speed of Adjustment k_c			
	δ	0.5	1.876	5	10
	0.03	21.257	21.049	20.828	20.698
	0.05	21.070	20.913	20.751	20.657
	0.07	20.884	20.778	20.675	20.616
	0.09	20.701	20.643	20.599	20.575
$k_r = 0.5$	0.1	20.609	20.576	20.562	20.555
	0.11	20.518	20.510	20.524	20.535
	0.13	20.337	20.377	20.449	20.494
	0.15	20.158	20.245	20.374	20.453
	0.17	19.981	20.114	20.299	20.413
	δ	0.5	1.876	5	10
	0.03	21.204	20.996	20.775	20.646
	0.05	21.017	20.860	20.699	20.605
	0.07	20.832	20.725	20.623	20.564
	0.09	20.648	20.591	20.547	20.523
$k_r = 5$	0.1	20.557	20.524	20.510	20.503
	0.11	20.466	20.458	20.472	20.483
	0.13	20.286	20.326	20.397	20.442
	0.15	20.107	20.194	20.322	20.401
	0.17	19.930	20.064	20.248	20.361
	δ	0.5	1.876	5	10
	0.03	21.209	21.001	20.780	20.652
	0.05	21.022	20.866	20.704	20.611
	0.07	20.837	20.731	20.628	20.570
	0.09	20.654	20.597	20.553	20.529
$k_r = 10$	0.1	20.562	20.530	20.515	20.508
	0.11	20.472	20.463	20.477	20.488
	0.13	20.291	20.331	20.402	20.447
	0.15	20.113	20.200	20.328	20.407
	0.17	19.935	20.069	20.253	20.366

Table 7

This table shows theoretical values for the one year, three-factor futures model. Parameters for the model are $\rho_{sc} = \rho_{sr} = \rho_{cr} = 0$, $S(t)$ = 20, r = .15, $\sigma_s = .393$, $\sigma_c = \sigma_r = .1$ and $\lambda = 0.198$

	δ	Speed of Adjustment k_c 0.5	1.876	5	10
	0.03	22.640	21.974	21.529	21.354
	0.05	22.286	21.777	21.444	21.312
	0.07	21.938	21.581	21.359	21.269
	0.09	21.596	21.387	21.274	21.227
$k_r = 0.5$	0.1	21.426	21.291	21.232	21.205
	0.11	21.258	21.195	21.190	21.184
	0.13	20.926	21.004	21.106	21.142
	0.15	20.600	20.816	21.022	21.100
	0.17	20.278	20.629	20.939	21.058
	δ	0.5	1.876	5	10
	0.03	22.467	21.807	21.365	21.192
	0.05	22.117	21.611	21.281	21.150
	0.07	21.771	21.417	21.196	21.107
	0.09	21.431	21.224	21.112	21.065
$k_r = 5$	0.1	21.263	21.129	21.070	21.044
	0.11	21.097	21.034	21.029	21.023
	0.13	20.767	20.845	20.945	20.981
	0.15	20.443	20.657	20.862	20.939
	0.17	20.124	20.472	20.779	20.897
	δ	0.5	1.876	5	10
	0.03	22.473	21.813	21.371	21.197
	0.05	22.122	21.617	21.286	21.155
	0.07	21.777	21.422	21.202	21.113
	0.09	21.437	21.230	21.118	21.071
$k_r = 10$	0.1	21.269	21.134	21.076	21.049
	0.11	21.102	21.039	21.034	21.028
	0.13	20.773	20.850	20.951	20.986
	0.15	20.448	20.662	20.867	20.945
	0.17	20.129	20.477	20.785	20.903

Table 8

This table shows theoretical values for the one month, three-factor futures model. Parameters for the model are $\rho_{cr} = \rho_{sr} = .5$, $\rho_{sc} = 0.766$, $S(t) = 20$, $r = .15$, $\sigma_s = .393$, $\sigma_c = \sigma_r = .1$ and $\lambda = 0.198$

	δ	Speed of Adjustment k_c: 0.5	1.876	5	10
	0.03	20.201	20.193	20.178	20.161
	0.05	20.168	20.161	20.151	20.138
	0.07	20.135	20.130	20.123	20.115
	0.09	20.102	20.099	20.096	20.093
$k_r = 0.5$	0.1	20.086	20.084	20.082	20.081
	0.11	20.069	20.068	20.069	20.070
	0.13	20.037	20.037	20.041	20.047
	0.15	20.004	20.006	20.014	20.025
	0.17	19.971	19.976	19.987	20.002
	δ	0.5	1.876	5	10
	0.03	20.197	20.189	20.175	20.158
	0.05	20.164	20.158	20.147	20.135
	0.07	20.132	20.127	20.120	20.112
	0.09	20.099	20.096	20.093	20.089
$k_r = 5$	0.1	20.082	20.080	20.079	20.078
	0.11	20.066	20.065	20.065	20.067
	0.13	20.033	20.034	20.038	20.044
	0.15	20.001	20.003	20.011	20.021
	0.17	19.968	19.972	19.983	19.999
	δ	0.5	1.876	5	10
	0.03	20.196	20.188	20.174	20.156
	0.05	20.163	20.157	20.146	20.134
	0.07	20.130	20.126	20.119	20.111
	0.09	20.098	20.095	20.092	20.088
$k_r = 10$	0.1	20.081	20.079	20.078	20.077
	0.11	20.065	20.064	20.064	20.065
	0.13	20.032	20.033	20.037	20.043
	0.15	19.999	20.002	20.010	20.020
	0.17	19.967	19.971	19.982	19.998

Table 9

This table shows theoretical values for the six month, three-factor futures model. Parameters for the model are $\rho_{cr} = \rho_{sr} = .5$, $\rho_{sc} = 0.766$, $S(t) = 20$, $r = .15$, $\sigma_s = .393$, $\sigma_c = \sigma_r = .1$ and $\lambda = 0.198$

	δ	Speed of Adjustment k_c 0.5	1.876	5	10
	0.03	21.228	21.005	20.794	20.682
	0.05	21.041	20.869	20.717	20.641
	0.07	20.856	20.734	20.641	20.600
	0.09	20.672	20.600	20.566	20.559
$k_r = 0.5$	0.1	20.581	20.533	20.528	20.538
	0.11	20.490	20.467	20.490	20.518
	0.13	20.309	20.334	20.415	20.477
	0.15	20.131	20.203	20.340	20.437
	0.17	19.953	20.072	20.266	20.396
	δ	0.5	1.876	5	10
	0.03	21.155	20.932	20.721	20.609
	0.05	20.968	20.796	20.645	20.568
	0.07	20.784	20.662	20.569	20.527
	0.09	20.600	20.528	20.494	20.486
$k_r = 5$	0.1	20.510	20.462	20.456	20.466
	0.11	20.419	20.396	20.419	20.446
	0.13	20.239	20.264	20.344	20.405
	0.15	20.061	20.133	20.269	20.365
	0.17	19.884	20.002	20.195	20.324
	δ	0.5	1.876	5	10
	0.03	21.151	20.928	20.717	20.605
	0.05	20.965	20.793	20.641	20.564
	0.07	20.780	20.659	20.566	20.523
	0.09	20.597	20.525	20.490	20.483
$k_r = 10$	0.1	20.506	20.458	20.453	20.462
	0.11	20.416	20.392	20.415	20.442
	0.13	20.236	20.260	20.340	20.402
	0.15	20.058	20.129	20.266	20.361
	0.17	19.881	19.999	20.191	20.321

Table 10

This table shows theoretical values for the one year, three-factor futures model. Parameters for the model are $\rho_{cr} = \rho_{sr} = .5$, $\rho_{sc} = 0.766$, $S(t) = 20$, $r = .15$, $\sigma_s = .393$, $\sigma_c = \sigma_r = .1$ and $\lambda = 0.198$

		Speed of Adjustment k_c			
	δ	0.5	1.876	5	10
	0.03	22.513	21.851	21.493	21.384
	0.05	22.161	21.655	21.408	21.341
	0.07	21.815	21.460	21.323	21.299
	0.09	21.474	21.267	21.239	21.256
$k_r = 0.5$	0.1	21.306	21.172	21.196	21.235
	0.11	21.139	21.076	21.154	21.214
	0.13	20.809	20.887	21.071	21.171
	0.15	20.484	20.699	20.987	21.129
	0.17	20.164	20.513	20.904	21.087
	δ	0.5	1.876	5	10
	0.03	22.242	21.583	21.224	21.114
	0.05	21.895	21.389	21.140	21.071
	0.07	21.553	21.197	21.056	21.029
	0.09	21.216	21.006	20.972	20.987
$k_r = 5$	0.1	21.050	20.912	20.931	20.966
	0.11	20.885	20.817	20.889	20.945
	0.13	20.559	20.630	20.806	20.904
	0.15	20.238	20.445	20.724	20.862
	0.17	19.922	20.261	20.642	20.820
	δ	0.5	1.876	5	10
	0.03	22.221	21.561	21.201	21.091
	0.05	21.874	21.367	21.117	21.048
	0.07	21.533	21.175	21.034	21.006
	0.09	21.196	20.985	20.950	20.964
$k_r = 10$	0.1	21.030	20.891	20.909	20.943
	0.11	20.865	20.796	20.867	20.923
	0.13	20.540	20.610	20.784	20.881
	0.15	20.219	20.424	20.702	20.839
	0.17	19.903	20.241	20.620	20.797

Table 11

This table shows theoretical values for the one month, three-factor futures model. Parameters for the model are $\rho_{cr} = \rho_{sr} = -0.5$, $\rho_{sc} = 0.766$, $S(t)$ = 20, r = .15, $\sigma_s = .393$, $\sigma_c = \sigma_r = .1$ and $\lambda = 0.198$

		Speed of Adjustment k_c			
	δ	0.5	1.876	5	10
	0.03	20.198	20.190	20.176	20.158
	0.05	20.165	20.159	20.148	20.135
	0.07	20.132	20.128	20.121	20.113
	0.09	20.099	20.097	20.093	20.090
$k_r = 0.5$	0.1	20.083	20.081	20.080	20.079
	0.11	20.067	20.066	20.066	20.067
	0.13	20.034	20.035	20.039	20.045
	0.15	20.001	20.004	20.011	20.022
	0.17	19.969	19.973	19.984	19.999
	δ	0.5	1.876	5	10
	0.03	20.195	20.187	20.173	20.155
	0.05	20.162	20.156	20.145	20.132
	0.07	20.129	20.125	20.118	20.110
	0.09	20.096	20.094	20.090	20.087
$k_r = 5$	0.1	20.080	20.078	20.077	20.076
	0.11	20.064	20.063	20.063	20.064
	0.13	20.031	20.032	20.036	20.042
	0.15	19.998	20.001	20.008	20.019
	0.17	19.966	19.970	19.981	19.996
	δ	0.5	1.876	5	10
	0.03	20.194	20.186	20.172	20.154
	0.05	20.161	20.155	20.144	20.132
	0.07	20.128	20.124	20.117	20.109
	0.09	20.096	20.093	20.089	20.086
$k_r = 10$	0.1	20.079	20.077	20.076	20.075
	0.11	20.063	20.062	20.062	20.063
	0.13	20.030	20.031	20.035	20.041
	0.15	19.997	20.000	20.007	20.018
	0.17	19.965	19.969	19.980	19.995

Table 12

This table shows theoretical values for the six month, three-factor futures model. Parameters for the model are $\rho_{cr} = \rho_{sr} = -0.5$, $\rho_{sc} = 0.766$, $S(t) = 20$, $r = .15$, $\sigma_s = .393$, $\sigma_c = \sigma_r = .1$ and $\lambda = 0.198$

	δ	Speed of Adjustment k_c 0.5	1.876	5	10
	0.03	21.140	20.916	20.703	20.590
	0.05	20.953	20.780	20.627	20.550
	0.07	20.769	20.646	20.552	20.509
	0.09	20.586	20.513	20.476	20.468
$k_r = 0.5$	0.1	20.495	20.446	20.439	20.448
	0.11	20.404	20.380	20.401	20.428
	0.13	20.225	20.248	20.327	20.387
	0.15	20.047	20.117	20.252	20.347
	0.17	19.870	19.987	20.178	20.306
	δ	0.5	1.876	5	10
	0.03	21.106	20.883	20.671	20.559
	0.05	20.920	20.748	20.595	20.518
	0.07	20.736	20.614	20.520	20.477
	0.09	20.553	20.480	20.445	20.437
$k_r = 5$	0.1	20.462	20.414	20.407	20.416
	0.11	20.372	20.348	20.370	20.396
	0.13	20.192	20.216	20.295	20.356
	0.15	20.015	20.085	20.221	20.315
	0.17	19.838	19.956	20.147	20.275
	δ	0.5	1.876	5	10
	0.03	21.120	20.897	20.686	20.573
	0.05	20.934	20.762	20.610	20.533
	0.07	20.750	20.628	20.534	20.492
	0.09	20.567	20.494	20.459	20.451
$k_r = 10$	0.1	20.476	20.428	20.422	20.431
	0.11	20.386	20.362	20.384	20.411
	0.13	20.206	20.230	20.309	20.370
	0.15	20.028	20.099	20.235	20.330
	0.17	19.852	19.969	20.161	20.289

Table 13

This table shows theoretical values for the one year, three-factor futures model. Parameters for the model are $\rho_{cr} = \rho_{sr} = -0.5$, $\rho_{sc} = 0.766$, $S(t)$ = 20, r = .15, $\sigma_s = .393$, $\sigma_c = \sigma_r = .1$ and $\lambda = 0.198$

	δ	Speed of Adjustment k_c: 0.5	1.876	5	10
	0.03	22.190	21.521	21.153	21.038
	0.05	21.844	21.328	21.069	20.996
	0.07	21.503	21.136	20.985	20.954
	0.09	21.167	20.946	20.902	20.912
$k_r = 0.5$	0.1	21.001	20.852	20.861	20.891
	0.11	20.837	20.758	20.819	20.870
	0.13	20.511	20.571	20.737	20.828
	0.15	20.191	20.387	20.655	20.787
	0.17	19.876	20.203	20.573	20.745
	δ	0.5	1.876	5	10
	0.03	22.120	21.459	21.097	20.984
	0.05	21.774	21.266	21.013	20.942
	0.07	21.434	21.075	20.930	20.900
	0.09	21.100	20.885	20.847	20.859
$k_r = 5$	0.1	20.934	20.791	20.805	20.838
	0.11	20.770	20.698	20.764	20.817
	0.13	20.446	20.512	20.682	20.775
	0.15	20.127	20.327	20.600	20.734
	0.17	19.812	20.145	20.518	20.693
	δ	0.5	1.876	5	10
	0.03	22.152	21.491	21.130	21.018
	0.05	21.806	21.298	21.046	20.976
	0.07	21.466	21.107	20.963	20.934
	0.09	21.130	20.917	20.879	20.892
$k_r = 10$	0.1	20.965	20.823	20.838	20.871
	0.11	20.800	20.729	20.797	20.850
	0.13	20.476	20.543	20.714	20.809
	0.15	20.156	20.358	20.632	20.767
	0.17	19.841	20.175	20.550	20.726

Table 14

This table shows the theoretical option prices for the one-factor futures model.

Option Parameters	Futures Price	Black-Scholes	One-factor Model
	15	0.021	0.005
$\sigma = 0.1$	20	2.038	1.964
$v^2 = 0.006$	25	6.659	6.659
	30	11.415	11.415
	15	1.272	0.841
$\sigma = 0.393$	20	3.866	3.320
$v^2 = 0.098$	25	7.543	7.131
	30	11.805	11.561

Other Parameter Values: r = 0.05, Strike Price = 18, k = 0.5, Maturity of One year

Option Parameters	Futures Price	Black-Scholes	One-factor Model
	15	0.002	0.001
$\sigma = 0.1$	20	1.990	1.973
$v^2 = 0.004$	25	6.827	6.827
	30	11.704	11.704
	15	0.681	0.525
$\sigma = 0.393$	20	3.167	2.956
$v^2 = 0.061$	25	7.160	7.043
	30	11.785	11.742

Other Parameter Values: r = 0.05, Strike Price = 18, k = 0.5, Maturity of Six months

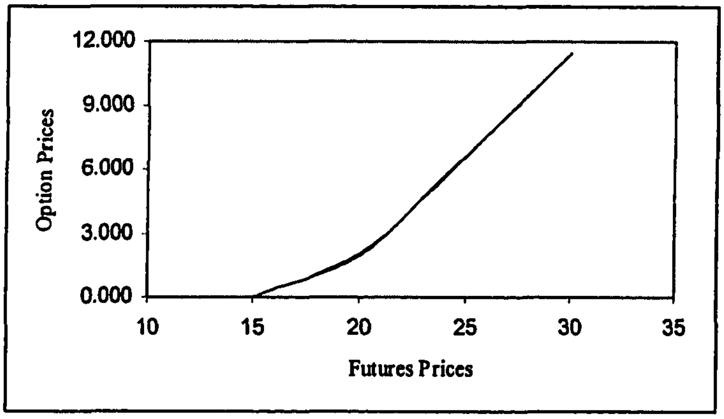
Option Parameters	Futures Price	Black-Scholes	One-factor Model
	15	0.000	0.000
$\sigma = 0.1$	20	1.992	1.992
$v^2 = 0.001$	25	6.971	6.971
	30	11.950	11.950
	15	0.042	0.038
$\sigma = 0.393$	20	2.196	2.184
$v^2 = 0.012$	25	6.972	6.972
	30	11.950	11.950

Other Parameter Values: r = 0.05, Strike Price = 18, k = 0.5, Maturity of One month

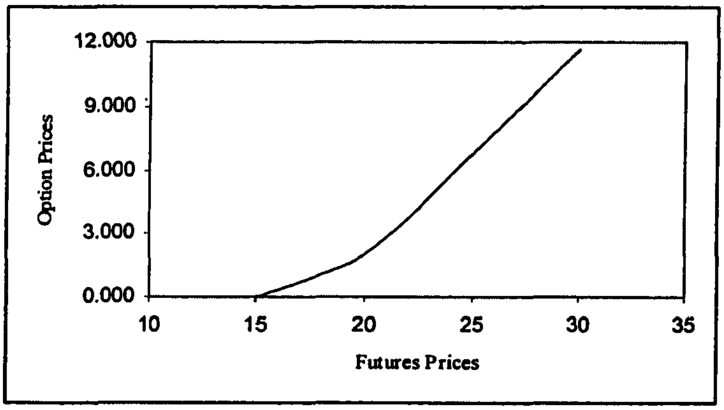
Figure 14

These graphs illustrate the difference between the Black-Scholes model and the one-factor model across maturities when $k = 0.5$, $\sigma = 0.1$ $r = 0.05$ and $X = 18$.

A. One year to maturity with $v^2 = 0.006$



B. Six months to maturity with $v^2 = 0.004$



C. One month to maturity with $v^2 = 0.001$

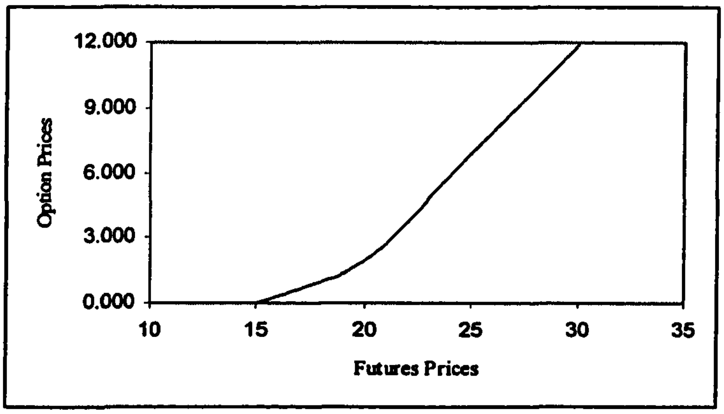
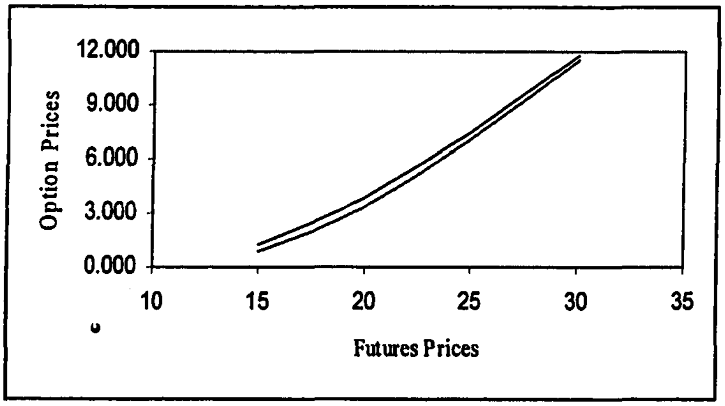


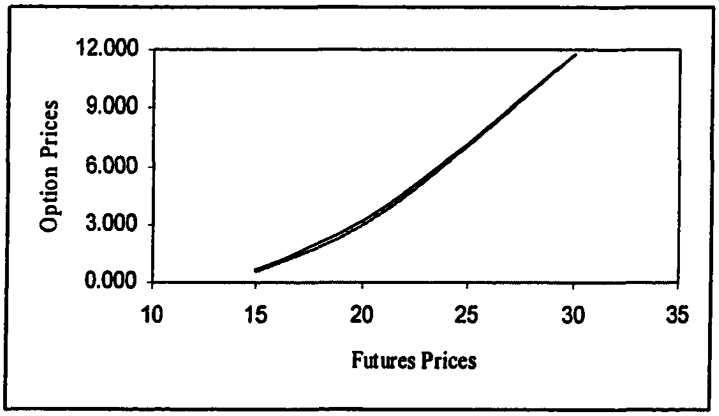
Figure 15

These graphs illustrate the difference between the Black-Scholes model and the one-factor model across maturities when $k = 0.5$, $\sigma = 0.393$, $r = 0.05$ and $X = 18$.

A. One year to maturity with $v^2 = 0.098$



B. Six months to maturity with $v^2 = 0.061$



C. One month to maturity with $v^2 = 0.012$

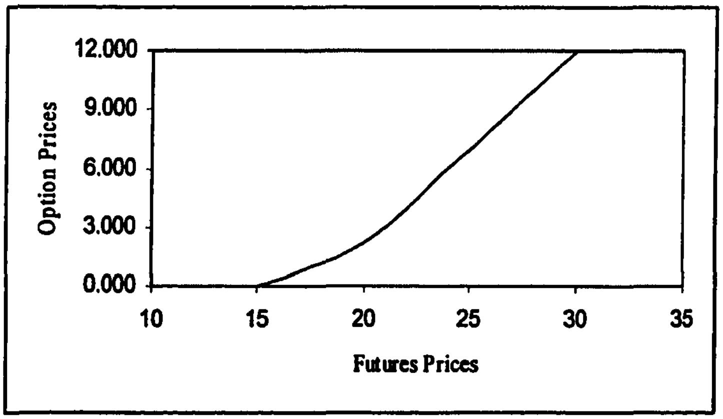


Table 15

This table shows the theoretical option prices for the one-factor futures model.

Option Parameters	Futures Price	Black-Scholes	One-factor Model
	15	0.021	0.001
$\sigma = 0.1$	20	2.038	1.930
$v^2 = 0.004$	25	6.659	6.659
	30	11.415	11.415
	15	1.272	0.569
$\sigma = 0.393$	20	3.866	2.961
$v^2 = 0.067$	25	7.543	6.909
	30	11.805	11.466

Other Parameter Values: r = 0.05, Strike Price = 18, k = 1.0, Maturity of One year

Option Parameters	Futures Price	Black-Scholes	One-factor Model
	15	0.002	0.000
$\sigma = 0.1$	20	1.990	1.963
$v^2 = 0.003$	25	6.827	6.827
	30	11.704	11.704
	15	0.681	0.406
$\sigma = 0.393$	20	3.167	2.786
$v^2 = 0.049$	25	7.160	6.965
	30	11.785	11.721

Other Parameter Values: r = 0.05, Strike Price = 18, k = 1.0, Maturity of Six months

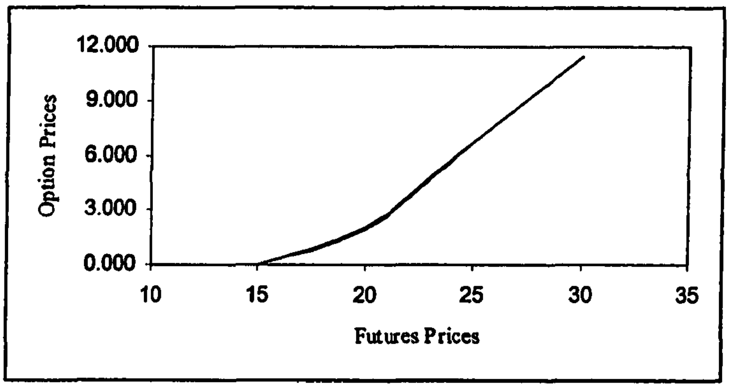
Option Parameters	Futures Price	Black-Scholes	One-factor Model
	15	0.000	0.000
$\sigma = 0.1$	20	1.992	1.992
$v^2 = 0.001$	25	6.971	6.971
	30	11.950	11.950
	15	0.042	0.035
$\sigma = 0.393$	20	2.196	2.174
$v^2 = 0.012$	25	6.972	6.972
	30	11.950	11.950

Other Parameter Values: r = 0.05, Strike Price = 18, k = 1.0, Maturity of One month

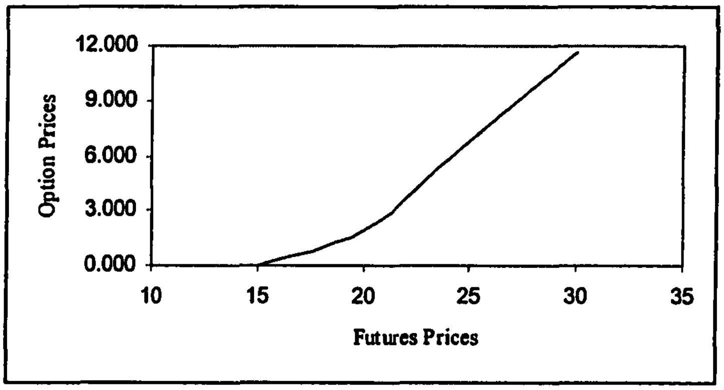
Figure 16

These graphs illustrate the difference between the Black-Scholes model and the one-factor model across maturities when $k = 1.0$, $\sigma = 0.1$, $r = 0.05$ and $X = 18$ $v^2 = 0.006$.

A. One year to maturity with $v^2 = 0.004$



B. Six months to maturity with $v^2 = 0.003$



C. One month to maturity with $v^2 = 0.001$

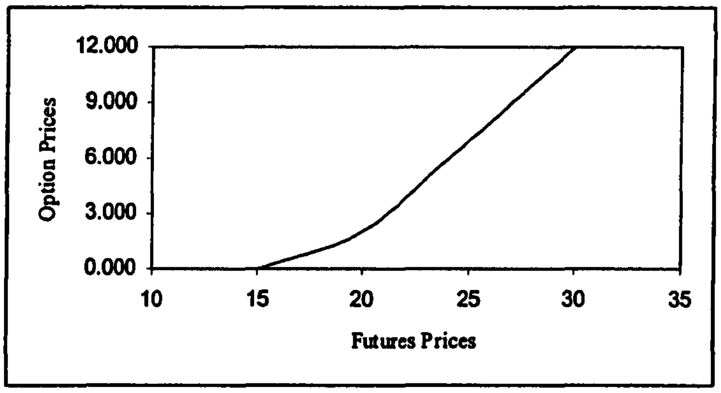
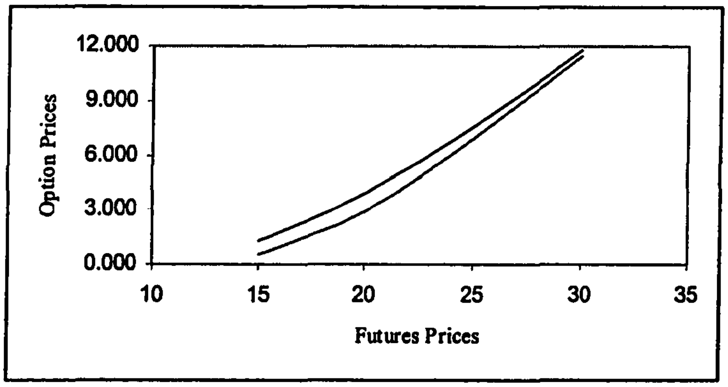


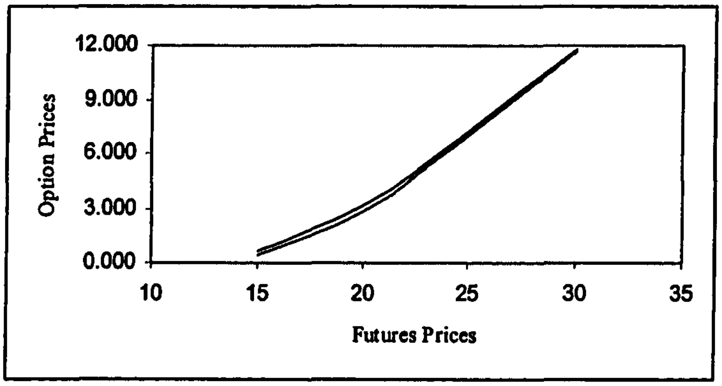
Figure 17

These graphs illustrate the difference between the Black-Scholes model and the one-factor model across maturities when $k = 1.0$, $\sigma = 0.393$, $r = 0.05$ and $X = 18$.

A. One year to maturity with $v^2 = 0.067$



B. Six months to maturity with $v^2 = 0.049$



C. One month to maturity with $v^2 = 0.012$

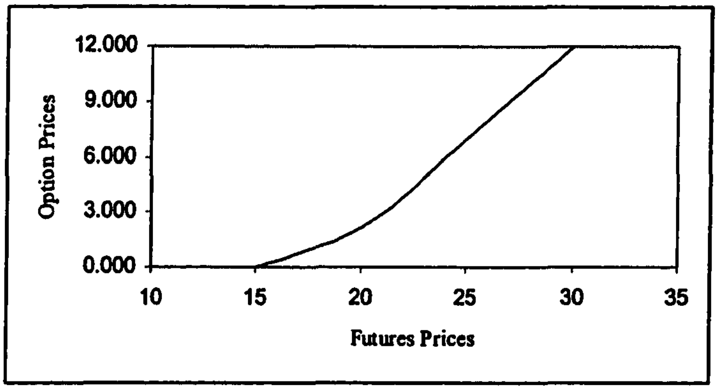


Table 16

This table shows the theoretical option prices for the one-factor futures model.

Option Parameters	Futures Price	Black-Scholes	One-factor Model
	15	0.021	0.000
$\sigma = 0.1$	20	2.038	1.903
$v^2 = 0.001$	25	6.659	6.659
	30	11.415	11.415
	15	1.272	0.061
$\sigma = 0.393$	20	3.866	2.150
$v^2 = 0.0154$	25	7.543	6.662
	30	11.805	11.415

Other Parameter Values: r = 0.05, Strike Price = 18, k = 5.0, Maturity of One year

Option Parameters	Futures Price	Black-Scholes	One-factor Model
	15	0.002	0.000
$\sigma = 0.1$	20	1.990	1.951
$v^2 = 0.0009$	25	6.827	6.827
	30	11.704	11.704
	15	0.681	0.062
$\sigma = 0.393$	20	3.167	2.202
$v^2 = 0.0153$	25	7.160	6.830
	30	11.785	11.704

Other Parameter Values: r = 0.05, Strike Price = 18, k = 5.0, Maturity of Six months

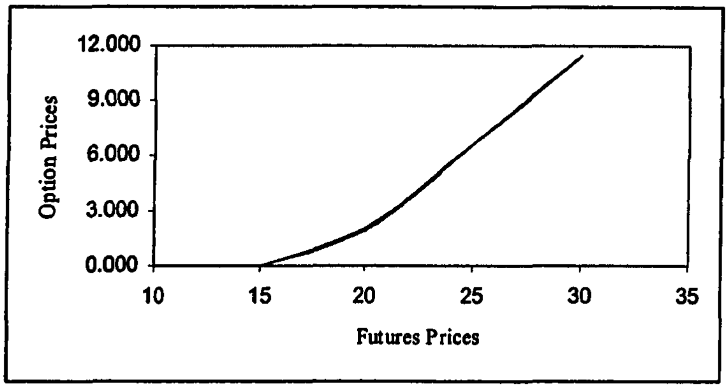
Option Parameters	Futures Price	Black-Scholes	One-factor Model
	15	0.000	0.000
$\sigma = 0.1$	20	1.992	1.992
$v^2 = 0.0005$	25	6.971	6.971
	30	11.950	11.950
	15	0.042	0.015
$\sigma = 0.393$	20	2.196	2.106
$v^2 = 0.008$	25	6.972	6.971
	30	11.950	11.950

Other Parameter Values: r = 0.05, Strike Price = 18, k = 5.0, Maturity of One month

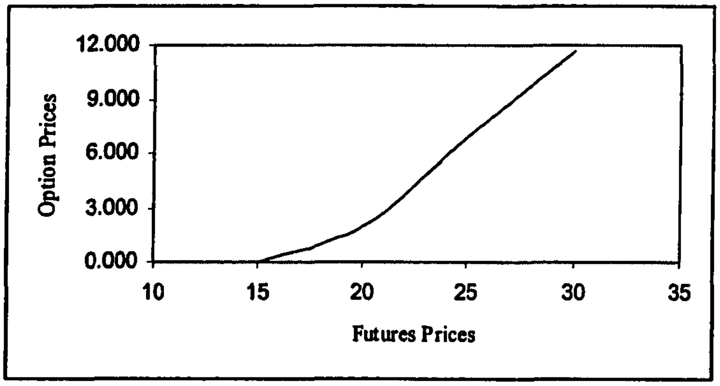
Figure 18

These graphs illustrate the difference between the Black-Scholes model and the one-factor model across maturities when $k = 5.0$, $\sigma = 0.1$, $r = 0.05$ and $X = 18$ $v^2 = 0.006$.

A. One year to maturity with $v^2 = 0.001$



B. Six months to maturity with $v^2 = 0.0009$



C. One month to maturity with $v^2 = 0.0005$

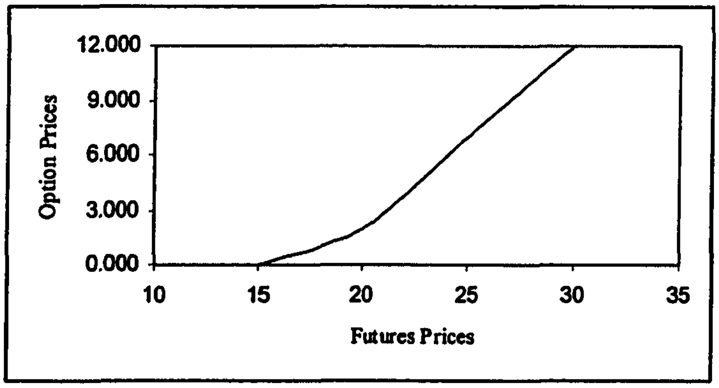
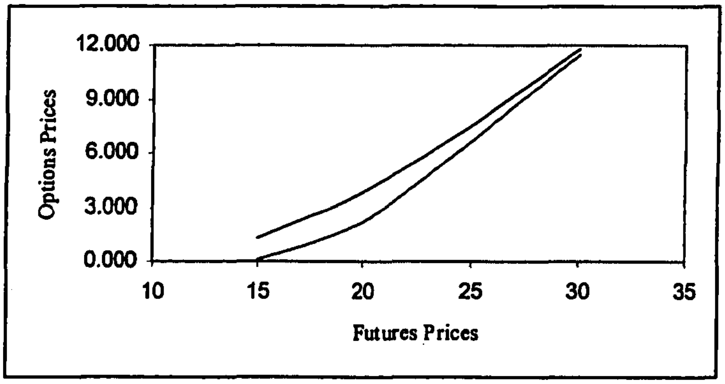


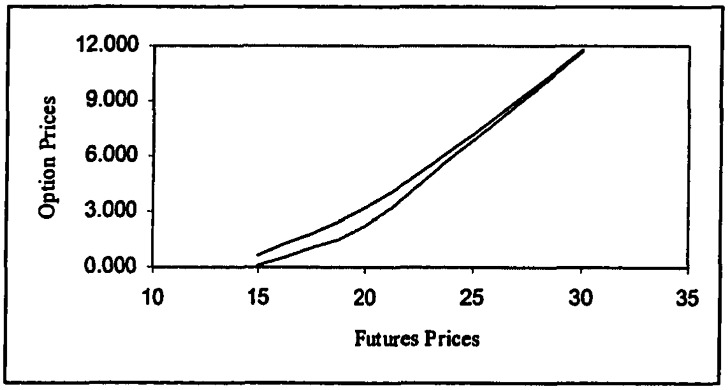
Figure 19

These graphs illustrate the difference between the Black-Scholes model and the one-factor model across maturities when $k = 5.0$, $\sigma = 0.393$, $r = 0.05$ and $X = 18$.

A. One year to maturity with $v^2 = 0.0154$



B. Six months to maturity with $v^2 = 0.0153$



C. One month to maturity with $v^2 = 0.008$

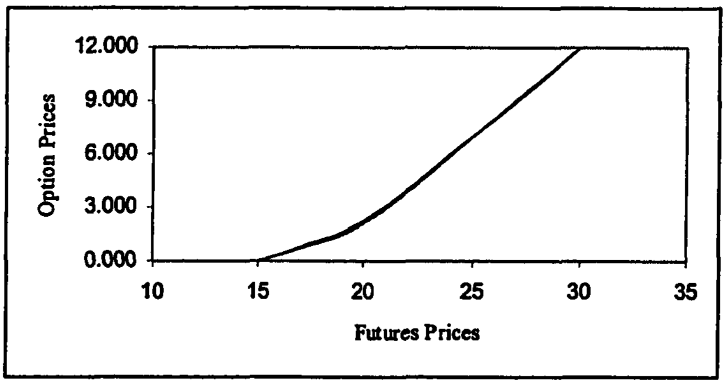


Table 17

This table shows the theoretical option prices for the two-factor futures model.

Option Parameters	Futures Price	Two-factor Model
$\sigma_s = 0.393$	15	1.289
$\sigma_c = 0.1$	20	3.886
$v^2 = 0.157$	25	7.559
$\rho_{sc} = 0.0$	30	11.816
$\sigma_s = 0.393$	15	1.104
$\sigma_c = 0.1$	20	3.656
$v^2 = 0.131$	25	7.376
$\rho_{sc} = 0.766$	30	11.698

Other Parameter Values: r = 0.05, Strike Price = 18, k = 0.5, Maturity of One year

Option Parameters	Futures Price	Two-factor Model
$\sigma_s = 0.393$	15	0.685
$\sigma_c = 0.1$	20	3.172
$v^2 = 0.078$	25	7.162
$\rho_{sc} = 0.0$	30	11.786
$\sigma_s = 0.393$	15	0.620
$\sigma_c = 0.1$	20	3.085
$v^2 = 0.071$	25	7.112
$\rho_{sc} = 0.766$	30	11.766

Other Parameter Values: r = 0.05, Strike Price = 18, k = 0.5, Maturity of Six months

Option Parameters	Futures Price	Two-factor Model
$\sigma_s = 0.393$	15	0.042
$\sigma_c = 0.1$	20	2.196
$v^2 = 0.012$	25	6.972
$\rho_{sc} = 0.0$	30	11.950
$\sigma_s = 0.393$	15	0.041
$\sigma_c = 0.1$	20	2.191
$v^2 = 0.013$	25	6.972
$\rho_{sc} = 0.766$	30	11.950

Other Parameter Values: r = 0.05, Strike Price = 18, k = 0.5, Maturity of One month

Table 18

This table shows the theoretical option prices for the two-factor futures model.

Option Parameters	Futures Price	Two-factor Model
$\sigma_s = 0.393$	15	1.284
$\sigma_c = 0.1,$	20	3.880
$v^2 = 0.156$	25	7.555
$\rho_{sc} = 0.0$	30	11.813
$\sigma_s = 0.393$	15	1.125
$\sigma_c = 0.1,$	20	3.682
$v^2 = 0.134$	25	7.396
$\rho_{sc} = 0.766$	30	11.711

Other Parameter Values: r = 0.05, Strike Price = 18, k = 1.0, Maturity of One year

Option Parameters	Futures Price	Two-factor Model
$\sigma_s = 0.393$	15	0.684
$\sigma_c = 0.1,$	20	3.171
$v^2 = 0.078$	25	7.162
$\rho_{sc} = 0.0$	30	11.786
$\sigma_s = 0.393$	15	0.624
$\sigma_c = 0.1,$	20	3.091
$v^2 = 0.071$	25	7.115
$\rho_{sc} = 0.766$	30	11.767

Other Parameter Values: r = 0.05, Strike Price = 18, k = 1.0, Maturity of Six months

Option Parameters	Futures Price	Two-factor Model
$\sigma_s = 0.393$	15	0.042
$\sigma_c = 0.1,$	20	2.196
$v^2 = 0.013$	25	6.972
$\rho_{sc} = 0.0$	30	11.950
$\sigma_s = 0.393$	15	0.041
$\sigma_c = 0.1,$	20	2.191
$v^2 = 0.013$	25	6.972
$\rho_{sc} = 0.766$	30	11.950

Other Parameter Values: r = 0.05, Strike Price = 18, k = 1.0, Maturity of One month

Table 19

This table shows the theoretical option prices for the two-factor futures model.

Option Parameters	Futures Price	Two-factor Model
$\sigma_s = 0.393$	15	1.274
$\sigma_c = 0.1$	20	3.868
$v^2 = 0.155$	25	7.545
$\rho_{sc} = 0.0$	30	11.807
$\sigma_s = 0.393$	15	1.206
$\sigma_c = 0.1$	20	3.783
$v^2 = 0.145$	25	7.476
$\rho_{sc} = 0.766$	30	11.761

Other Parameter Values: r = 0.05, Strike Price = 18, k = 5.0, Maturity of One year

Option Parameters	Futures Price	Two-factor Model
$\sigma_s = 0.393$	15	0.682
$\sigma_c = 0.1$	20	3.169
$v^2 = 0.077$	25	7.160
$\rho_{sc} = 0.0$	30	11.785
$\sigma_s = 0.393$	15	0.647
$\sigma_c = 0.1$	20	3.121
$v^2 = 0.074$	25	7.133
$\rho_{sc} = 0.766$	30	11.774

Other Parameter Values: r = 0.05, Strike Price = 18, k = 5.0, Maturity of Six months

Option Parameters	Futures Price	Two-factor Model
$\sigma_s = 0.393$	15	0.0424
$\sigma_c = 0.1$	20	2.196
$v^2 = 0.013$	25	6.972
$\rho_{sc} = 0.0$	30	11.950
$\sigma_s = 0.393$	15	0.0401
$\sigma_c = 0.1$	20	2.192
$v^2 = 0.013$	25	6.972
$\rho_{sc} = 0.766$	30	11.950

Other Parameter Values: r = 0.05, Strike Price = 18, k = 5.0, Maturity of One month

Chapter 6

Conclusion

It has been noted that over the last few years energy markets have exhibited turbulent behavior. In particular, Crude oil prices tripled, natural gas prices increased three hundred and fifty percent, and heating oil once sold for over $100 per barrel. These volatile market conditions have caused an increased interest in the use of financial instruments, such as futures, and options on futures. These instruments are the main vehicles for hedging price risk, and the ability to price these instruments is now an increasingly important problem in financial economics.

In this dissertation, we develop models of the stochastic behavior of commodity prices that take into account mean reversion, in terms of their ability to price commodity contingent claims. These pricing equations are found under a risk neutral economy using equivalent martingale measures. Our analysis begins with the derivation of three closed form solutions for futures/forward contracts. In all three models, the logarithm of the futures/forward price is linear in the state variables. Moreover, the difference between the models depends on the volatility term. The one-factor model implies that the volatility of futures prices will converge to a value of zero and the futures price will converge to a fixed value as maturity increases. The two- and three-factor models, however, imply that the futures price volatility will decrease but converge to a fixed value different from zero, while the futures price converges to some fixed rate of growth.

After developing solutions for future/forward contracts, it is shown that each of these futures contracts have an implicit price dynamic. Development of these price dynamics reveal that futures prices follow a martingale process with an exponentially dampened volatility. The stochastic differential for the futures price is then used to price options written on the underlying futures contract. These pricing formulas incorporate discounts in the option prices due to the term structure of volatility implicit in the futures price. The discounts applied to the one-factor option prices are greater than the discounts for the two-factor model, since the one factor futures price volatility converges to zero and not some finite number.

To conclude our analysis we consider the option pricing model for jump-diffusion developed by Hilliard and Reis (1998). Our analysis indicates that the option pricing model posited by Hilliard and Reis is founded on mathematical irregularities. To begin, the authors assert that the form of the futures price is invariant to the presence of a jump process. This is incorrect in that the presence of a jump process in the futures price dynamics transforms the distribution of futures prices into a mixed Poisson Gaussian distribution. This is not the same as a standard normal distribution that is used when no jump process is present. Furthermore, Hilliard and Reis state that the jump process in the spot price is the same as the jump process in the futures price. The jump process for the futures price cannot be the same as the spot price since the futures price is a function of the spot price. Lastly, Hilliard and Reis unilaterally impose the general equilibrium results from Bates' (1988) model in their analysis. This is a problem since Bates does not consider a model that includes a mean reverting convenience yield or a stochastic interest. In short, Hilliard and Reis' jump diffusion option model is built on unstable foundations and the results from their analysis are tenuous at best.

References

Anderson, B. D. O. and J. B. Moore (1979) *Optimal Filtering*, Englewood Cliffs. NJ: Prentice Hall.

Bates, D. S. (1988) "Pricing Options Under Jump-Diffusion Processes," Wharton School Working paper.

Bates, D. S. (1991) "The Crash of '87: Was It Expected? The Evidence from Options Markets," *Journal of Finance* 66:3, 1009-1044.

Bjerksund, P. (1991) "Contingent Claim Evaluation when the Convenience Yield is Stochastic: Analytical Results," *Working Paper.*

Bessembinder, H. J. F. Coughenour, P. J. Seguin, and M. M. Smoller (1995) "Mean Reversion in Equilibrium Asset Prices: Evidence from the Futures Term Structure," *Journal of Finance* 50:1, 361-375.

Black, F. (1976) "The Pricing of Commodity Contracts," *Journal of Financial Economics* 3, 167-179

Black, F. and M. Scholes (1973) "The Pricing of Options and Corporate Liabilities," *Journal of Political Economy* 81: 3, 637-654.

Brennan, M. J. (1958) "The Theory of Storage," *American Economic Review* 48:1, 50-72.

Brennan, M. J. (1991) "The Price of Convenience and the Valuation of Commodity Contingent Claims," in D. Lund and B Oksendal, Eds.: *Stochastic Models and Option Values*, North Holland.

Brennan, M. J. and E. S. Schwartz (1979) "A Continuous Time Approach to the Pricing of Bonds," *Journal of Banking and Finance* 3, 133-155.

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* 17:3, 301-329.

Brennan, M. J. and E. S. Schwartz (1985) "Evaluating Natural Resource Investments," *Journal of Business* 58:2, 135-157.

Cootner, P. H., 1964, *The Random Character of Stock Markets* M.I.T. Press. Cambridge, Mass.

Cortazar, G. and E. S. Schwartz (1994) "The Evaluation of Commodity Contingent Claims," *Journal of Derivatives* 1, 27-39.

Cox, J. C., J. E. Ingersoll, and S. A. Ross, (1981) "The Relation between Forwards and Futures Prices," *Journal of Financial Economics* 9 321-346.

Cox, J. C., J. E. Ingersoll, and S. A. Ross, (1985a) "An Intertemporal General Equilibrium Model of Asset Prices," *Econometrica* 53:2, 363-384.

Cox, J. C., J. E. Ingersoll, and S. A. Ross, (1985b) "A Theory of Term Structure of Interest Rates," *Econometrica* 53:2, 385-407.

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

Fama, E. and K. French (1987) "Commodity Futures Prices: Some Evidence on Forecast Power, Premiums, and the Theory of Storage," *Journal of Business* 60, 55-73.

Fama, E. and K. French (1988) "Business Cycles and the Behavior of Metals Prices," *Journal of Finance* 43, 1075-1094.

Gibson, R. and E. S. Schwartz (1990) "Stochastic Convenience Yield and the Pricing of Oil Contingent Claims," *Journal of Finance* 45:3, 959-976.

Hamiltion, J., (1991) *Time Series Analysis*, Princeton NJ: Princeton University Press.

Harrison, M. and D. Kreps (1979) "Martingales and Arbitrage in Multiperiod Securities Markets," *Journal of Economic Theory* 20, 381-408.

Harrison, M. and S. Pliska (1981) "Martingales and Stochastic Integrals in Theory of Continuous Trading," *Stochastic Processes and Their Applications* 11, 313-316.

Harvey, A. (1989) *Forecasting, Structural Time-Series Models and the Kalman Filter*, New York: Cambridge University Press.

Heath, D. R. Jarrow and A. Morton (1992) "Bond Pricing and the Term Structure of Interest Rates: A New Methodology for Contingent Claim Valuation," *Econometrica* 60, 77-105.

Hillard, J. E. and J. Reis (1998) "Valuation of Commodity Futures and Options under Stochastic Convenience Yields, Interest Rates, and Jump Diffusion in the Spot Price," *Journal of Financial and Quantitative Analysis* 33:1, 61-86.

Hoel, P. G. S. C. Port, and C. J. Stone (1972) *Introduction to Stochastic Processes*, Houghton Mifflin Company.

Hull, J. (2000), *Options, and Other Derivatives*, Fourth Edition Prentice Hall.

Jarrow, R. A. and G. S. Oldfield (1981) "Forward Contracts and Futures Contracts," *Journal of Financial Economics* 9, 373-382.

Kushner, A. J. (1995) *Numerical Methods for Stochastic Control Problems in Continuous Time*, New York Springer-Verlag.

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

Merton, R. C. (1973a) "Appendix: Continuous-Time Speculative Processes," *SIAM Review* 15, 34-38.

Merton, R. C. (1973b) "Theory of Rational Option Pricing," *Bell Journal of Economics and Management Science* 4, 141-183.

Merton, R. C. (1976) "Option Pricing when the Underlying Stock Returns are Discontinuous," *Journal of Financial Economics* 3, 125-44.

Merton, R. C. (1990) *Continuous Time Finance*, Blackwell Publishers.

Merton, R. C. and P. A. Samuelson (1974) "Fallacy of Log-Normal Approximation of Optimal Portfolio Decision Making over Many Periods," *Journal Financial Economics* 1, 67-97

Miltersen K. R. and E. S. Schwartz, (1998) "Pricing of Options on Commodity Futures with Stochastic Term Structures of Convenience Yields and Interest Rates," *Journal of Financial and Quantitative Analysis* 33:1, 33-59.

Neftci, S. N. (2000) *An Introduction to the Mathematics of Financial Derivatives*, Academic Press.

Press, J. S. (1967) "A compound Events Model for Security Prices," *Journal of Business* 40, 317-335.

Rendleman, K. and C. Carabini (1979) "The Valuation of Options on Futures Contracts," *Journal of Finance*, December, 1319-1340.

Samuelson, P.A. (1965a) "Rational Theory of Warrant Pricing," *Industrial Management Review* 6, 13-31.

Samuelson, P.A. (1965b) "Proof that Properly Anticipated Prices Fluctuate Randomly," *Industrial Management Review* 6, 41-49.

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

Schwartz, E. S (1982) "The Pricing of Commodity-Linked Bonds," *Journal of Finance* 37:2, 525-539.

Schwartz, E. S (1997) "The Stochastic Behavior of Commodity Prices: Implications for Valuation and Hedging," *Journal of Finance* 52:3, 923-973.

Telser, L. G. (1958) "Futures Trading and the Storage of Cotton and Wheat," *Journal of Political Economy* 66, 233-244.

Working, H. (1948) "Theory of Inverse Carry Charge in Futures Markets," *Journal of Farm Economics* 30, 1-28.

Appendix A

Dynamic Factor Analysis

This appendix is composed of two sections, the first describes the Kalman filter and the second describes the expectations maximization (EM) algorithm. The EM algorithm utilizes the Kalman filter in the estimation procedure, and is therefore, discussed first.

A.1 Kalman filter

Introduction

In 1960, R. E. Kalman provided an alternative method of formulating the least squares filtering problem using state-space methods. Engineers, especially in the field of navigation, were quick to see the Kalman technique as a practical solution to a number of problems that were previously considered intractable using the Wiener methods. The Wiener solution of the optimal filter problem is a filter weighting solution. In effect this tells us how the past values of the input should be weighted in order to determine the present value of the output, that is, the optimal estimate. Unfortunately, the Weiner solution does not lend itself very well to problems with large amounts of data. The two main features of the Kalman formulation and solution of the problem are (1) vector modeling of the random processes under consideration and (2) recursive processing of the noisy measurement (input) data.

When working with practical problems involving discrete data, it is important that our methods be computationally feasible as well as mathematically correct. A simple example will illustrate this. Consider the problem of estimating the mean of some unknown constant based on sequence of noisy measurements. Assume that our estimate is to be the sample mean and that we wish to refine our estimate with each new measurement as it becomes available. That is, think of processing the data on-line. Let the measurement sequence be denoted as $z_1, z_2, \ldots\ldots, z_n$ where the subscript denotes the time at which the measurement is taken. One method of processing the data would be to store each measurement as it becomes available and then compute the sample mean in accordance with the Weiner algorithm (in words):

1. First measurement z_1: Store z_1 and estimate the mean as:

$$\hat{m}_1 = z_1$$

2. Second measurement z_2: Store z_2 along with z_1 and estimate the mean as

$$\hat{m}_2 = \frac{z_1 + z_2}{2}$$

3. Third measurement z_3 : Store z_3 along with z_1 and z_2 estimate the mean as

$$\hat{m}_3 = \frac{z_1 + z_2 + z_3}{3}$$

4. And so forth.

Clearly, this would yield the correct sequence of sample means as the experiment progresses. It should be clear that the amount of memory needed to store the measurement keeps increasing with time, and also the number of arithmetic operations needed to form the estimate increases correspondingly. This would lead to obvious problems when the total amount of data is large. Thus, consider a simple variation in the computational procedure in which each new estimate is formed as a blend of the old estimate and the current measurement. To be specific, consider the following algorithm:

1. First measurement z_1: Store z_1 and estimate the mean as:

$$\hat{m}_1 = z_1$$

store $\hat{m}_1$ and discard z_1.

2. Second measurement z_2: Compute the estimate as a weighted sum of the previous estimate $\hat{m}_1$ and the current measurement z_2:

$$\hat{m}_2 = \frac{1}{2}\hat{m}_1 + \frac{1}{2}z_2$$

store $\hat{m}_2$ and discard z_2.

3. Third measurement z_3: Compute the estimate as a weighted sum of $\hat{m}_2$ and z_3:

$$\hat{m}_3 = \frac{2}{3}\hat{m}_2 + \frac{1}{3}z_3$$

store $\hat{m}_3$ and discard z_3.

4. And so forth. It should be obvious that at the nth stage the weighted sum is

$$\hat{m}_n = \frac{(n-1)}{n}\hat{m}_{n-1} + \frac{1}{n}z_n \quad .$$

Clearly, the above procedure yields the same identical sequence of estimates as before, but without the need to store all the previous measurements. I simply use the result of the previous step to help obtain the estimate at the current step of the process. In this way, the previous computational effort is used to good advantage and not wasted. The second algorithm can proceed ad infinitum without a growing memory problem. Eventually, of course, as n becomes extremely large, a round-off problem might be encountered. However, this is to be expected with either of the two algorithms.

The second algorithm is a simple example of a recursive mode of operation. The key element in any recursive procedure is the use of the results of the previous step to aid in obtaining the desire result for the current step. This is one of the main features of Kalman filtering, and one that clearly distinguishes it from the weighting-function (Weiner) approach, which requires arithmetic operations on all past data.

In order to apply the recursive philosophy of the Kalman filter to estimation of a stochastic process, it is first necessary that both the process and the measurement noise be modeled in state space form. The state space form is an enormously powerful tool, which opens the way to handling a wide variety of time series models. Once a model has been put in state space form, the Kalman filter may be applied. The general state space form applies to a multivariate time series, $\mathbf{y}_t$ containing N elements. This time series represents our observations of the world. These observable variables are related to an *mx1* vector, $\mathbf{x}_t$ known as the state vector, via a measurement equation.

$$\mathbf{y}_t = \mathbf{M}\mathbf{x}_t + \mathbf{v}_t \tag{A.1}$$

where $\mathbf{M}$ is an *Nxm* matrix and $\mathbf{v}_t$ is an *Nx1* vector of serially uncorrelated disturbances with mean zero and covariance matrix $\mathbf{R}_t$. That is,

$$E(\mathbf{v}_t) = 0 \text{ and } Var(\mathbf{v}_t) = \mathbf{R}_t \ . \tag{A.2}$$

The elements of the state vector $\mathbf{x}_t$ are not directly observable and are determined by the state or transition equations

$$\mathbf{x}_{t+1} = \mathbf{A}\mathbf{x}_t + \mathbf{D}\mathbf{w}_t \tag{A.3}$$

where $\mathbf{A}$ is an *mxm* transition matrix and $\mathbf{w}_t$ is an *mx1* vector of serially uncorrelated disturbances with mean zero and covariance matrix $\mathbf{Q}_t$. That is,

$$E(\mathbf{w}_t) = 0 \text{ and } Var(\mathbf{w}_t) = \mathbf{Q}_t . \tag{A.4}$$

The inclusion of the matrix $\mathbf{D}$ in front of the disturbance term, is to some extent, arbitrary. The disturbance term could always be redefined so as to have a covariance matrix $\mathbf{DQ_tD'}$. Nevertheless, the representation in (A.3) is often more natural when $\mathbf{v}_t$ is identified with a particular set of disturbances in the model.

The specification is completed by assuming that the initial vector $\mathbf{x_0}$ has mean μ and covariance matrix Σ, where

$$\Sigma = E[(\mathbf{x}_0 - \boldsymbol{\mu})(\mathbf{x}_0 - \boldsymbol{\mu})'] \tag{A.5}$$

Note that this state space configuration is different from the conventional autoregressive series because $\mathbf{x}_t$ is not observable. $\mathbf{v}_t$ and $\mathbf{w}_t$ can be thought of as observation and model noise respectively. **M, R, A, D** and **Q** are assumed to be non-stochastic.

In the trivial case when **M** = $\mathbf{I_n}$, then equation (A.4) reduces to $\mathbf{y}_t = \mathbf{x}_t + \mathbf{v}_t$. Under the assumption that $\mathbf{v}_t = \begin{bmatrix} 0 \\ . \\ 0 \end{bmatrix}$, $\mathbf{x}_t = \mathbf{M}^{-1}\mathbf{y}_t$ and is observable. When M is not the identity matrix, then the measurement equation for N = 2 and m = 3 is given by:

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$$

Expanding the first element of $\mathbf{y}_t$ yields,

$$y_1 = a_{11}x_1 + a_{12}x_2 + a_{13}x_{13} + v_1 \tag{A.6}$$

I can use these confounded observations on $\mathbf{y}_t$ to determine the mean of the three state variables, $\mathbf{x}_1$, $\mathbf{x}_2$ and $\mathbf{x}_3$. Determining the mean of each state variable at time *t* is the goal of the Kalman filter. The Kalman filter estimate is denoted by:

$$E(\mathbf{x}_1 \mid \mathbf{y}_0, \mathbf{y}_1, \ldots\ldots, \mathbf{y}_t) = \hat{\mathbf{x}}_{t,t}.$$

Deriving the Kalman filter depends upon several properties of conditional probabilities. A review of these properties follows. The results from this review will be utilized in the derivation of the Kalman filter.

Review of Conditional Probabilities

Suppose I have two jointly distributed random vectors **x** and **y**. For simplicity also assume the random variables in this section have zero means. The dimension of **x** is *3x1* and y is *2x1*. What is the "best" estimate of **x** in terms of y? The question is simple econometrics and is solved by:

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} \beta_1^1 & \beta_2^1 \\ \beta_1^2 & \beta_2^2 \\ \beta_1^3 & \beta_2^3 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

or $\hat{\mathbf{x}} = \mathbf{by}$. The variance of the estimation error is given by

$$E[(\mathbf{x} - \mathbf{by})'(\mathbf{x} - \mathbf{by})] \qquad \text{(A.7)}$$

which is a scalar. Because I want the "best" estimate of **x**, I will find the minimum of the estimation error by taking the derivative of equation (A.7) with respect to **b** and setting it equal to zero. This requires passing a derivative through the expectation operator. I know a derivative can pass through the expectation operator by the following example.

Assume $f(z) = 3z^2 \quad 0 \le z \le 1$. The expected value of *f*(*z*) is

$$E(z) = \int_0^1 z f(z) dz = \int_0^1 z(3z^2) dz = \frac{3}{4} z^4 \Big|_0^1 = \frac{3}{4}.$$

The derivative of the expected value is

$$\frac{\partial E(z)}{\partial z} = \frac{\partial \left[\frac{3}{4} z^4 \right]}{\partial z} = 3z^3 \Big|_0^1 = 3.$$

Therefore, taking the expected value and then the derivative of *f(z)* yields an answer of 3.

To verify that a derivative can pass through and expectation operator, I will reverse the order of the operations. I will calculate the derivative first and then integrate or apply the expectations operator. If I obtain the same result using both methods, then I am assured that the derivative can pass through an expectation operator.

$$\frac{\partial\left[\int z3z^2dz\right]}{\partial z}=\frac{\partial\left[\int 3z^3dz\right]}{\partial z}=\int 9z^2dz=3z^3\Big|_0^1=3$$

The result obtained using both methods is identical, and therefore, this example shows that the derivative can pass through the expectations operator. Now, I will return to the original goal of minimizing the estimation error so that I can obtain the best linear estimate for $\mathbf{x}$.

The variance of the estimation error in equation (A.7) can be rewritten as

$$E(\mathbf{x'x}+\mathbf{x'by}+\mathbf{y'b'x}+\mathbf{y'b'by}). \tag{A.8}$$

Minimizing with respect to $\mathbf{b}$ by taking the derivative of (A.8) and setting it equal to zero yields:

$$E(-2\mathbf{xy'}+2\mathbf{byy'})=\mathbf{0}.$$

Solving for $\mathbf{b}$ yields $\mathbf{b}=E(\mathbf{xy'})E(\mathbf{yy'})^{-1}=\Sigma_{xy}\Sigma_{yy}^{-1}$, where Σ_{xy} and Σ_{yy} are covariance matrices since $\mathbf{x}$ and $\mathbf{y}$ have zero means. Our best guess for $\mathbf{x}$ is

$$E(\mathbf{xy'})E(\mathbf{yy'})^{-1}\mathbf{y}=\mathbf{by}=\hat{\mathbf{x}} \tag{A.9}$$

or

$$\hat{\mathbf{x}}=\Sigma_{xy}\Sigma_{yy}^{-1}\mathbf{y}. \tag{A.10}$$

The estimation error is denoted $\tilde{\mathbf{x}}=\mathbf{x}-\hat{\mathbf{x}}$ and its variance-covariance is given by

$$E(\tilde{\mathbf{x}}\tilde{\mathbf{x}}') = E[(\mathbf{x} - \hat{\mathbf{x}})(\mathbf{x} - \hat{\mathbf{x}})']$$

which can be expanded as:

$$E[\mathbf{x}\mathbf{x}' - 2\mathbf{x}\hat{\mathbf{x}}' + \hat{\mathbf{x}}\hat{\mathbf{x}}'].$$

Utilizing (A.10), this expression for the variance of the estimation error can be rewritten as:

$$E[(\mathbf{x} - \boldsymbol{\Sigma}_{\mathbf{xy}}\boldsymbol{\Sigma}_{\mathbf{yy}}^{-1}\mathbf{y})(\mathbf{x} - \boldsymbol{\Sigma}_{\mathbf{xy}}\boldsymbol{\Sigma}_{\mathbf{yy}}^{-1}\mathbf{y})']$$

or

$$E[\mathbf{x}\mathbf{x}' + \boldsymbol{\Sigma}_{\mathbf{xy}}\boldsymbol{\Sigma}_{\mathbf{yy}}^{-1}\mathbf{y}\mathbf{y}'\boldsymbol{\Sigma}_{\mathbf{yy}}^{-1}\boldsymbol{\Sigma}_{\mathbf{xy}} - 2\boldsymbol{\Sigma}_{\mathbf{xy}}\boldsymbol{\Sigma}_{\mathbf{yy}}^{-1}\mathbf{y}\mathbf{x}'].$$

Earlier, I made the assumption that all random variables are distributed with zero mean. Therefore, $E(\mathbf{y}\mathbf{y}'\boldsymbol{\Sigma}_{\mathbf{yy}}^{-1}) = \mathbf{I}$ and $\boldsymbol{\Sigma}_{\mathbf{xy}} = E(\mathbf{y}\mathbf{x}')$. Using these facts and distributing the expectations operator reduces the above expression to:

$$E(\mathbf{x}\mathbf{x}') - \boldsymbol{\Sigma}_{\mathbf{xy}}\boldsymbol{\Sigma}_{\mathbf{yy}}^{-1}\boldsymbol{\Sigma}_{\mathbf{yx}} = E(\tilde{\mathbf{x}}\tilde{\mathbf{x}}')$$

or

$$\boldsymbol{\Sigma}_{\mathbf{xx}} - \boldsymbol{\Sigma}_{\mathbf{xy}}\boldsymbol{\Sigma}_{\mathbf{yy}}^{-1}\boldsymbol{\Sigma}_{\mathbf{yx}} = E(\tilde{\mathbf{x}}\tilde{\mathbf{x}}'). \tag{A.11}$$

The best linear estimate would ensure that the $Cov(\tilde{\mathbf{x}},\mathbf{y}) = \mathbf{0}$. Does my derived estimate exhibit this quality? Utilizing the fact that these random variables have zero mean, this covariance is given by:

$$Cov(\tilde{\mathbf{x}},\mathbf{y}) = \mathbf{E}(\mathbf{x} - \hat{\mathbf{x}},\mathbf{y}') = E(\mathbf{x}\mathbf{y}' - \boldsymbol{\Sigma}_{\mathbf{xy}}\boldsymbol{\Sigma}_{\mathbf{yy}}^{-1}\mathbf{y}\mathbf{y}') = \boldsymbol{\Sigma}_{\mathbf{xy}} - \boldsymbol{\Sigma}_{\mathbf{xy}} = 0.$$

Therefore, the least squares estimator is orthogonal to the estimation error.

Consider the case where **y** and **z** are uncorrelated, but they both impact **x**. The mean of **x** is given by $E(\mathbf{x} \mid \mathbf{y},\mathbf{z})$ which can decomposed into:

$$E(\mathbf{x} \mid \mathbf{y},\mathbf{z}) = E(\mathbf{x} \mid \mathbf{y}) + E(\mathbf{x} \mid \mathbf{z})$$

because y and z are uncorrelated. Let me illustrate the observation above by setting

$$\mathbf{w} = (\mathbf{y}',\mathbf{z}')'. \tag{A.12}$$

Applying (A.10) to determine $E(\mathbf{x} \mid \mathbf{w})$ yields

$$\hat{\mathbf{x}} = \boldsymbol{\Sigma}_{\mathbf{xw}}\boldsymbol{\Sigma}_{\mathbf{ww}}^{-1}\mathbf{w} \tag{A.13}$$

where

$$\boldsymbol{\Sigma}_{\mathbf{ww}} = \begin{bmatrix} \boldsymbol{\Sigma}_{\mathbf{yy}} & \boldsymbol{\Sigma}_{\mathbf{yz}} \\ \boldsymbol{\Sigma}_{\mathbf{zy}} & \boldsymbol{\Sigma}_{\mathbf{zz}} \end{bmatrix} = \begin{bmatrix} \boldsymbol{\Sigma}_{\mathbf{yy}} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Sigma}_{\mathbf{zz}} \end{bmatrix}$$

and

$$\boldsymbol{\Sigma}_{\mathbf{xw}} = E(\mathbf{xw}') = E[\mathbf{x}(\mathbf{y}',\mathbf{z}')] = E[\mathbf{xy}';\mathbf{xz}'] = [\boldsymbol{\Sigma}_{\mathbf{xy}} : \boldsymbol{\Sigma}_{\mathbf{xz}}]. \tag{A.14}$$

Note that the cross products are zero because I have defined y and z to be uncorrelated.

Using (A.12), (A.13) and (A.14) the best estimate for x can be written in the spirit of my prior derivation as:

$$\hat{\mathbf{x}} = [\boldsymbol{\Sigma}_{\mathbf{xy}} : \boldsymbol{\Sigma}_{\mathbf{xz}}] \begin{bmatrix} \boldsymbol{\Sigma}_{\mathbf{yy}} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Sigma}_{\mathbf{zz}} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{y} \\ \mathbf{z} \end{bmatrix}$$

or

$$\hat{\mathbf{x}} = \boldsymbol{\Sigma}_{\mathbf{xy}}\boldsymbol{\Sigma}_{\mathbf{yy}}^{-1}\mathbf{y} + \boldsymbol{\Sigma}_{\mathbf{xz}}\boldsymbol{\Sigma}_{\mathbf{zz}}^{-1}\mathbf{z}.$$

The equation above illustrates that $E(\mathbf{x} \mid \mathbf{y},\mathbf{z}) = E(\mathbf{x} \mid \mathbf{y}) + E(\mathbf{x} \mid \mathbf{z})$ when y and z are uncorrelated.

The last item to consider before deriving the Kalman filter is the observation that

$$E(\mathbf{y} \mid \mathbf{x}_1,\mathbf{x}_2) = E(\mathbf{y} \mid \tilde{\mathbf{x}}_1,\mathbf{x}_2)$$

where $\mathbf{x}_1$ has been orthogonalized with respect to $\mathbf{x}_2$ to yield $\tilde{\mathbf{x}}_1$. The first half of establishing the observation made above is to recall that from ordinary least squares (OLS) that our best guess at $E(\mathbf{y} \mid \mathbf{x}_1, \mathbf{x}_2)$ is given by:

$$\frac{\sum x_2^2 \sum x_1 y - \sum x_1 x_2 \sum x_2 y}{\sum x_1^2 \sum x_2^2 - \left(\sum x_1 x_2\right)^2}\mathbf{x}_1 + \frac{\sum x_1^2 \sum x_2 y - \sum x_1 x_2 \sum x_1 y}{\sum x_1^2 \sum x_2^2 - \left(\sum x_1 x_2\right)^2}\mathbf{x}_2 \tag{A.15}$$

and establishing the second half of the observation is a matter of showing that the mathematical expression that estimates $E(\mathbf{y} \mid \mathbf{x}_1, \mathbf{x}_2)$ also estimates $E(\mathbf{y} \mid \tilde{\mathbf{x}}_1, \mathbf{x}_2)$.

Assume $\mathbf{y} = \boldsymbol{\beta}_1 \mathbf{x}_1 + \boldsymbol{\beta}_2 \mathbf{x}_2 + \mathbf{e}$ where $\mathbf{x}_1$ and $\mathbf{x}_2$ are uncorrelated. I first obtain $\hat{\mathbf{y}}$ under these conditions and then move on to the more realistic case where $\mathbf{x}_1$ and $\mathbf{x}_2$ may be correlated. Recall that the vector x is in deviation form. The regression coefficients for this model are estimated by:

$$\begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \end{bmatrix} = (\mathbf{x}'\mathbf{x})^{-1}\mathbf{x}'\mathbf{y} = \frac{1}{\sum x_1^2 \sum x_2^2 - \left(\sum x_1 x_2\right)^2} \begin{bmatrix} \sum x_2^2 & -\sum x_1 x_2 \\ -\sum x_1 x_2 & \sum x_1^2 \end{bmatrix} \begin{bmatrix} \sum x_1 y \\ \sum x_2 y \end{bmatrix}$$

and since $\sum x_1 x_2 = 0$, $\hat{\beta}_1$ and $\hat{\beta}_2$ reduce to:

$$\hat{\beta}_1 = \frac{\sum x_2^2 \sum x_1 y}{\sum x_2^2 \sum x_1^2} = \frac{\sum x_1 y}{\sum x_1^2}$$

$$\hat{\beta}_2 = \frac{\sum x_1^2 \sum x_2 y}{\sum x_2^2 \sum x_1^2} = \frac{\sum x_2 y}{\sum x_2^2}$$

Therefore the estimate of the mean of y conditional on $\mathbf{x}_1$ and $\mathbf{x}_2$ is simply:

$$\hat{\mathbf{y}} = \hat{\beta}_1 \mathbf{x}_1 + \hat{\beta}_2 \mathbf{x}_2 = \frac{\sum x_1 y}{\sum x_1^2}\mathbf{x}_1 + \frac{\sum x_2 y}{\sum x_2^2}\mathbf{x}_2 \tag{A.16}$$

which corresponds to my best guess for the quantity $\{E(\mathbf{y} \mid \mathbf{x}_1) + E(\mathbf{y} \mid \mathbf{x}_2)\}$.

When the constraint that $\mathbf{x}_1$ and $\mathbf{x}_2$ are uncorrelated is relaxed, the estimator for β_1 and β_2 are given by:

$$\begin{bmatrix}\hat{\beta}_1 \\ \hat{\beta}_2\end{bmatrix} = (\mathbf{x}'\mathbf{x})^{-1}\mathbf{x}'\mathbf{y} = \frac{1}{\sum x_1^2 \sum x_2^2 - \left(\sum x_1 x_2\right)^2}\begin{bmatrix}\sum x_2^2 & -\sum x_1 x_2 \\ -\sum x_1 x_2 & \sum x_1^2\end{bmatrix}\begin{bmatrix}\sum x_1 y \\ \sum x_2 y\end{bmatrix}$$

or

$$\frac{\sum x_2^2 \sum x_1 y - \sum x_1 x_2 \sum x_2 y}{\sum x_1^2 \sum x_2^2 - \left(\sum x_1 x_2\right)^2}\mathbf{x}_1 + \frac{\sum x_1^2 \sum x_2 y - \sum x_1 x_2 \sum x_1 y}{\sum x_1^2 \sum x_2^2 - \left(\sum x_1 x_2\right)^2}\mathbf{x}_2 .$$

Which is my best guess at $E(\mathbf{y} \mid \mathbf{x}_1, \mathbf{x}_2)$. Note that this is the same result obtained in (A.15) but it cannot be reduced to (A.16) because of the correlation between $\mathbf{x}_1$ and $\mathbf{x}_2$. Therefore, I have shown the first half of establishing that the mathematical expression that estimates $E(\mathbf{y} \mid \mathbf{x}_1, \mathbf{x}_2)$ also estimates $E(\mathbf{y} \mid \tilde{\mathbf{x}}_1, \mathbf{x}_2)$. I now focus on the second half, which is to establish that these two mathematical expressions are equivalent.

Define the results of orthogonalizing $\mathbf{x}_1$ with respect to $\mathbf{x}_2$ is $\tilde{\mathbf{x}}_1$:

$$\tilde{\mathbf{x}}_1 = \mathbf{x}_1 - \gamma \mathbf{x}_2$$

where

$$\gamma = \frac{\sum x_1 x_2}{\sum x_2^2}. \tag{A.17}$$

The mean of y conditional on $\tilde{\mathbf{x}}_1$ and $\mathbf{x}_2$ is given by:

$$E(\mathbf{y} \mid \tilde{\mathbf{x}}_1, \mathbf{x}_2) = \frac{\sum \tilde{x}_1 y}{\sum \tilde{x}_1^2}\tilde{\mathbf{x}}_1 + \frac{\sum x_2 y}{\sum x_2^2}\mathbf{x}_2 \tag{A.18}$$

by extrapolating from (A.16) since $\tilde{\mathbf{x}}_1$ and $\mathbf{x}_2$ are uncorrelated. Recall that it is my goal to show that the expression above is mathematically equivalent to (A.15).

Substituting $\tilde{\mathbf{x}}_1 = \mathbf{x}_1 - \gamma\mathbf{x}_2$ into equation (A.18) yields:

$$\frac{\sum(x_1 - \gamma x_2)y}{\sum(x_1 - \gamma x_2)^2}(\mathbf{x}_1 - \gamma\mathbf{x}_2) + \frac{\sum x_2 y}{\sum x_2^2}\mathbf{x}_2 \tag{A.19}$$

which is my best guess at $E(\mathrm{y} \mid \tilde{\mathbf{x}}_1, \mathbf{x}_2)$. Substituting the definition of γ from (A.17) into (A.19) yields:

$$\frac{\sum\left(x_1 - \frac{\sum x_1 x_2}{\sum x_2^2}x_2\right)y}{\sum\left(x_1 - \frac{\sum x_1 x_2}{\sum x_2^2}x_2\right)^2}\left(\mathbf{x}_1 - \frac{\sum x_1 x_2}{\sum x_2^2}\mathbf{x}_2\right) + \frac{\sum x_2 y}{\sum x_2^2}\mathbf{x}_2$$

which is my best guess at $E(\mathrm{y} \mid \tilde{\mathbf{x}}_1, \mathbf{x}_2)$. Multiplying the numerator and denominator of the first term by $\sum x_2^2$, distributing the summation sign and simplifying yields:

$$\frac{\sum x_2^2 \sum x_1 y - \sum x_1 x_2 \sum x_2 y}{\sum x_2^2 \sum x_1^2 - \left(\sum x_1 x_2\right)^2}\left(\mathbf{x}_1 - \frac{\sum x_1 x_2}{\sum x_2^2}\mathbf{x}_2\right) + \frac{\sum x_2 y}{\sum x_2^2}\mathbf{x}_2$$

which is the operational equivalent to $E(\mathrm{y} \mid \tilde{\mathbf{x}}_1, \mathbf{x}_2)$.

Combining terms of $\mathbf{x}_2$ yields:

$$\frac{\sum x_2^2 \sum x_1 y - \sum x_1 x_2 \sum x_2 y}{\sum x_2^2 \sum x_1^2 - \left(\sum x_1 x_2\right)^2}\mathbf{x}_1 + \mathbf{x}_2\left[\frac{\sum x_2 y}{\sum x_2^2} - \frac{\sum x_1 x_2}{\sum x_2^2}\left(\frac{\sum x_2^2 \sum x_1 y - \sum x_1 x_2 \sum x_2 y}{\sum x_2^2 \sum x_1^2 - \left(\sum x_1 x_2\right)^2}\right)\right] \tag{A.20}$$

My goal is to be able to write (A.20) as (A.15). The coefficient on $\mathbf{x}_1$ are identical in both equations and do not require any work. On the other hand, the coefficients on $\mathbf{x}_2$ are not identical. I will now try to rewrite the second term of (A.20) as the coefficient of $\mathbf{x}_2$ in (A.15). Rewriting the coefficient of $\mathbf{x}_2$ in (A.20) in terms of a common denominator yields:

$$\left[\frac{1}{\sum x_2^2}\right]\left[\frac{1}{\sum x_1^2 \sum x_2^2 - \left(\sum x_1 x_2\right)^2}\right]\left[\sum x_2 y \sum x_1^2 \sum x_2^2 - \sum x_1 x_2 \sum x_2^2 \sum x_1 y\right]$$

which can be further reduced to:

$$\frac{\sum x_1^2 \sum x_2 y - \sum x_1 x_2 \sum x_1 y}{\sum x_1^2 \sum x_2^2 - \left(\sum x_1 x_2\right)^2}$$

and is exactly the coefficient of $\mathbf{x}_2$ in (A.15). Therefore, I have been able to reduce (A.20) so that it is identical to (A.15). That is, the expression that would be used to operationally interpret

$$E(\mathbf{y} \mid \tilde{\mathbf{x}}_1, \mathbf{x}_2) \text{ and } E(\mathbf{y} \mid \mathbf{x}_1, \mathbf{x}_2) \tag{A.21}$$

are the same. I will apply the same aspects of this analysis to my derivation of the Kalman filter.

Derivation of the Kalman filter

The goal of the Kalman filter is to estimate the conditional mean of $\mathbf{x}_t$. Define

$$\hat{\mathbf{x}}_{t,t} = E(\mathbf{x}_t \mid \mathbf{y}_0, \ldots, \mathbf{y}_t)$$

as the estimate made at time t of the state variable at time t. Define the new information of the innovation contained in $\mathbf{y}_t$ as:

$$\tilde{\mathbf{y}}_{t,t-1} = \mathbf{y}_t - \hat{\mathbf{y}}_{t,t-1} \tag{A.22}$$

where $\hat{\mathbf{y}}_{t,t-1}$ is the forecast made at time t-1 for y at time t. The Kalman filtered estimate can be partitioned into:

$$\hat{\mathbf{x}}_{t,t} = E(\mathbf{x}_t \mid \mathbf{y}_0,\dots,\mathbf{y}_{t-1},\tilde{\mathbf{y}}_{t,t-1}) \tag{A.23}$$

which states that mean of $\mathbf{x}_t$ is conditional on historical information and the innovation or the orthogonalized component $\tilde{\mathbf{y}}_{t,t-1}$. Because the innovation is orthogonal to the historical information, equation (A.23) can be decomposed into:

$$\hat{\mathbf{x}}_{t,t} = E(\mathbf{x}_t \mid \mathbf{y}_0,\dots,\mathbf{y}_{t-1}) + E(\mathbf{x}_t \mid \tilde{\mathbf{y}}_{t,t-1}). \tag{A.24}$$

Equation (A.24) states that the forecast for $\mathbf{x}_t$ is dependent on the old information available and some new information contained in $\mathbf{y}_t, \tilde{\mathbf{y}}_{t,t-1}$. Clearly, the analogy can be drawn from the first few pages of this appendix. The forecast for $\mathbf{x}_t$ will be a weighting scheme consisting of old and new information that are independent or each other.

I have completed all the necessary preliminary work and now focus on the derivation of the Kalman filter, by definition, I can rewrite (A.22) as:

$$\tilde{\mathbf{y}}_{t,t-1} = \mathbf{y}_t - E(\mathbf{y}_t \mid \mathbf{y}_0,\dots,\mathbf{y}_{t-1}).$$

My goal is to be able to rewrite the filtered estimate as a function of a previous estimate and the expected value of $\mathbf{x}_t$ conditional on the innovation. That is, I want to rewrite (A24) as:

$$\hat{\mathbf{x}}_{t,t} = \mathbf{A}\hat{\mathbf{x}}_{t-1,t-1} + E(\mathbf{x}_t \mid \tilde{\mathbf{y}}_{t,t-1}) \tag{A.25}$$

To effect this refinement, I must first refer back to the original transition equation. Recall that the transition equation was defined as:

$$\mathbf{x}_{t+1} = \mathbf{A}\mathbf{x}_t + \mathbf{D}\mathbf{w}_t$$

and therefore,

$$\mathbf{x}_t = \mathbf{A}\mathbf{x}_{t-1} + \mathbf{D}\mathbf{w}_{t-1}$$

must be true. The expectation of which yields:

$$\hat{\mathbf{x}}_{t,t-1} = \mathbf{A}\hat{\mathbf{x}}_{t-1,t-1} + \mathbf{0} \tag{A.26}$$

since $\mathbf{w}_t$ is distributed with a mean of 0. Using (A.26), I can rewrite the mean of $\mathbf{x}_t$ conditional on $\mathbf{y}_0,\ldots\ldots,\mathbf{y}_{t-1}$ as:

$$\mathbf{A}E(\mathbf{x}_{t-1} \mid \mathbf{y}_0,\ldots\ldots\ldots,\mathbf{y}_{t-1}) + \mathbf{0}. \tag{A.27}$$

The filtered estimate becomes:

$$\hat{\mathbf{x}}_{t,t} = \mathbf{A}\hat{\mathbf{x}}_{t-1,t-1} + E(\mathbf{x}_t \mid \tilde{\mathbf{y}}_{t,t-1}) \tag{A.28}$$

which accomplishes our goal. Equation (A.28) states that I am able to partition the forecast into last periods forecast weighted by the transition matrix **A** and a new forecast based on last periods innovation.

The forecast error of $\hat{\mathbf{x}}_{t,t-1}$ is given by:

$$\tilde{\mathbf{x}}_{t,t-1} = \mathbf{x}_t - \hat{\mathbf{x}}_{t,t-1}. \tag{A.29}$$

Using (A.26) and the definition of the transition equation, the right hand side of (A.29) can be rewritten as:

$$\tilde{\mathbf{x}}_{t,t-1} = \mathbf{A}\mathbf{x}_{t-1} + \mathbf{D}\mathbf{w}_{t-1} - \mathbf{A}\hat{\mathbf{x}}_{t-1,t-1}$$

$$\tilde{\mathbf{x}}_{t,t-1} = \mathbf{A}(\mathbf{x}_{t-1} - \hat{\mathbf{x}}_{t-1,t-1}) + \mathbf{D}\mathbf{w}_{t-1}.$$

The variance of the forecast of $\tilde{\mathbf{x}}_{t,t-1}$ is defined as $\mathbf{P}_{t,t-1}$ and is given by:

$$E\left[\left(\mathbf{A}(\tilde{\mathbf{x}}_{t-1,t-1}) + \mathbf{D}\mathbf{w}_{t-1}\right)\left(\mathbf{A}(\tilde{\mathbf{x}}_{t-1,t-1}) + \mathbf{D}\mathbf{w}_{t-1}\right)'\right] \equiv \mathbf{P}_{t,t-1} \tag{A.30}$$

Expanding the left hand side of (A.30) yields,

$$\mathbf{A}\mathbf{P}_{t-1,t-1}\mathbf{A}' + \mathbf{D}\mathbf{Q}_{t-1}\mathbf{D}' = \mathbf{P}_{t,t-1} \tag{A.31}$$

since the $\tilde{\mathbf{x}}_{t-1,t-1}$ and $\mathbf{w}_t$ are uncorrelated.

The second term of (A.25) states that the forecast or estimate of $\hat{\mathbf{x}}_{t,t}$ depends on $E(\mathbf{x}_t \mid \tilde{\mathbf{y}}_{t,t-1})$. Using results from (A.10), I can rewrite this conditional expectation as:

$$E(\mathbf{x}_t \mid \tilde{\mathbf{y}}_{t,t-1}) = \mathbf{\Sigma}_{x\tilde{y}}\mathbf{\Sigma}_{\tilde{y}\tilde{y}}^{-1}\tilde{\mathbf{y}}_{t,t-1}. \tag{A.32}$$

Equation (A.32) states that my filtered estimate is dependent on the covariance of the state variable with the innovation and the variance of the innovation itself. I will now rewrite each part of the right hand side in terms of what I know, matrices **M, A, D, R** and **Q** so that the estimate is tractable. Restating the definition of the innovation in equation (A.22) is helpful at this point and is given by:

$$\tilde{\mathbf{y}}_{t,t-1} = \mathbf{y}_t - \hat{\mathbf{y}}_{t,t-1}.$$

Utilizing the original description of the measurement equation and substituting for $\hat{\mathbf{y}}_{t,t-1}$ (A.23) can be rewritten as:

$$\tilde{\mathbf{y}}_{t,t-1} = \mathbf{y}_t - \mathbf{M}\hat{\mathbf{x}}_{t,t-1}.$$

Further substitution for $\mathbf{y}_{t,}$ yields:

$$\tilde{\mathbf{y}}_{t,t-1} = \mathbf{M}\mathbf{x}_t + \mathbf{v}_t - \mathbf{M}\hat{\mathbf{x}}_{t,t-1}$$

and combining terms:

$$\tilde{\mathbf{y}}_{t,t-1} = \mathbf{M}(\mathbf{x}_t - \hat{\mathbf{x}}_{t,t-1}) + \mathbf{v}_t \tag{A.33}$$

and finally,

$$\tilde{\mathbf{y}}_{t,t-1} = \mathbf{M}\tilde{\mathbf{x}}_{t,t-1} + \mathbf{v}_t. \tag{A.34}$$

Equation (A.34) implies that the innovation is dependent on the estimation error of $\mathbf{x}_t$ which is intuitive given our state space system.

The variance-covariance matrix of (A.34) is given by:

$$E(\tilde{\mathbf{y}}_{t,t-1}\tilde{\mathbf{y}}'_{t,t-1}) = E\left[(\mathbf{M}\tilde{\mathbf{x}}_{t,t-1} + \mathbf{v}_t)(\mathbf{M}\tilde{\mathbf{x}}_{t,t-1} + \mathbf{v}_t)'\right] = E[\mathbf{M}\tilde{\mathbf{x}}_{t,t-1}\tilde{\mathbf{x}}'_{t,t-1}\mathbf{M}' + \mathbf{v}_t\mathbf{v}'_t]$$

Where the cross products disappear since $\tilde{\mathbf{x}}_t$ and $\mathbf{v}_t$ are independent. This equation can be further simplified by invoking a definition from (A.30):

$$= \mathbf{M}\mathbf{P}_{t,t-1}\mathbf{M}' + \mathbf{R} \tag{A.35}$$

That is, I have now defined the inverse of the middle term of (A.32).

$$\mathbf{\Sigma}_{yy}^{-1} = [\mathbf{M}\mathbf{P}_{t,t-1}\mathbf{M}' + \mathbf{R}]^{-1} \tag{A.36}$$

Working on the first term of (A.32), $\mathbf{\Sigma}_{x\tilde{y}}$, and using (A.33):

$$E(\mathbf{x}_t\tilde{\mathbf{y}}'_{t,t-1}) = E\left[\mathbf{x}_t(\mathbf{M}(\mathbf{x}_t - \hat{\mathbf{x}}_{t,t-1}) + \mathbf{v}_t)'\right] = E\left[\mathbf{x}_t(\mathbf{M}\tilde{\mathbf{x}}_{t,t-1} + \mathbf{v}_t)'\right].$$

Rewriting the expressions above yields:

$$E[(\mathbf{x}_t\tilde{\mathbf{x}}'_{t,t-1})\mathbf{M}' + \mathbf{x}_t\mathbf{v}'_t] = E[(\mathbf{x}_t\tilde{\mathbf{x}}'_{t,t-1})\mathbf{M}']$$

since $\mathbf{x}_t$ and $\mathbf{v}_t$ are orthogonal. Rewriting $\mathbf{x}_t$ in terms of $\tilde{\mathbf{x}}_{t,t-1}$ and $\hat{\mathbf{x}}_{t,t-1}$ yields:

$$= E[(\tilde{\mathbf{x}}_{t,t-1} + \hat{\mathbf{x}}_{t,t-1})(\tilde{\mathbf{x}}'_{t,t-1})\mathbf{M}'].$$

Finally, the covariance between the Kalman and the innovation is:

$$= E[(\tilde{\mathbf{x}}_{t,t-1}\tilde{\mathbf{x}}'_{t,t-1})\mathbf{M}'] = \mathbf{P}_{t,t-1}\mathbf{M}' \tag{A.37}$$

since a standard property of least squares is that the estimator and error are orthogonal.

I am now ready to define the Kalman filter. Recall from (A.24) that the Kalman filter estimate is:

$$\hat{\mathbf{x}}_{t,t} = E(\mathbf{x}_t \mid \mathbf{y}_0,\ldots\ldots,\mathbf{y}_{t-1}) + E(\mathbf{x}_t \mid \tilde{\mathbf{y}}_{t,t-1}).$$

Substituting results from (A.10) and (A.24), the Kalman filtered estimate can be rewritten as:

$$\hat{\mathbf{x}}_{t,t} = \hat{\mathbf{x}}_{t,t-1} + \Sigma_{x\tilde{y}}\Sigma_{\tilde{y}\tilde{y}}^{-1}\tilde{\mathbf{y}}_{t,t-1}.$$

Further substitution using results from (A.36) and (A.37) yields:

$$\hat{\mathbf{x}}_{t,t} = \hat{\mathbf{x}}_{t,t-1} + (\mathbf{P}_{t,t-1}\mathbf{M}')[\mathbf{M}\mathbf{P}_{t,t-1}\mathbf{M}' + \mathbf{R}]^{-1}\tilde{\mathbf{y}}_{t,t-1}. \quad \text{(A.38)}$$

Equation (A.38) computes the Kalman filter estimate and is usually written in terms of a Kalman gain matrix $\mathbf{K}_t$, which is defined as:

$$\mathbf{K}_t = (\mathbf{P}_{t,t-1}\mathbf{M}')[\mathbf{M}\mathbf{P}_{t,t-1}\mathbf{M}' + \mathbf{R}]^{-1}. \quad \text{(A.39)}$$

Finally,

$$\hat{\mathbf{x}}_{t,t} = \hat{\mathbf{x}}_{t,t-1} + \mathbf{K}_t(\mathbf{y}_t - \mathbf{M}\hat{\mathbf{x}}_{t,t-1}) \quad \text{(A.40)}$$

which is the final version of the Kalman filter estimate.

That is, the filtered estimate of the state vector is the sum of the predicted value of x at time *t* conditional on information available at time *t-1* and a correction term, which is the product of the gain matrix $\mathbf{K}_t$ and the observation residual, which provides new information on the evolution of the system.

It is imperative that I find the variance of my estimation error for use in the EM algorithm.

$$\tilde{\mathbf{x}}_{t,t} = \mathbf{x}_t - \hat{\mathbf{x}}_{t,t}$$

From (A.40):

$$\tilde{\mathbf{x}}_{t,t} = \mathbf{x}_t - \hat{\mathbf{x}}_{t,t-1} - \mathbf{K}_t(\mathbf{y}_t - \hat{\mathbf{y}}_{t,t-1}).$$

Using (A.22), this can be rewritten:

$$\tilde{\mathbf{x}}_{t,t} = \tilde{\mathbf{x}}_{t,t-1} - \mathbf{K}_t(\tilde{\mathbf{y}}_{t,t-1}).$$

The variance of the estimation error, $E(\tilde{\mathbf{x}}_{t,t}\tilde{\mathbf{x}}'_{t,t})$ is

$$E(\tilde{\mathbf{x}}_{t,t-1}\tilde{\mathbf{x}}'_{t,t-1})+\mathbf{K}_t E(\tilde{\mathbf{y}}_{t,t-1}\tilde{\mathbf{y}}'_{t,t-1})\mathbf{K}'_t-\mathbf{K}_t E(\tilde{\mathbf{y}}_{t,t-1}\tilde{\mathbf{x}}'_{t,t-1})-E(\tilde{\mathbf{x}}_{t,t-1}\tilde{\mathbf{y}}'_{t,t-1})\mathbf{K}'_t$$

Using the results from (A.35) and (A.37) this reduces to:

$$E(\tilde{\mathbf{x}}_{t,t-1}\tilde{\mathbf{x}}'_{t,t-1})+\mathbf{K}_t(\mathbf{M}\mathbf{P}_{t,t-1}\mathbf{M}'+\mathbf{R})\mathbf{K}'_t-\mathbf{K}_t(\mathbf{M}\mathbf{P}'_{t,t-1})-(\mathbf{P}_{t,t-1}\mathbf{M}')\mathbf{K}'_t \qquad \text{(A.41)}$$

This complex expression for the variance of the estimation can be simplified using the definition of the Kalman gain matrix. The second term of (A.41) reduces to:

$$\mathbf{K}_t(\mathbf{M}\mathbf{P}_{t,t-1}\mathbf{M}'+\mathbf{R})\mathbf{K}'_t=\mathbf{K}_t(\mathbf{M}\mathbf{P}_{t,t-1}\mathbf{M}'+\mathbf{R})(\mathbf{M}\mathbf{P}_{t,t-1}\mathbf{M}'+\mathbf{R})^{-1}\mathbf{M}\mathbf{P}'_{t,t-1}=\mathbf{K}_t\mathbf{M}\mathbf{P}'_{t,t-1}$$

after substitution from (A.39). In addition, note that by definition, the first term in (A.41) is $\mathbf{P}_{t,t-1}$. Therefore, (A.41) reduces to:

$$\mathbf{P}_{t,t-1}+\mathbf{K}_t\mathbf{M}\mathbf{P}'_{t,t-1}-\mathbf{P}_{t,t-1}\mathbf{M}'\mathbf{K}'_t-\mathbf{K}_t\mathbf{M}\mathbf{P}'_{t,t-1}=\mathbf{P}_{t,t-1}-\mathbf{P}_{t,t-1}\mathbf{K}_t\mathbf{M}$$

or simply:

$$\mathbf{P}_{t,t-1}(\mathbf{I}-\mathbf{K}_t\mathbf{M}). \qquad \text{(A.42)}$$

The derivation of the Kalman filter is complete. The Kalman filter can best be summarized pictorially. Figure A.1 describes the algorithm, which I have just derived.

Figure A.1

1. *Enter prior estimate and error covariance*

$$\mathbf{K}_t = \left(\mathbf{P}_{t,t-1}\mathbf{M}'\right)\left[\mathbf{M}\mathbf{P}_{t,t-1}\mathbf{M}' + \mathbf{R}\right]^{-1}$$

5. *Compute Kalman Gain*

$$\hat{\mathbf{x}}_{t+1,t} = \mathbf{A}\hat{\mathbf{x}}_{t,t}$$

$$\mathbf{P}_{t+1,t} = \mathbf{A}\mathbf{P}_{t,t}\mathbf{A}' + \mathbf{Q}_t$$

4. *project ahead*

2. *Update estimate with innovation*

$$\hat{\mathbf{x}}_{t,t} = \hat{\mathbf{x}}_{t,t-1} + \mathbf{K}_t\left(\mathbf{y}_t - \mathbf{M}\hat{\mathbf{x}}_{t,t-1}\right)$$

3. *Compute error covariance for updated estimate*

$$\mathbf{P}_{t,t-1}\left(\mathbf{I} - \mathbf{K}_t\mathbf{M}\right)$$

A simple scalar example is provided to clarify the derivation and to illustrate the recursive nature of the Kalman filter.

Consider the following system.

$$y_t = x_t + v_t$$
$$x_{t+1} = x_t + w_t$$

where in this simple system A = M = D =1. In addition assume

$$V(w_t) = Q = 1 \quad V(v_t) = R = 2 \qquad \mathrm{V}(\mathrm{x}_0) = 1$$

The Kalman filter estimate given by (A.40) is:

$$\hat{\mathbf{x}}_{t,t} = \hat{\mathbf{x}}_{t,t-1} + \mathbf{K}_t\left(\mathbf{y}_t - \mathbf{M}\hat{\mathbf{x}}_{t,t-1}\right)$$

or using (A.38) and the fact that M=1:

$$\hat{\mathbf{x}}_{t,t} = \hat{\mathbf{x}}_{t,t-1} + \mathbf{P}_{t,t-1}\left(\mathbf{P}_{t,t-1} + \mathbf{R}\right)^{-1}\tilde{\mathbf{y}}_{t,t-1}.$$

The variance of the estimation error, $\mathbf{P}_{t,t-1}$ from (A.31) is:

$$\mathbf{P}_{t,t-1} = \mathbf{A}\mathbf{P}_{t-1,t-1}\mathbf{A}' + \mathbf{D}\mathbf{Q}_{t-1}\mathbf{D}' = \mathbf{P}_{t-1,t-1} + \mathbf{Q}_{t-1}$$

when A=D=1 in my scalar example.

Using the equation above but changing notation from vector to scalar which is appropriate for this example, I can derive Kalman filter estimates beginning at time $t=1$.

Time $t = 1$

$$\hat{x}_{1,1} = \frac{1}{2}\hat{x}_{1,0} + \frac{1}{2}y_1$$

Time $t = 2$

$$P_{1,0} = P_{0,0} + Q = 1 + 1 = 2$$

$$\hat{x}_{1,1} = \hat{x}_{1,0} + (2)(2+2)^{-1}(y_1 - \hat{x}_{1,0})$$

$$\hat{x}_{1,1} = \hat{x}_{1,0} + \frac{1}{2}(y_1 - \hat{x}_{1,0})$$

$$\hat{x}_{2,2} = \hat{x}_{2,1} + K_2(y_2 - 1\hat{x}_{2,1})$$

$$\hat{x}_{2,1} = A\hat{x}_{1,1} = 1\hat{x}_{1,1}$$

$$\hat{x}_{2,2} = 1\left(\frac{1}{2}\hat{x}_{1,0} + \frac{1}{2}y_1\right) + K_2\left(y_2 - 1\left(\frac{1}{2}\hat{x}_{1,0} + \frac{1}{2}y_1\right)\right)$$

$$K_2 = P_{t,t-1}(P_{t,t-1} + R)^{-1} = P_{2,1}(P_{2,1} + R)^{-1}$$

$$P_{2,1} = P_{1,1}(I - KM) + Q = 2\left(1 - \frac{1}{2}\right) + 1 = 2$$

$$K_2 = 2(2+2)^{-1} = \frac{1}{2}$$

$$\hat{x}_{2,2} = \frac{1}{2}\hat{x}_{1,0} + \frac{1}{2}y_1 + \frac{1}{2}y_2 - \frac{1}{4}\hat{x}_{1,0} - \frac{1}{4}y_1 = \frac{1}{4}\hat{x}_{1,0} + \frac{1}{4}y_1 + \frac{1}{2}y_2$$

Time $t = 3$

$$\hat{x}_{3,3} = \hat{x}_{3,2} + P_{3,2}(P_{3,2} + R)^{-1}(y_3 - \hat{x}_{3,2})$$

$$\hat{x}_{3,2} = A\hat{x}_{2,2} = 1\hat{x}_{2,2}$$

$$\hat{x}_{3,3} = \hat{x}_{2,2} + P_{3,2}(P_{3,2} + R)^{-1}(y_3 - \hat{x}_{2,2})$$

$$P_{3,2} = P_{2,2} + Q$$

$$P_{2,2} = P_{2,1}(I - KM) = 2\left(1 - \frac{1}{2}1\right) = 2 - 1 = 1$$

$$P_{3,2} = 1 + 1 = 2$$

$$\hat{x}_{3,3} = \hat{x}_{2,2} + 2(2+2)^{-1}(y_3 - \hat{x}_{2,2}) = \frac{1}{8}x_{1,0} + \frac{1}{8}y_1 + \frac{1}{4}y_2 + \frac{1}{2}y_3$$

It is important to note that as I iterate through the Kalman filter, the most recent observation is weighted more heavily and the initial estimate is given the least weight. Intuitively, this is what you would expect to yield the most accurate estimate.

A.2 The EM Algorithm

Building on the state space model and the Kalman filter discussed in Part I, the EM algorithm used in this research will now be discussed. The state space model developed earlier is repeated here in scalar format.

Measurement Equation: $y_t = ax_t + v_t$

Transition Equation: $x_{t+1} = \phi x_t + w_t$

Where v_t and w_t are distributed with mean zero and $V(v_t) = r$ and $V(w_t) = q$. The initial value of x_0 may be assumed to be a random variable with mean μ_0 and σ_0.

The estimation of the parameters involved in the specification of the state space model above can be accomplished using maximum likelihood under the assumption that $x_0, w_1, \ldots\ldots, w_T$ and $v_1, \ldots\ldots, v_T$ are jointly normal and uncorrelated variables. The EM algorithm proceeds by successively maximizing the conditional expectation of the likelihood function.

Consider the joint likelihood function defined $L(x,v)$ given by:

$$\frac{1}{\sqrt{2\pi\sigma_0}} e^{-\frac{(x_0-\mu_0)^2}{2\sigma_0}} \left[\frac{1}{\sqrt{2\pi q}}\right]^{T} e^{-\frac{\Sigma(x_t-\phi x_{t-1})^2}{2q}} \left[\frac{1}{\sqrt{2\pi r}}\right]^{T} e^{-\frac{\Sigma(y_t-ax_t)^2}{2r}} \quad \text{(A.43)}$$

Let θ represent the parameters $\phi, r, q, \sigma_0, \mu_0$. Clearly, it is much easier to maximize the natural logarithm of $L(x,v,\theta)$ than the function directly. Consider:

$$\ln L \propto -\frac{1}{2}\ln\sigma_0 - \frac{T}{2}\ln q - \frac{T}{2}\ln r - \frac{1}{2}\sum\frac{(x_t - \phi x_{t-1})^2}{q} - \frac{1}{2}\sum\frac{(y_t - ax_t)^2}{r} - \frac{1}{2}\frac{(x_0 - \mu_0)^2}{\sigma_0} \tag{A.44}$$

While $\ln L(x,v,\theta)$ involves the parameters in a convenient form, they still cannot be maximized directly since x_t is not observed. The role of the Kalman should now be clear. It is the Kalman filter output that allows maximization of (A.44). Recall that the Kalman filter documents otherwise unobservable variables. Using the Kalman filter output, I am able to maximize (A.44) and solve for the appropriate parameters.

If θ is θ_t, and E_t denotes the expectation under θ_t, then the function:

$$E_t\left[\ln L(x,v;\theta \mid y)\right] \tag{A.45}$$

can be written in terms of the Kalman filter output. From here on , the time subscripts in (A.45) are assumed and will not be explicitly noted.

Working on the expectation of (A.45) term by term, I will rewrite (A.44) so that I can obtain the maximum likelihood estimators for the unobservable parameters.

Evaluating the expectation of the first term of (A.44) involving x_t:

$$E\left[-\frac{1}{2}\sum\frac{(x_t - \phi x_{t-1})^2}{q}\right] = -\frac{1}{2q}E\left[\sum\left(\hat{x}_t^T + \tilde{x}_t^T - \phi(\hat{x}_{t-1}^T + \tilde{x}_{t-1}^T)\right)^2\right] \tag{A.46}$$

by involving $\tilde{x}_t^T = x_t^T - \hat{x}_t^T$, the Kalman error defined in Part I of this appendix. Recall that $\hat{x}_t^T$ is the expectation of x at t given $y_1,\ldots\ldots,y_T$. $\hat{x}$ and $\tilde{x}_t$ are independent of one another because a standard property of a least squares estimator is that it is orthogonal to its estimation error. Consequently, (A.46) can be written as:

$$-\frac{1}{2q}E\sum\left\{\begin{aligned}&(\hat{x}_t^T)^2 + (\tilde{x}_t^T)^2 + \phi^2(\hat{x}_{t-1}^T)^2 + \phi^2(\tilde{x}_{t-1}^T)^2 + 2\hat{x}_t^T\tilde{x}_t^T - 2\phi\hat{x}_t^T\hat{x}_{t-1}^T\\&-2\phi\hat{x}_t^T\tilde{x}_{t-1}^T - 2\phi\tilde{x}_t^T\hat{x}_{t-1}^T - 2\phi\tilde{x}_t^T\tilde{x}_{t-1}^T + 2\phi^2\hat{x}_{t-1}^T\tilde{x}_{t-1}^T\end{aligned}\right\} \tag{A.47}$$

To illustrate the point that a least squares estimator is orthogonal to its estimation error, consider the following:

$$E\left(\tilde{x}_t^T \hat{x}_t^T\right) = E\left((x_t - \hat{x}_t^T)\hat{x}_t^T\right)$$

the OLS analog of which is:

$$\left(\mathbf{y} - \mathbf{x}\hat{\boldsymbol{\beta}}\right)' \mathbf{x}\hat{\boldsymbol{\beta}} = \mathbf{y}'\mathbf{x}\hat{\boldsymbol{\beta}} - \hat{\boldsymbol{\beta}}'\mathbf{x}'\mathbf{x}\hat{\boldsymbol{\beta}}$$

$$\mathbf{y}'\mathbf{x}\hat{\boldsymbol{\beta}} - \left((\mathbf{x}'\mathbf{x})^{-1}\mathbf{x}'\mathbf{y}\right)\mathbf{x}'\mathbf{x}\hat{\boldsymbol{\beta}} = \mathbf{y}'\mathbf{x}\hat{\boldsymbol{\beta}} - \mathbf{y}'\mathbf{x}\hat{\boldsymbol{\beta}} = 0$$

where $\mathbf{y}'$ is *1xT* and $\mathbf{x}$ is *Txk* and $\hat{\boldsymbol{\beta}}$ is *kx1*.

Invoking this property, (A.47) reduces to:

$$-\frac{1}{2q} E\sum\left\{(\hat{x}_t^T)^2 + (\tilde{x}_t^T)^2 + \phi^2(\hat{x}_{t-1}^T)^2 + \phi^2(\tilde{x}_{t-1}^T)^2 - 2\phi\hat{x}_t^T\hat{x}_{t-1}^T - 2\phi\tilde{x}_t^T\tilde{x}_{t-1}^T\right\}. \quad \text{(A.48)}$$

Passing the expectation operator through the summation sign in (A.48) yields:

$$-\frac{1}{2q}\sum\begin{Bmatrix} E(\hat{x}_t^T)^2 + E(\tilde{x}_t^T)^2 + \phi^2 E(\hat{x}_{t-1}^T)^2 + \phi^2 E(\tilde{x}_{t-1}^T)^2 - 2\phi E(\hat{x}_t^T\hat{x}_{t-1}^T) \\ -2\phi E(\tilde{x}_t^T\tilde{x}_{t-1}^T) \end{Bmatrix}. \quad \text{(A.49)}$$

Realizing that $E(\hat{x}_t^T) = \hat{x}_t^T$ and applying the summation across terms reduces (A.49) to:

$$-\frac{1}{2q}\begin{Bmatrix} \sum(\hat{x}_t^T)^2 + \sum E(\tilde{x}_t^T)^2 + \phi^2\sum(\hat{x}_{t-1}^T)^2 + \phi^2\sum E(\tilde{x}_{t-1}^T)^2 - 2\phi\sum(\hat{x}_t^T\hat{x}_{t-1}^T) \\ -2\phi\sum E(\tilde{x}_t^T\tilde{x}_{t-1}^T) \end{Bmatrix} \quad \text{(A.50)}$$

where now terms are written in terms of variance and covariances of the Kalman filter estimate and error across time. In summary, I have started with equation (A.46), the expectation of the first term of (A.44) involving x_t and rewrote it in terms of what is known, the Kalman filter output. The benefit of (A.50) is that I can maximize it with respect to the parameter of interest. In much the same fashion, I will be rewriting the remaining two terms of (A.44) and ultimately solving for the parameters, ϕ, q and r.

Similarly, evaluating the second term of (A.44) involving x_t yields:

$$E\left[-\frac{1}{2}\sum\frac{(y_t-ax_t^T)^2}{r}\right]=-\frac{1}{2r}\sum E(y_t-ax_t^T)^2=-\frac{1}{2r}\sum E(y_t-a\hat{x}_t^T-a\tilde{x}_t^T)^2 \quad (A.51)$$

where I have taken the expectations operator through the summation sign and utilized the definition of the Kalman error. Expanding the last equation in (A.51) yields:

$$-\frac{1}{2r}\sum E\left[(y_t-a\hat{x}_t^T)^2+a^2(\tilde{x}_t^T)^2-2y_t a\tilde{x}_t^T+2a^2\hat{x}_t^T\tilde{x}_t^T\right] \quad (A.52)$$

Utilizing the fact that:

(1) $\hat{x}$ and $\tilde{x}$ are independent of one another,

(2) $E(\tilde{x})=0$,

(3) $E(\hat{x}_t^T)=\hat{x}_t^T$,

and

(4) $E\left[-2y_t a\tilde{x}_t^T\right]=-2y_t aE\left[\tilde{x}_t^T\right]$ since a and y_t are known,

equation (A.52) can be reduced to:

$$-\frac{1}{2r}\left[\sum(y_t-a\hat{x}_t^T)^2+a^2\sum E\left[\tilde{x}_t^T\right]^2\right]. \quad (A.53)$$

This term of (A.44) has also been rewritten in terms of the Kalman filter output, specifically the variance of the Kalman filter error. Therefore, this term can also be maximized with respect to the parameter of interest.

Evaluating the last term of (A.44) yields:

$$E\left[-\frac{1}{2}\frac{(x_0-\mu_0)^2}{\sigma_0}\right]=-\frac{1}{2\sigma_0}E\left[\hat{x}_0^T+\tilde{x}_0^T-\mu_0\right]^2$$

$$= -\frac{1}{2\sigma_0} E\left[\hat{x}_0^T - \mu_0\right]^2 - \frac{1}{2\sigma_0} E\left[\tilde{x}_0^T \tilde{x}_0^T + 2\hat{x}_0^T \tilde{x}_0^T - 2\mu_0 \tilde{x}_0^T\right] \tag{A.54}$$

by substitution of $\tilde{x}_t^T = x_t^T - \hat{x}_t^T$. Utilizing $E(\tilde{x}_t^T) = 0$ and the fact that $\tilde{x}_0^T$ is independent of $\hat{x}_0^T$ and μ_0, equation (A.54) can be rewritten in terms of variances:

$$-\frac{1}{2\sigma_0}\left\{E\left[\hat{x}_0^T - \mu_0\right]^2 + E\left[\tilde{x}_0^T \tilde{x}_0^T\right]\right\} \tag{A.55}$$

This term of (A.44) has also been rewritten in terms of the variance of the Kalman filter error.

Therefore, the last terms of (A.44) have been transformed from terms that were intractable into terms that can be directly maximized. These terms can be combined, maximized and solved for the appropriate parameter(s) because they are written in terms of what is known, the output from the Kalman filter. By combining equations (A.50), (A.53) and (A.55), and using them in place of (A.44), I can describe the natural log of the likelihood function in terms which can be statistically evaluated when maximized with respect to the parameters of interest. Consequently, the function can be solved for the parameters, ϕ, q and r.

Taking the partial derivative of (A.50), (A.53) and (A.55) with respect to ϕ yields:

$$-\frac{1}{2q}\left\{2\phi \sum \left(\hat{x}_{t-1}^T\right)^2 + 2\phi \sum E\left(\tilde{x}_{t-1}^T\right)^2 - 2\sum \left(\hat{x}_t^T \hat{x}_{t-1}^T\right) - 2\sum \left(\tilde{x}_{t-1}^T \tilde{x}_t^T\right)\right\} \tag{A.56}$$

Setting (A.56) equal to zero and solving for ϕ I obtain:

$$\phi\left[\sum \left(\hat{x}_{t-1}^T\right)^2 + \sum E\left(\tilde{x}_{t-1}^T\right)^2\right] = \sum \left(\hat{x}_t^T \hat{x}_{t-1}^T\right) + \sum \left(\tilde{x}_{t-1}^T \tilde{x}_{t-1}^T\right)$$

or more simply,

$$\phi = \frac{\sum\left(\hat{x}_t^T \hat{x}_{t-1}^T\right) + \sum\left(\tilde{x}_t^T \tilde{x}_{t-1}^T\right)}{\left[\sum\left(\hat{x}_t^T\right)^2 + \sum\left(\tilde{x}_{t-1}^T\right)^2\right]}. \tag{A.57}$$

Equation (A.57) states that the coefficients of the lagged state variables in the transition equations are dependent on the variance and covariance of the Kalman filter estimate and the error across time.

Similarly, taking the derivatives of (A.50), (A.53) and (4.55) with respect to q and setting it equal to zero yields:

$$-\frac{T}{2q} + \frac{1}{2q^2}\left\{\begin{matrix}\sum(\hat{x}_t^T)^2 + \sum E(\tilde{x}_t^T)^2 + \phi^2\left[\sum(\hat{x}_{t-1}^T)^2 + \sum E(\tilde{x}_{t-1}^T)^2\right] \\ -2\phi\left[\sum(\hat{x}_t^T \hat{x}_{t-1}^T) + \sum E(\tilde{x}_{t-1}^T \tilde{x}_t^T)\right]\end{matrix}\right\} = 0.$$

Substituting (A.57) into the expression above yields:

$$-\frac{T}{2q} + \frac{1}{2q}\left\{\begin{matrix}\sum(\hat{x}_t^T)^2 + \sum E(\tilde{x}_t^T)^2 + \phi\left[\sum\left(\hat{x}_t^T \hat{x}_{t-1}^T\right) + \sum\left(\tilde{x}_t^T \tilde{x}_{t-1}^T\right)\right] \\ -2\phi\left[\sum(\hat{x}_t^T \hat{x}_{t-1}^T) + \sum E(\tilde{x}_{t-1}^T \tilde{x}_t^T)\right]\end{matrix}\right\} = 0.$$

Combining terms and multiplying by $2q^2$ yields:

$$-Tq + \left\{\sum(\hat{x}_t^T)^2 + \sum E(\tilde{x}_t^T)^2 - \phi\left[\sum(\hat{x}_t^T \hat{x}_{t-1}^T) + \sum E(\tilde{x}_{t-1}^T \tilde{x}_t^T)\right]\right\} = 0.$$

Finally solving for q:

$$q = \frac{\sum(\hat{x}_t^T)^2 + \sum E(\tilde{x}_t^T)^2 - \phi\left[\sum(\hat{x}_t^T \hat{x}_{t-1}^T) + \sum E(\tilde{x}_{t-1}^T \tilde{x}_t^T)\right]}{T}.$$

This equation states that the variance of the transition equation is dependent on the variance of the Kalman filter estimate and the covariance of the Kalman filter estimation error across time.

Lastly, the derivative of (A.50), (A.53) and (A.55) with respect to r is given by:

$$\frac{T}{2r} + \frac{1}{2r^2}\left[\sum(y_t - a\hat{x}_t^T)^2 + a^2 \sum E\left[\tilde{x}_t^T\right]^2\right].$$

Setting the expression equal to zero and solving for r yields:

$$r = \frac{\left|\sum\left[(y_t - a\hat{x}_t^T)^2\right] + a^2 \sum E\left[\tilde{x}_t^T\right]^2\right|}{T}.$$

This equation states that the variance of the measurement equation is dependent on the variance of the Kalman filter estimation error.

In summary, I have solved for the three parameters of interest by rewriting terms of (A.44) in terms of the Kalman filter output and maximizing the resulting equation. The term from (A.54) involving σ_0 and μ_0 involves only a single observation. Therefore, I can regard x_0 as fixed, so that $\sigma_0 = 0$ and $\mu_0 = x_0$.

The overall procedure can be regarded as simply alternating between the Kalman and the multivariate normal maximum likelihood equations. Summarizing the overall procedure:

1. Initialize μ_0, ϕ, q, r and set $\sigma_0 = 0$.

2. Use the Kalman to calculate the Kalman estimate, variance of the estimation error and the covariance of the estimation error across time: $\hat{x}_t^T, E\left(\tilde{x}_t^T \tilde{x}_t^T\right)$ and $E\left(\tilde{x}_t^T \tilde{x}_{t-1}^T\right)$ respectively.

3. Evaluate the following log likelihood with the initial parameters from (1) and the results from the Kalman:

$$-\frac{1}{2}\sum V(\varepsilon_t) - \frac{1}{2}\sum\left[\frac{\left(y_t - a\hat{x}_t^T\right)^2}{V(\varepsilon_t)}\right]. \quad \text{(A.58)}$$

This is also known as the score. The evaluation of (A.58) follows from the fact that:

$$y_t = ax_t + v_t \text{ and } x_t = \phi x_{t-1} + w_{t-1}$$

with ε_t defined as:

$$\varepsilon_t = y_t - a\hat{x}_t^T$$

$$\varepsilon_t = ax_t + v_t - a\hat{x}_t^T$$

$$\varepsilon_t = a\left(x_t - \hat{x}_t^T\right) + v_t .$$

Clearly, I want to minimize our prediction error, $x_t - \hat{x}_t^T$ and the estimation error, v_t. Both errors are summarized by ε_t which has a mean of zero and a variance of:

$$E\left[\varepsilon_t^2\right] = a^2 E\left[\left(x_t - \hat{x}_t^T\right)^2\right] + V\left[v_t\right] = a^2 E\left[\tilde{x}_t^T \tilde{x}_t^T\right] + V\left[v_t\right] = a^2 E\left[\tilde{x}_t^T \tilde{x}_t^T\right] + r$$

The likelihood function I want to maximize is:

$$\prod_1^T \left(\frac{1}{\sqrt{2\pi V(\varepsilon_t)}}\right) e^{-\frac{1}{2}\sum \frac{(y_t - a\hat{x}_t^T)^2}{V(\varepsilon_t)}}$$

The relevant log likelihood function is then proportional to:

$$-\frac{1}{2}\sum V(\varepsilon_t) - \frac{1}{2}\sum \frac{(y_t - a\hat{x}_t^T)^2}{V(\varepsilon_t)}$$

Which is a parsimonious way of writing the log likelihood function (A.44) since ε_t is the "total error".

4. Update ϕ, q and r using the maximizing equations (A.50), (A.53) and (A.55) respectively.

5. Use the new parameters to get new Kalman filter output $\hat{x}_t^T, E\left(\tilde{x}_t^T \tilde{x}_t'^T\right)$ and $E\left(\tilde{x}_t^T \tilde{x}_{t-1}'^T\right)$.

6. Recalculate the log likelihood function and determine if it improves enough to continue. If the score has not improved "much", then it is said that the EM algorithm has converged. For example, the convergence criterion used in this research was set at .0001. That is, the first four decimal places of the score had to remain unchanged to attain convergence.

The remaining portion of this appendix is a simple example of the EM algorithm. This example presents an iterative computation of maximum likelihood estimates when the observations can be viewed as incomplete data. Each iteration consists of an expectation step followed by a maximization step, hence the EM algorithm. The attraction of the EM algorithm is due to its simplicity when applied to a wide range of examples. In particular, when the underlying complete data come from an exponential distribution whose maximum likelihood estimators are easily computed. Therefore, the maximization step of the EM algorithm is also easily computed.

Consider 197 animals, which have characteristics that are distributed multinomial into four categories, so that the observed data consists of:

$$(y_1, y_2, y_3, y_4) = (125, 18, 20, 34).$$

A genetic model for the population specifies cell probabilities as:

$$\left(\frac{1}{2}+\frac{1}{4}\pi, \frac{1}{4}-\frac{1}{4}\pi, \frac{1}{4}-\frac{1}{4}\pi, \frac{1}{4}\pi,\right) \text{ where } 0 \le \pi \le 1.$$

To illustrate the EM algorithm, I represent (y_1, y_2, y_3, y_4) as incomplete data from a five category multinomial population were the cell probabilities are:

$$\left(\frac{1}{2}, \frac{1}{4}\pi, \frac{1}{4}-\frac{1}{4}\pi, \frac{1}{4}-\frac{1}{4}\pi, \frac{1}{4}\pi\right).$$

The idea being to split the first of the four original categories into two categories. Thus, the complete data consists of:

$$(x_1, x_2, x_3, x_4, x_5)$$

where

$$y_1 = x_1 + x_2,\ y_2 = x_3,\ y_3 = x_4,\ y_4 = x_5$$

The complete data specification of the probability density function (pdf) is:

$$f(\bullet \mid \pi) = \frac{(x_1 + x_2 + x_3 + x_4 + x_5)!}{x_1! x_2! x_3! x_4! x_5!} \left(\frac{1}{2}\right)^{x_1} \left(\frac{1}{4}\pi\right)^{x_2} \left(\frac{1}{4} - \frac{1}{4}\pi\right)^{x_3} \left(\frac{1}{4} - \frac{1}{4}\pi\right)^{x_4} \left(\frac{1}{4}\pi\right)^{x_5} \quad \text{(A.59)}$$

Clearly, this pdf is similar to the binomial pdf which is given by:

$$f(x) = \frac{n!}{x!(n-x)!} p^x q^{n-x}.$$

Summing over the multinomial to obtain the expected value will take place over values of x_1 and x_2. In particular, the set (125,0), (124,1), ..., (0,125). Recall that x_3, x_4, and x_5 are known. For the multinomial, the expected value is given by:

$$\sum (x_1 + x_2) f(x_1, x_2 \mid x_3, x_4, x_5, \pi)$$

The expectation step estimates the sufficient statistics of the complete data $(x_1, x_2, x_3, x_4, x_5)$ given the observed data (y_1, y_2, y_3, y_4). This means that before I can maximize $f(x_1, x_2, x_3, x_4, x_5)$, I need to determine the $E(x_1)$ and $E(x_2)$ and insert that into $f(x_1, x_2, x_3, x_4, x_5)$. That is, (A.59) will be given by:

$$f(\bullet \mid \pi) = \frac{(x_1 + x_2 + x_3 + x_4 + x_5)!}{x_1! x_2! x_3! x_4! x_5!} \left(\frac{1}{2}\right)^{\hat{x}_1} \left(\frac{1}{4}\pi\right)^{\hat{x}_2} \left(\frac{1}{4} - \frac{1}{4}\pi\right)^{x_3} \left(\frac{1}{4} - \frac{1}{4}\pi\right)^{x_4} \left(\frac{1}{4}\pi\right)^{x_5}$$

where $E(x_1) = \hat{x}_1$ and $E(x_2) = \hat{x}_2$.

If I need $E(x)$ from $\sum xf(x)$ with $f(x) = C\binom{n}{x}p^x q^{n-x}$, I would find the maximum likelihood estimator of p and multiply by n. The derivation follows. The binomial is give by:

$$f(x) = C\binom{n}{x}p^x q^{n-x}.$$

The natural logarithm of which is:

$$\ln f(x) = \ln C\binom{n}{x} + x\ln p + (n-x)\ln(1-p).$$

Differentiating with respect to p:

$$\frac{\partial \ln f(x)}{\partial p} = \frac{x}{p} + \frac{(n-x)}{(1-p)}(-1)$$

and setting equal to zero:

$$\frac{x}{\hat{p}} = \frac{(n-x)}{(1-\hat{p})}.$$

Rearranging terms:

$$(1-\hat{p})x = (n-x)\hat{p}$$

and finally the estimate of p is:

$$\hat{p} = \frac{x}{n}.$$

Then my best guess at the expected value of x would be $n\hat{p}$.

Given this result, how would I allocate 125 between x_1 and x_2 given this expected value framework:

$$\sum (x_1 + x_2) f(x_1, x_2 \mid \pi, x_3, x_4, x_5)?$$

Maximizing the natural log of the conditional probability density function from (A.59) yields:

$$\ln f(\bullet \mid \pi) = \ln \frac{n!}{x_1!..x_5!} + x_1 \ln p_1 + (125 - x_1) \ln p_2 + x_3 \ln\left(\frac{1}{4} - \frac{1}{4}\pi\right)$$
$$+ x_4 \ln\left(\frac{1}{4} - \frac{1}{4}\pi\right) + x_5 \ln\left(\frac{1}{4}\pi\right)$$

where p_1 and p_2 are the probability of observing x_1 and x_2 respectively if y_1 is observed. Therefore, $p_1 = 1 - p_2$. Differentiating with respect to p_1 and setting this equal to zero yields:

$$\frac{x_1}{\hat{p}_1} + \frac{(125 - x_1)}{(1 - \hat{p}_1)} = 0$$

$$\frac{x_1}{\hat{p}_1} = \frac{(125 - x_1)}{(1 - \hat{p}_1)}$$

$$x_1(1 - \hat{p}_1) = (125 - x_1)\hat{p}_1$$

$$x_1 = 125\hat{p}_1$$

$$\hat{p}_1 = \frac{x_1}{125} \qquad \hat{p}_2 = 1 - \frac{x_1}{125}$$

My best guess at $E(x_1)$ would be $125\hat{p}_1$, and my best guess at $E(x_2)$ would be $125\hat{p}_2$. It is easy to see the similarity to the binomial distribution in these results. From the original specification, I know that:

$$\hat{p}_1 = \frac{1/2}{1/2 + 1/4\,\hat{\pi}}$$

and

$$\hat{p}_2 = \frac{\frac{1}{4}\hat{\pi}}{\frac{1}{2} + \frac{1}{4}\hat{\pi}}.$$

The first iteration will be based on the initial value of $\hat{\pi}$. Using this initial estimate, I can calculate my best guess at the expected value of x_1 and x_2 as:

$$E(x_1) = 125\left[\frac{\frac{1}{2}}{\frac{1}{2} + \frac{1}{4}\hat{\pi}}\right]$$

$$E(x_2) = 125\left[\frac{\frac{1}{4}\hat{\pi}}{\frac{1}{2} + \frac{1}{4}\hat{\pi}}\right].$$

With $\hat{x}_1 \approx E(x_1)$ and $\hat{x}_2 \approx E(x_2)$ in place, I can now maximize $f(x_1, x_2, x_3, x_4, x_5)$ with respect to π. To obtain another $\hat{\pi}$ I compute the following:

$$f(x \mid \pi) = \frac{n!}{x_1!\ldots x_5!}\left(\frac{1}{2}\right)^{\hat{x}_1}\left(\frac{1}{4}\pi\right)^{\hat{x}_2}\left(\frac{1}{4} - \frac{1}{4}\pi\right)^{18}\left(\frac{1}{4} - \frac{1}{4}\pi\right)^{20}\left(\frac{1}{4}\pi\right)^{34}$$

Maximizing $\ln f(x \mid \pi)$ begins with:

$$\ln\frac{n!}{x_1!\ldots x_5!} + \hat{x}_1 \ln\frac{1}{2} + \hat{x}_2 \ln\left(\frac{1}{4}\pi\right) + 18\ln\left(\frac{1}{4} - \frac{1}{4}\pi\right) + 20\ln\left(\frac{1}{4} - \frac{1}{4}\pi\right) + 34\ln\left(\frac{1}{4}\pi\right)$$

and taking the derivative:

$$\frac{\partial \ln f(x \mid \pi)}{\partial \pi} = \frac{\hat{x}_2\left(\frac{1}{4}\right)}{\frac{1}{4}\pi} + \frac{18\left(-\frac{1}{4}\right)}{\frac{1}{4} - \frac{1}{4}\pi} + \frac{20\left(-\frac{1}{4}\right)}{\frac{1}{4} - \frac{1}{4}\pi} + \frac{34\left(\frac{1}{4}\right)}{\frac{1}{4}\pi}.$$

Now setting it equal to zero and solving for π:

$$(1 - \pi)\hat{x}_2 - 18\pi - 20\pi + 34(1 - \pi) = 0$$

$$\pi = \frac{\hat{x}_2 + 34}{\hat{x}_2 + 18 + 20 + 34}.$$

The EM algorithm for this example is defined by cycling back and forth between the expression for $\hat{x}_2$ and π where π_a is the initial estimate for π:

$$\hat{x}_2^a = 125\left[\frac{\frac{1}{4}\hat{\pi}_a}{\frac{1}{2}+\frac{1}{4}\hat{\pi}_a}\right]$$

$$\pi_b = \frac{\hat{x}_2^a + 34}{\hat{x}_2^a + 18 + 20 + 34}$$

$$\hat{x}_2^b = 125\left[\frac{\frac{1}{4}\hat{\pi}_b}{\frac{1}{2}+\frac{1}{4}\hat{\pi}_b}\right]$$

$$\pi_c = \frac{\hat{x}_2^b + 34}{\hat{x}_2^b + 18 + 20 + 34}.$$

If I substitute:

$$\hat{x}_2^p = 125\left[\frac{\frac{1}{4}\hat{\pi}_p}{\frac{1}{2}+\frac{1}{4}\hat{\pi}_p}\right]$$

into

$$\pi_{p+1} = \frac{\hat{x}_2^p + 34}{\hat{x}_2^p + 18 + 20 + 34}$$

and let $\pi_{p+1} = \pi_p$ then the result is a quadratic equation in π.

$$\pi = \frac{125\left[\frac{\frac{1}{4}\pi}{\frac{1}{2}+\frac{1}{4}\pi}\right] + 34}{125\left[\frac{\frac{1}{4}\pi}{\frac{1}{2}+\frac{1}{4}\pi}\right] + 18 + 20 + 34}$$

or

$$\pi = \frac{125\left[\frac{1}{4}\pi\right] + 34\left(\frac{1}{2} + \frac{1}{4}\pi\right)}{125\left[\frac{1}{4}\pi\right] + (18 + 20 + 34)\left(\frac{1}{2} + \frac{1}{4}\pi\right)}$$

Combining terms of π:

$$197\pi^2 - 15\pi - 68 = 0.$$

Using the quadratic formula yields:

$$\frac{15 \pm \sqrt{225 + 53584}}{394} \approx .6268.$$

I have just solved for the maximum likelihood estimate; .6268 is the actual π. Testing the accuracy of the EM algorithm against this value, I set the initial estimate of π, π_a to equal to ½. Using the scheme outlined above, five cycles of the EM algorithm yield the following results:

$\pi_b = .60824$
$\pi_c = .624321$
$\pi_d = .6264888$
$\pi_e = .626777$
$\pi_f = .626815$

After only three cycles, the algorithm's output is accurate to within three decimal places. This highly encouraging result is only a simplistic example of how powerful the EM algorithm is at finding maximum likelihood estimates.

If the distribution is unimodal and mild regularity conditions hold, then the parameter yielded by the EM algorithm obtains either a global or local maximum. Successive steps in the iteration always increase the likelihood function. Maximum likelihood estimates of the parameters using the EM algorithm are consistent and asymptotically normal.

Appendix B

Arbitrage Pricing Equations

This appendix shows the derivation for the valuation equations obtained through standard arbitrage portfolio methods.

B.1 One factor model

The commodity spot price follows a mean reverting process

$$dS(t) = k(\mu - \ln S(t))S(t)dt + \sigma S(t)dZ_s(t) \tag{B1}$$

where k is the speed of adjustment parameter, μ is the long run expected return for the spot commodity, σ is the diffusion coefficient, and $dZ_s(t)$ is the increment of a standard Brownian motion. Assume the price *F*(S,*t*) of futures contract is a twice continuously differentiable function of *S*(*t*), we can use Ito's lemma to define the instantaneous price change. The notation on the *F*(S,*t*) is suppressed for the remainder of the analysis

$$dF = F_S dS(t) + \frac{1}{2}F_{SS}[dS(t)]^2 - F_\tau dt \quad . \tag{B2}$$

where τ = (*T*-*t*) is the term to maturity. Substituting in (B2) for $dS(t)$ and $[dS(t)]^2$ yields

$$dF = F_S(k(\mu - \ln S(t))S(t)dt + \sigma S(t)dZ_s(t)) + \frac{1}{2}F_{SS}\sigma^2[S(t)]^2 dt - F_\tau dt \quad . \tag{B3}$$

Rearranging terms in (B4) yields

$$dF = \left(\frac{1}{2}F_{SS}\sigma^2[S(t)]^2 + F_S k(\mu - \ln S(t))S(t) - F_\tau\right)dt + \sigma S(t)F_S dZ_s(t) \quad . \tag{B4}$$

Expressing as the instantaneous return yields

$$\frac{dF}{F} = \gamma dt + s dZ_s(t) \tag{B5}$$

where $\gamma = \left(\frac{1}{2}F_{SS}\sigma^2[S(t)]^2 + F_S k(\mu - \ln S(t))S(t) - F_\tau\right)/F$ and $s = \sigma S(t)F_S / F$.

Under standard arbitrage assumptions, we can construct an arbitrage portfolio with two time varying futures contracts. For ease of exposition, we express them only in terms of their respective maturities, $F(1)$ and $F(2)$. The arbitrage portfolio is given as

$$P = x_1 F(1) + x_2 F(2). \qquad \text{(B6)}$$

We may express the instantaneous return for the portfolio as

$$\frac{dP}{P} = x_1 \frac{dF(1)}{F(1)} + x_2 \frac{dF(2)}{F(2)}. \qquad \text{(B7)}$$

Substituting the corresponding instantaneous returns for the futures contracts from equation (B5) into (B7) yields

$$\frac{dP}{P} = x_1(\gamma_1 dt + s_1 dZ_s(t)) + x_2(\gamma_2 dt + s_2 dZ_s(t)).$$

Rearranging terms in the expression above

$$\frac{dP}{P} = (x_1\gamma_1 + x_2\gamma_2)dt + (x_1 s_1 + x_2 s_2)dz_t. \qquad \text{(B8)}$$

For the rate of return on the portfolio to non-stochastic (riskless) the coefficients on $dZ_s(t)$ must be equal to zero. That is

$$x_1 s_1 + x_2 s_2 = 0. \qquad \text{(B9)}$$

In addition, with no initial investment the rate of return must be equal

$$x_1\gamma_1 + x_2\gamma_2 = 0. \qquad \text{(B10)}$$

From the arbitrage pricing theory (APT) we know the following

$$\sum x_i = 0 \qquad \sum x_i s_i = 0 \qquad \sum x_i \gamma_i = 0.$$

With the above APT assumptions we way write equations (B9) and (B10) in matrix notation

$$\begin{bmatrix} \gamma_1 & \gamma_2 \\ s_1 & s_2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\mathbf{Ax} = \mathbf{0} .$$

The linear system above is a homogeneous system. One of two solutions exist for this system. The first is the trivial solution where **x** = 0. The other is the non-trivial solution. The existence of a non-trivial solution implies the row vectors of A are linear dependent. Thus, algebraically it follows

$$\gamma = \lambda s \ . \tag{B11}$$

where γ and s are 1x2 vectors. Substituting in the partial differential equation, (B11), for γ and s yields

$$\left(\frac{1}{2}F_{SS}\sigma^2[S(t)]^2 + F_S k(\mu - \ln S(t))S(t) - F_\tau\right)/F = \lambda\sigma S(t)F_S/F . \tag{B12}$$

Simplifying,

$$\frac{1}{2}F_{SS}\sigma^2[S(t)]^2 + F_S k(\mu - \lambda\sigma - \ln S(t))S(t) - F_\tau = 0 \ . \tag{B13}$$

B.2 Gibson and Schwartz (1990) two factor model

The presentation is the two factor model derived Gibson and Schwartz (1990). The first factor is the spot price of the commodity and the second factor is the instantaneous convenience yield, δ. These factors are assumed to follow the joint stochastic process:

$$dS = \mu S dt + \sigma_1 S dZ_1 \tag{B14}$$

$$d\delta = k(\alpha - \delta)dt + \sigma_2 dZ_2 \tag{B15}$$

where the increments to the standard Brownian motion are correlated

$$dZ_1 dZ_2 = \rho dt, \tag{B16}$$

α is the long run convenience yield, k is the speed of adjustment for δ around its long run mean. Equation (B14) is a standard geometric Brownian motion characterizing the commodity spot price, while the stochastic convenience yield described in equation (B15), follows an Ornstein-Uhlenbeck stochastic process. The system describe above is similar to Brennan and Schwartz (1979) and Brennan and Schwartz (1982).

Assume the price $B(S,\delta,t)$ of an contingent claim is a twice continuously differentiable function of S and δ, we can use Ito's lemma to define the instantaneous price change.

$$dB = B_S dS + B_\delta d\delta - B_\tau dt + \frac{1}{2} B_{SS} dS^2 + \frac{1}{2} B_{\delta\delta} d\delta^2 + B_{s\delta} dS d\delta. \tag{B17}$$

Substituting in equation (B17), the expressions for dS, $d\delta$, dS^2 and $d\delta^2$ we obtain

$$= B_S(\mu S dt + \sigma_1 S dZ_1) + B_\delta(k(\alpha - \delta)dt + \sigma_2 dZ_2) - B_\tau dt + \frac{1}{2} B_{SS}\sigma_1^2 S^2 dt$$
$$+ \frac{1}{2} b_{\delta\delta}\sigma_2^2 dt + b_{s\delta} S\rho\sigma_1\sigma_2 dt$$

where τ is the time to maturity.

Rearranging the above

$$= \left[B_S \mu S + B_\delta k(\alpha - \delta) - B_\tau + \frac{1}{2} B_{SS}\sigma_1^2 S^2 + \frac{1}{2} B_{\delta\delta}\sigma_2^2 + B_{s\delta} S\rho\sigma_1\sigma_2 \right] dt$$
$$+ \sigma_1 S B_S dZ_1 + \sigma_2 B_\delta dZ_2 \quad . \tag{B18}$$

Now rewriting the above in the form of the instantaneous return we get

$$\frac{dB}{B} = \gamma dt + s_1 dZ_1 + s_2 dZ_2 \tag{B19}$$

where $\gamma = \left[B_S \mu S + B_\delta k(\alpha - \delta) - B_\tau + \frac{1}{2} B_{SS} \sigma_1^2 S^2 + \frac{1}{2} B_{\delta\delta} \sigma_2^2 + B_{s\delta} S \rho \sigma_1 \sigma_2 \right] / B$

$$s_1 = \frac{\sigma_1 S B_S}{B}$$

and

$$s_2 = \frac{\sigma_2 B_\delta}{B}.$$

Under standard arbitrage assumptions, we can construct an arbitrage portfolio with three time varying contingent claims. For ease of exposition, I suppress the notation for the contingent claims and express them only in terms of their respective maturities, $B(1)$, $B(2)$ and $B(3)$. The arbitrage portfolio is given as

$$P = x_1 B(1) + x_2 B(2) + x_3 B(3). \tag{B20}$$

We may express the instantaneous return for the portfolio as

$$\frac{dP}{P} = x_1 \frac{dB(1)}{B(1)} + x_2 \frac{dB(2)}{B(2)} + x_3 \frac{dB(3)}{B(3)}. \tag{B21}$$

Substituting the corresponding instantaneous returns for the contingent claims from equation (B19) into (B21) yields

$$\frac{dP}{P} = x_1(\gamma_1 dt + s_{11} dZ_1 + s_{21} dZ_2) + x_2(\gamma_2 dt + s_{12} dZ_1 + s_{22} dZ_2) + x_3(\gamma_3 dt + s_{13} dZ_1 + s_{23} dZ_2).$$

Rearranging terms in the expression above

$$\frac{dP}{P} = (x_1\gamma_1 + x_2\gamma_2 + x_3\gamma_3)dt + (x_1 s_{11} + x_2 s_{12} + x_3 s_{13})dZ_1 + (x_1 s_{21} + x_2 s_{22} + x_3 s_{23})dZ_2. \tag{B22}$$

For the rate of return on the portfolio to non-stochastic (riskless) the coefficients on dZ_1 and dZ_2 must be equal to zero. That is

$$x_1 s_{11} + x_2 s_{12} + x_3 s_{13} = 0 \tag{B23}$$

$$x_1 s_{21} + x_2 s_{22} + x_3 s_{23} = 0. \tag{B24}$$

In addition, if the rate of return is non-stochastic then it must be equal to the riskless rate of return, r

$$x_1 \gamma_1 + x_2 \gamma_2 + x_3 \gamma_3 = r. \tag{B25}$$

From the arbitrage pricing theory (APT) we know the following

$$\sum x_i = 1 \qquad \sum x_i s_{1i} = \sum x_i s_{2i} = 0$$

$$\sum x_i \gamma_i = r \quad \text{or} \quad \sum x_i (\gamma_i - r) = 0.$$

With the above APT assumptions we way write equations (B23), (B24) and (B25) in matrix notation

$$\begin{bmatrix} \gamma_1 - r & \gamma_2 - r & \gamma_3 - r \\ s_{11} & s_{12} & s_{13} \\ s_{21} & s_{22} & s_{23} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\mathbf{Ax} = \mathbf{0}.$$

One plausible solution to the above is the trivial solution

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

This solution, however, is not feasible because we require some investment. The other solution to the homogeneous system is the nontrivial solution. For a homogeneous system to have a nontrivial solution the coefficient matrix must equal zero. This implies that the row vectors in the coefficient matrix are linear dependent. Thus, we can express as,

$$\gamma - r = \lambda_1 s_1 + \lambda_2 s_2 \tag{B26}$$

where γ, s_1 and s_2 are 1x3 vectors.

Now substitution of γ, s_1 and s_2 into equation (B26) yields

$$\left\{\left[B_S\mu S + B_\delta k(\alpha-\delta) - B_\tau + \frac{1}{2}B_{SS}\sigma_1^2 S^2 + \frac{1}{2}B_{\delta\delta}\sigma_2^2 + B_{s\delta}S\rho\sigma_1\sigma_2\right]/B\right\} - r =$$
$$\lambda_1 \frac{\sigma_1 S B_S}{B} + \lambda_2 \frac{\sigma_2 B_\delta}{B}.$$

(B27)

We can reduce the above by solving for λ_1. The spot contract ($B(S,0)=S$), must also satisfy equation (B26). The total expected return μ_s to the owner of the commodity derives from two sources, namely the convenience yield δ and the expected spot price change μ. One can define the market price per unit of spot price risk λ_1 by solving the partial differential equation for *S*. Thus, if $B(S,0)=S$ and the partial derivatives $B_S=1$ and $B_\delta=0$, then we may rewrite (B26) as follows

$$\mu_S - r = \lambda_1\left(\frac{S1\sigma_1}{S}\right) + \lambda_2(0)$$

$$\mu_S - r = \lambda_1\sigma_1$$

$$\lambda_1 = \frac{\mu_S - r}{\sigma_1}.$$

Further, if the total expected return of the spot price derives from two sources, the convenience yield and the expected price change μ, then we can derive λ_1 as

$$\lambda_1 = \frac{(\mu+\delta)-r}{\sigma_1}. \quad \text{(B28)}$$

Substituting equation (B28) into equation (B27) yields

$$\left[B_S\mu S + B_\delta k(\alpha-\delta) - B_\tau + \frac{1}{2}B_{SS}\sigma_1^2 S^2 + \frac{1}{2}B_{\delta\delta}\sigma_2^2 + B_{s\delta}S\rho\sigma_1\sigma_2\right] - rB =$$

$$\left(\frac{(\mu+\delta)-\mathrm{r}}{\sigma_1}\right)\sigma_1 SB_S + \lambda_2\sigma_2 B_\delta .$$

Moving the terms on the right hand side to the left hand side and rearranging yields

$$\frac{1}{2}B_{SS}\sigma_1^2 S^2 + B_{s\delta}S\rho\sigma_1\sigma_2 + \frac{1}{2}B_{\delta\delta}\sigma_2^2 + B_S S(\mathrm{r}-\delta) + B_\delta\big(k(\alpha-\delta) - \lambda_2\sigma_2\big) - B_\tau - rB = 0 \; . \tag{B29}$$

Equation (B29) is the valuation equation. For a futures contract the partial differential equation is

$$\frac{1}{2}F_{SS}\sigma_1^2 S^2 + F_{s\delta}S\rho\sigma_1\sigma_2 + \frac{1}{2}F_{\delta\delta}\sigma_2^2 + F_S S(\mathrm{r}-\delta) + F_\delta\big(k(\alpha-\delta) - \lambda_2\sigma_2\big) - F_\tau = 0 \; . \tag{B30}$$

Appendix C

Risk Neutral Pricing Methods: Partial Differential Equations and Equivalent Martingales

With the use of stochastic calculus, modern finance has developed two major methods for pricing contingent claims. The first and better known method utilizes a partial differential equation. The second method (which is used in my analysis) requires transforming underlying processes into martingales. Both methods are addressed in this appendix.

C.1 Partial Differential Equations

We begin with a general discussion to the partial differential equations approach to pricing contingent claims. The derivations for the various futures pricing models used in the analysis are shown in appendix B. These derivations show how the dynamics of the underlying state variables and a no arbitrage condition produce a partial differential equation that the futures price must follow. Given this expression, we use some necessary boundary conditions to determine the pricing formula for a particular futures price. The purpose of this section is to intuitively describe the above method of solution used for pricing a contingent claim.

The intent of our analysis is to determine the price of a derivative security and how its behaves over time. For our purpose here we will determine the value of a call option whose price is solely determined by an underlying stock price. With this said, for us to say something about a call option, we must know how the underlying stock price behaves. Therefore, we begin by positing a model that characterizes the dynamics of the underlying stock price, $S(t)$, and from there we determine the dynamic behavior of the call option. Accordingly, we assume that the stochastic differential of the stock price, $dS(t)$, obeys the stochastic differential equation

$$dS(t) = \mu S(t)dt + \sigma S(t)dZ(t) \qquad t \in [0,\infty), \tag{C1}$$

where μ is the instantaneous return on the stock and σ is the diffusion coefficient. $dZ(t)$ is the standard increment of a Brownian motion. Suppose the option price can be written as a twice-continuously differentiable function of the stock price and time, namely $F(S,t)$. If the stock price follows the dynamics described in (C1), then the option's return dynamics can be written in a similar form as

$$dF(t) = \mu_F F(S,t)dt + \sigma_F F(S,t)dZ(t), \tag{C2}$$

where μ_F is the instantaneous return on the option and σ_F is its diffusion coefficient. Note, $F(S(t),t)$ is the notation used to denote the derivative security's price written on the underlying stock. $dF(t)$ is the total change in the option's price over an infinitesimal interval, and F_t is the partial derivative of $F(S(t),t)$ with respect to time.

Using Ito's lemma and the stock price dynamics given in (C1), we may alternatively express the increment for the derivative security as

$$dF(t) = F_s dS(t) + \frac{1}{2}F_{ss}[dS(t)]^2 + F_t dt.$$

Substituting in for the stochastic differentials and rearranging yields,

$$dF(t) = F_t dt + \frac{1}{2}F_{ss}\sigma^2 dt + F_s(\mu S(t)dt + \sigma S(t)dZ(t))$$

$$dF(t) = \left[F_s \mu S(t) + \frac{1}{2}F_{ss}\sigma^2[S(t)]^2 + F_t\right]dt + F_s \sigma S(t)dW(t). \tag{C3}$$

If we know the functional form for $F(S(t),t)$ then we could take the partial derivatives F_t, F_s and F_{ss} to obtain the exact stochastic differential equation that governs the dynamics of the option price. The functional form is not known, however, we can determine it.

Given the expressions in (C2) and (C3) we can define the drift and diffusion terms for the option in (C2) as

$$\mu_F = \left[\frac{1}{2}F_{ss}\sigma^2 S^2 + \mu S F_s + F_t\right] \Big/ F, \tag{C4}$$

$$\sigma_F = \sigma S F_s / F \tag{C5}$$

Notice from equation (C3), the same Weiner increment that drives $S(t)$ also describes the movements in $F(S(t),t)$. We would expect this as $F(S(t),t)$ derives its value from $S(t)$. Since $dF(t)$ and $dS(t)$ have the same source of underlying uncertainty we can form a risk free portfolio in continuous time. Let $P(t)$ dollars be invested in the stock, $S(t)$, the option, $F(S(t),t)$, and a riskless asset with a return r per unit time. The portfolio's investment strategy is in portions $w_1, w_2,$ and w_3 of the assets above respectively, where $\sum_{j=1}^{3} w_j = 1$. The return dynamics to the portfolio are expressed as:

$$dP(t) = \mu_P P(t)dt + \sigma_P P(t)dZ(t), \tag{C6}$$

where μ_P is the instantaneous return on the portfolio and σ_P is the diffusion coefficient for the portfolio. The value of this portfolio changes as time passes due to changes in $F(S(t),t)$ and of $S(t)$. The drift and diffusion terms for the portfolio are linear combinations of the drift and diffusion terms of the individual assets held in the portfolio. The drift term is

$$\mu_P = w_1\mu + w_2\mu_F + w_3\mu_r .$$

Now using the constraint $w_1 + w_2 + w_3 = 1$ and the fact that the expected return on the riskless asset is equal to r we obtain

$$\mu_P = w_1\mu + w_2\mu_F + (1 - w_1 + w_2)r ,$$

$$\mu_P = w_1(\mu - r) + w_2(\mu_F - r) + r \tag{C7}$$

The diffusion coefficient for the portfolio is

$$\sigma_P = w_1\sigma + w_2\sigma_F + w_3\sigma_r .$$

By definition $\sigma_r = 0$ so we are left with

$$\sigma_P = w_1\sigma + w_2\sigma_F \tag{C8}$$

If the portfolio in expression (C6) is risk-free, this means $\sigma_P = 0$. If the portfolio is riskless and there are no arbitrage opportunities then the drift term for this portfolio must equal the risk-free rate of return. That is, $\mu_P = r$. Therefore, we can write expressions (C7) and (C8) as

$$\mu_P = w_1(\mu - r) + w_2(\mu_F - r) + r = r$$

$$\sigma_P = w_1\sigma + w_2\sigma_F = 0 ,$$

or

$$\mu_P - r = w_1(\mu - r) + w_2(\mu_F - r) + r = 0 \tag{C9}$$

$$\sigma_P = w_1\sigma + w_2\sigma_F = 0 , \tag{C10}$$

We may express the above system of equations in the following form

$$\begin{bmatrix} \mu - r & \mu_F - r \\ \sigma & \sigma_F \end{bmatrix} \begin{bmatrix} w_1^* \\ w_2^* \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \tag{C11}$$

The system of equations in (C11) is known as a homogeneous equation system. There are two possible solutions to this system. The first is called the trivial solution where the weights are all equal to zero, $w_i's = 0$. This solution, however, is not a viable solution since by definition there must be some investment. The second solution and more appropriate is the non-trivial solution where the portfolio weights are all non zero, $w_i's \neq 0$. The only way to obtain a nontrivial solution from a homogeneous system of equations is if the coefficient matrix is singular. That is,

$$\begin{vmatrix} \mu - r & \mu_F - r \\ \sigma & \sigma_F \end{vmatrix} = 0.$$

The above condition implies that the row vector $[\mu - r, \mu_F - r]$ is a multiple of the row vector $[\sigma, \sigma_F]$; consequently one of these two equations is redundant. Dealing with the first equation we see

$$w_1^*(\mu - r) + w_2^*(\mu_F - r) = 0,$$

$$w_1^*(\mu - r) = -w_2^*(\mu_F - r),$$

$$w_1^* = -w_2^* \frac{(\mu_F - r)}{(\mu - r)}. \qquad \text{(C12)}$$

Now substituting this expression into the second equation we have

$$\left(-w_2^* \frac{(\mu_F - r)}{(\mu - r)}\right)\sigma + w_2^* \sigma_F = 0,$$

$$w_2^*\left[\left(-\frac{(\mu_F - r)}{(\mu - r)}\right)\sigma + \sigma_F\right] = 0. \qquad \text{(C13)}$$

We are looking for the non-trivial solution $w_i^* \neq 0$. The only way the equation above equals zero is if the expression inside the brackets equals zero. This implies

$$\left(-\frac{(\mu_F - r)}{(\mu - r)}\right)\sigma + \sigma_F = 0,$$

$$\left(\frac{(\mu_F - r)}{(\mu - r)}\right)\sigma = \sigma_F,$$

$$\sigma(\mu_F - r) = \sigma_F(\mu - r),$$

$$\frac{(\mu_F - r)}{\sigma_F} = \frac{(\mu - r)}{\sigma}. \quad \text{(C14)}$$

The non-trivial solution for the homogeneous system of equations in expression (12) is

$$w_i^* = \frac{\gamma_i - r}{s_i}, \quad \text{(C15)}$$

where $\gamma_i = [\mu, \mu_F]$ and $s_i = [\sigma, \sigma_F]$. To illustrate the above consider the expression

$$w_1^* = -w_2^* \frac{(\mu_F - r)}{(\mu - r)}.$$

Substitute this expression into

$$w_1^*\sigma + w_2^*\sigma_F = 0,$$

$$-w_2^* \frac{(\mu_F - r)}{(\mu - r)}\sigma + w_2^*\sigma_F = 0. \quad \text{(C16)}$$

If $w_2^* = \frac{(\mu_F - r)}{\sigma_F}$, then expression (15) becomes

$$-\left(\frac{(\mu_F - r)}{\sigma_F}\right)\frac{(\mu_F - r)}{(\mu - r)}\sigma + \left(\frac{(\mu_F - r)}{\sigma_F}\right)\sigma_F = 0,$$

$$-\frac{(\mu_F - r)(\mu_F - r)}{\sigma_F(\mu - r)}\sigma + \left(\frac{(\mu_F - r)}{\sigma_F}\right)\sigma_F = 0,$$

$$-\frac{(\mu_F - r)^2}{\sigma_F^2} + \left(\frac{(\mu_F - r)}{\sigma_F}\right)\frac{(\mu - r)}{\sigma} = 0. \quad \text{(C17)}$$

Recall from expression (C14), that

$$\frac{(\mu_F - r)}{\sigma_F} = \frac{(\mu - r)}{\sigma}.$$

Therefore, expression (16) reduces to

$$-\frac{(\mu_F - r)^2}{\sigma_F^2} - \frac{(\mu - r)^2}{\sigma^2} = 0.$$

The result above shows that the optimal weighting scheme,

$$w_i^* = \frac{\gamma_i - r}{s_i}, \qquad \text{(C18)}$$

Now substituting expressions (C4) and (C5) into expression (C18) yields

$$\frac{\mu - r}{\sigma} = \frac{\left[\frac{1}{2}F_{ss}\sigma^2[S(t)]^2 + \mu S(t)F_s + F_t\right]/F(S(t),t) - r}{\sigma S(t)F_s/F(S(t),t)},$$

$$\frac{(\mu - r)\sigma S(t)F_s}{\sigma} = \frac{1}{2}F_{ss}\sigma^2[S(t)]^2 + \mu S(t)F_s + F_t - F(S(t),t)r,$$

$$(\mu - r)S(t)F_s = \frac{1}{2}F_{ss}\sigma^2[S(t)]^2 + \mu S(t)F_s + F_t - F(S(t),t)r,$$

$$\frac{1}{2}F_{ss}\sigma^2[S(t)]^2 + F_s rS(t) + F_t - F(S(t),t)r = 0. \qquad \text{(C19)}$$

Equation (C19) is the partial differential equation for pricing call option written on the underlying stock whose price follows (C1). The solution to this equation, $F(S(t),t)$, is the unknown we wish to find. While the exact form of $F(S(t),t)$ is not known, we do know that the linear combination of the partial derivatives is equal to zero (expression (C19)). In addition, at time T we know that $F(S(t),T)$ must equal some known value $G(S(t),T)$, which is called a boundary condition. This value indicates some plausible condition that must be satisfied by $F(S(t),T)$ at time T. Once the necessary boundary conditions are stated, standard methods of solution exists for solving (C19). One well known example of the partial differential equation pricing method is the Black-Scholes option pricing model. While the example is illustrated for pricing options on stocks the analysis can be easily generalized to price other contingent claims written on any underlying asset.

The astute reader should notice that the solution, $F(S(t),t)$, to expression (C19) is also the solution to equation (C3). The question is, what do we gain from using expression (C19) over equation (C3). One, we invoke a no arbitrage condition to find a deterministic equation describing the expected movement of the derivative security. Two, as mentioned above, standard well known techniques exist for solving these equations. However, it is possible to solve the stochastic differential equation in equation (C3) and this method is discussed next.

C.2 *Equivalent Martingale Measures*

Recent methods of derivative asset pricing do not necessarily exploit partial differential equations implied by arbitrage free portfolios. They implicitly use a non-arbitrage condition to convert prices of such assets into martingales. This is done through risk-adjusting the probability distributions of an underlying diffusion process using the tools provided by the Girsanov theorem. Once the distribution for the process is transformed the asset becomes a martingale. Thus, we will show that the fair market value of the commodity today will equal its risk-adjusted expected value.

Changing the means

There are two ways one can change the mean of a random variable. The first method involves operating on the realizations of the random variable. The second, and counterintuitive method, leaves the realizations of the random variable unchanged, and operates on the probabilities associated with the realizations of the random variable. Both operations lead to a change in the original mean, while preserving other characteristics of the original random variable.

Example 1: Operating on the individual realizations

The most common method used in econometrics and statistics for changing the mean of a random variable is to simply add a constant to the random variable. For example, let X denote a random variable. The expected value of X is

$$E[X]=0. \tag{C20}$$

We can alter the mean of X by creating a new variable Z. Let $Z = X + a$. Given the new random variable Z, the expectation will be such that

$$E[Z]=E[X]+a=a. \tag{C21}$$

A simple example will illustrate the point above. Let X have the following distribution

$$x_1 = 10 \quad f(x_1)=\frac{1}{3}$$
$$x_2 = -3 \quad f(x_2)=\frac{1}{3}$$
$$x_3 = -1 \quad f(x_3)=\frac{1}{3}$$

We can calculate the expected value of X as a weighted average of its possible values:

$$E[X]=\frac{1}{3}[10]+\frac{1}{3}[-3]+\frac{1}{3}[-1]=2. \tag{C22}$$

Now, suppose we would like to change the mean of X using the method outlined above. More precisely, suppose we would like to calculate a new random variable with the same variance but with a new mean of one. We call this random variable Z and let

$$Z = X - 1. \tag{C23}$$

Using the formula in (C22) but with Z instead of X we have

$$E[X] = \frac{1}{3}[10-1] + \frac{1}{3}[-3-1] + \frac{1}{3}[-1-1] = 1. \tag{C24}$$

Did the variance stay the same? The variance for X is

$$V[X] = \frac{1}{3}[10-2]^2 + \frac{1}{3}[-3-2]^2 + \frac{1}{3}[-1-2]^2 = \frac{98}{3}. \tag{C25}$$

The variance of Z is

$$V[Z] = \frac{1}{3}[9-1]^2 + \frac{1}{3}[-4-1]^2 + \frac{1}{3}[-2-1]^2 = \frac{98}{3}. \tag{C26}$$

We were able to alter the mean for the random variable while keeping the variance unchanged. However, we can accomplish this objective in a different fashion. That is, instead of operating on the random variable itself we can alter the mean by transforming the distribution of the random variable.

Example 2: Operating on the Probability Distribution

Consider the first example. X is defined

$$x_1 = 10 \quad f(x_1) = \frac{1}{3}$$
$$x_2 = -3 \quad f(x_2) = \frac{1}{3}$$
$$x_3 = -1 \quad f(x_3) = \frac{1}{3}$$

It is clear that $E(X) = 2$ and the $V(X) = \frac{98}{3}$. Now we want to transform X so that its mean becomes one, while leaving the variance unchanged. To find a new set of probabilities we can use the following information

$$E[X] = f(x_1^*)[10] + f(x_2^*)[-3] + f(x_3^*)[-1] = 1 \qquad \text{(C27)}$$

$$V[X] = f(x_1^*)[10-1]^2 + f(x_2^*)[-3-1]^2 + f(x_3^*)[-1-1]^2 = \frac{98}{3} \qquad \text{(C28)}$$

$$f(x_1^*) + f(x_2^*) + f(x_3^*) = 1 \quad . \qquad \text{(C29)}$$

The system of equations can be rewritten as

$$\begin{bmatrix} 10 & -3 & -1 \\ 81 & 16 & 4 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} f(x_1^*) \\ f(x_2^*) \\ f(x_3^*) \end{bmatrix} = \begin{bmatrix} 1 \\ 98/3 \\ 1 \end{bmatrix}$$

Solving the system above for the probabilities yields

$$f(x_1^*) = \frac{122}{149},\ f(x_2^*) = \frac{22}{39} \text{ and } f(x_3^*) = \frac{5}{33}.$$

Now calculating the mean of X under the new probabilities yields

$$E^*[X] = \frac{122}{149}[10] + \frac{22}{39}[-3] + \frac{5}{33}[-1] = 1. \qquad \text{(C30)}$$

The variance of X under the new probabilities is

$$V^*[X] = \frac{122}{149}[10-1]^2 + \frac{22}{39}[-3-1]^2 + \frac{5}{33}[-1-1]^2 = \frac{98}{3}. \qquad \text{(C31)}$$

Just as we did in the first example, we transformed the mean of X. The key here is we accomplished our objective by transforming the probability distribution for X, and not its realizations.

One may wonder how the above method is useful in valuing financial assets. Consider the following. Let r be the risk-free rate of return. A typical risky asset $S(t)$ must offer a rate of return R greater than r on average otherwise there will be no reason to hold the risky asset. We may write this as

$$E_t[S(t+1)] > (1+r)S(t). \tag{C32}$$

On the average, the risky asset will appreciate faster than the growth of a risk-free investment. This inequality can be rewritten as

$$\frac{1}{(1+r)} E_t[S(t+1)] > S(t). \tag{C33}$$

Here the left-hand side represents the expected future price discounted at the risk-free rate. For some $\lambda > 0$,

$$\frac{1}{(1+r)} E_t[S(t+1)] = S(t)(1+\lambda). \tag{C34}$$

Note that the positive constant can be interpreted as a risk premium. Transforming the above

$$\frac{E_t[S(t+1)]}{S(t)} = (1+r)(1+\lambda). \tag{C35}$$

The term on the left-hand side of this equation represents the expected gross return, $E_t[1 + R]$. This means that

$$E_t(1+R) = (1+r)(1+\lambda), \tag{C36}$$

which says that the expected return on a risky asset must exceed the risk-free return approximately by λ:

$$E_t(R) \cong r+\lambda, \tag{C37}$$

in the case where r and λ are small enough that the cross product term can be ignored.

Under these conditions, λ is the risk premium for holding the asset for one period, and $\frac{1}{1+r}$ is the risk-free discount factor. The problem of the financial analyst is to obtain the fair market value of the asset today. That is, the analyst would like to calculate $S(t)$. One way to do this is to exploit the relation

$$E_t\left[\frac{1}{(1+R)}S(t+1)\right]=S(t) \tag{C38}$$

by calculating the expectation on the left-hand side. Evaluating the expectation in (C38), however, requires knowledge of the distribution of R, and this requires knowledge of the asset's risk premium λ. The problem is the investor rarely, if ever, knows this value of λ before obtaining the fair market value of the asset. Therefore, implementing (C38) will go nowhere in terms of calculating $S(t)$. .

In theory, we would like to price assets according to expression (C38). We have seen that the inability to properly identify an asset's expected return leaves us to find an alternative method for pricing assets. Examination of expression (C37) shows that the expected return of an asset is a function of the risk-free rate of return and a risk premium. If we are capable of purging the latent risk premium from the asset's expected return, then the risky asset's expected return would equal the known risk-free rate of return. Subsequently, equation (C38) would then be a viable method for pricing a risky asset.

As we discussed earlier, we know it is possible to alter the mean of a random variable by altering its probability distribution. Asset returns are stochastic and presumed to follow a log-normal distribution. If we possessed knowledge of the return distribution, we could then use an equivalent probability measure to evaluate the return of the asset. The transformation would allow us to purge the risk premium from the asset's expected return. The idea is to alter the distribution of R to equate the expected return to the risk free rate of return. In doing so we have taken the risk premium out of the asset and we may now use the known risk-free rate to discount the transformed forecast of the next period's asset price. The new expression is

$$E_t^*\left[\frac{1}{(1+r)}S(t+1)\right]=S(t). \tag{C39}$$

The process provides us with an $I(t)$-adapted forecast of today's asset price. This value is equal to the fair market value of the asset. That is, this price reflects what the asset should sell for today. The forecasted fair market value could be different from the actual price in the market. If so arbitrage opportunities exist and agents would to act opportunistically.

We can demonstrate the notion above with a heuristic example. Consider a risky asset whose fair market price is $S(t)=100$. Over the next interval of time, dt, the assets price will take on one of three values $S(t+dt)=100$, $S(t+dt)=110$ or $S(t+dt)=120$. Each realization for the asset's spot price is equally likely. In addition to the risky asset, there is a risk-free asset that returns five percent over an interval of dt. Given the terminal payoffs for $S(t)$ and its probability distribution we know the expect payoff for $S(t+dt)$ is

$$E_t[S(t+dt)]=100\left(\frac{1}{3}\right)+110\left(\frac{1}{3}\right)+120\left(\frac{1}{3}\right)=110. \tag{C40}$$

We see the expected return on the risky asset is

$$E_t\left[\frac{S(t+dt)}{S(t)}\right]=\frac{100-100}{100}\left(\frac{1}{3}\right)+\frac{110-100}{100}\left(\frac{1}{3}\right)+\frac{120-100}{100}\left(\frac{1}{3}\right)=.10. \tag{C41}$$

Given the actual expected return, we may find the actual value of the asset using expression (C38). The asset price is equal to the next period expected price discounted by it expected return. This is

$$S(t) = \frac{E_t[S(t+dt)]}{1.10} = \frac{110}{1.10} = 100. \tag{C42}$$

Notice in order to find the fair market value of $S(t)$ we needed to know the true probability distribution for $S(t)$ in order to determine its expected return. In a financial market, it is rare for agents to know the expected return before we know the true value for $S(t)$. Therefore utilizing (C42) will typically go nowhere.

One the other hand, we can alter the mean of $S(t)$ without having to use the expected return for $S(t)$. That is, we will find a risk-adjusted probability distribution for $S(t)$ then find the risk neutral forecast for $S(t)$ and discount it by the known risk-free rate of return. This too should yield a fair market value for $S(t)$. To find the risk-adjusted probabilities we need to solve the following set of equations

$$0f(s_1^*) + 10f(s_2^*) + 20f(s_3^*) = 5 \tag{C43}$$

$$25f(s_1^*) + 25f(s_2^*) + 225f(s_3^*) = 66 \tag{C44}$$

$$f(s_1^*) + f(s_2^*) + f(s_3^*) = 1. \tag{C45}$$

Equation (C43) is the condition that the new mean of $S(t)$ under the risk-adjusted probabilities must equal the risk-free return. Equation (C44) states that the variance of the asset remains unchanged. Lastly, Expression (C45) is the constraint that the probabilities must sum to one.

Solving the system of equations for the risk adjusted probabilities yields

$$f(s_1^*) = .705,\ f(s_2^*) = .09,\ f(s_3^*) = .205.$$

The risk adjusted forecast for $S(t+dt)$ is

$$E_t^*[S(t+dt)] = 100(.705) + 110(.09) + 120(.205) = 105. \qquad \text{(C46)}$$

Discounting the forecast by the risk-free rate yields

$$S(t) = \frac{E_t^*[S(t+dt)]}{1.05} = \frac{105}{1.05} = 100. \qquad \text{(C47)}$$

The solution in (C47) is equal to the solution in (C42). We have found the fair market value for $S(t)$ and we did it using a risk-adjusted probability distribution. In addition, the variance of the asset has remained the same. That is,

$$V_t^*[S(t+dt)] = (100-105)^2(.705) + (110-105)^2(.09) + (120-105)^2(.205) = 66.$$

Thus far, all our examples have been for discrete random variables. Naturally, one would question if we can alter the probability distribution for a continuous random variable. The answer is yes.

Consider a normally distributed random variable $z(t)$:

$$z(t) \sim N(0,1). \qquad \text{(C48)}$$

The state space is continuous and the probability density $f(z(t))$ of this random variable is given by the well known expression

$$f(z(t)) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(z(t))^2}. \qquad \text{(C49)}$$

Suppose we are interested in the probability that $z(t)$ falls near a specific value $\bar{z}$. Then, this probability can be expressed by first choosing a small interval $\Delta > 0$, and next by calculating the integral of the normal density over the region in question

$$P\left(\bar{z} - \frac{1}{2}\Delta < z(t) < \bar{z} + \frac{1}{2}\Delta\right) = \int_{\bar{z}-\frac{1}{2}\Delta}^{\bar{z}+\frac{1}{2}\Delta} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(z(t))^2} dz(t). \qquad \text{(C50)}$$

Now, if the region around $\bar{z}$ is small then $f(z(t))$ will not change very much as $z(t)$ varies from $\bar{z}-\frac{1}{2}\Delta$ to $\bar{z}+\frac{1}{2}\Delta$. This means we can approximate $f(z(t))$ by $f(\bar{z})$ during this interval and write the integral on the right-hand side as

$$\int_{\bar{z}-\frac{1}{2}\Delta}^{\bar{z}+\frac{1}{2}\Delta} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(z(t))^2} dz(t) \cong \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\bar{z}^2} \int_{\bar{z}-\frac{1}{2}\Delta}^{\bar{z}+\frac{1}{2}\Delta} dz(t)$$

$$\cong \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\bar{z}^2} \Delta \qquad \text{(C51)}$$

The probability in equation (C51) is a "mass" represented by a rectangle with base Δ and height $f(\bar{z}(t))$. Visualized this way, probability corresponds to a measure that is associated with possible values of $z(t)$ in small intervals. Probabilities are called measures because they are mappings from arbitrary sets to nonnegative real numbers. For infinitesimal Δ, which we write as $dz(t)$, these measures are denoted by the symbol $dP(z(t))$, or

$$dP(\bar{z}) = P\left(\bar{z}-\frac{1}{2}dz(t) < z(t) < \bar{z}+\frac{1}{2}dz(t)\right). \qquad \text{(C52)}$$

This can be read as the probability that the random variable $z(t)$ will fall within a small interval centered on $\bar{z}$ and of infinitesimal length $dz(t)$. The sum of all such probabilities will then be given by adding these $dP(z(t))$ for various values of $\bar{z}$. Formally, this is expressed by the use of the integral

$$\int_{-\infty}^{\infty} dP(z(t)) = 1. \qquad \text{(C53)}$$

Given expression (C48), the probability measure for $z(t)$ is denoted as

$$dP(z(t)) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(z(t))^2} dz(t). \tag{C54}$$

Therefore we state that we have a normally distributed variable, $z(t)$, with a probability measure $dP(z(t))$.

The question remains, how do we transform the probability distribution of $z(t)$ to alter its mean. Consider the function

$$\xi(z(t)) = e^{z(t)\mu - \frac{1}{2}\mu^2}. \tag{C55}$$

If we multiply $dP(z(t))$ by $\xi(z(t))$, we obtain

$$dP^*(z(t)) = dP(z(t))\xi(z(t)) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(z(t))^2 + z(t)\mu - \frac{1}{2}\mu^2} dz(t),$$

$$dP^*(z(t)) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(z(t)-\mu)^2} dz(t). \tag{C56}$$

Integrating over expression (C51) yields

$$\int_{-\infty}^{\infty} dP^*(z(t)) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(z(t)-\mu)^2} dz(t) = 1.$$

From the integral above, we see that expression (C51) is also a probability measure. It turns out that by multiplying $dP(z(t))$ by $\xi(z(t))$, and then switching to P^*, we succeeded in changing the mean of $z(t)$. Note, that in this particular case, the multiplication by $\xi(z(t))$ preserved the shape of the probability measure. In fact, expression (C56) is still a bell shaped, Gaussian curve with the same variance. But $dP(z(t))$ and $dP^*(z(t))$ are different measures. They have different means and they assign different weights to intervals on the z-axis.

To illustrate the above discussion, we calculate the probability measures for $dP(z(t))$ and $dP^*(z(t))$ at various points for $z(t)$. In particular, we consider intervals for $z(t)$ equal to .01 or $\Delta = .01$, and we let $z(t) = 0$, $z(t) = 1.5$, and $z(t) = 3$. Under the true probability measure for $z(t)$ we have

$$dP(z(t)) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(z(t))^2} (.01).$$

Therefore, the values for $dP(z(t))$ at $z(t) = 0$, $z(t) = 1.5$, and $z(t) = 3$ are

$$dP(z(t) = 0) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(0)^2} (.01) = 0.0039904,$$

$$dP(z(t) = 1.5) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(1.5)^2} (.01) = 0.0012955,$$

$$dP(z(t) = 3) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(3)^2} (.01) = 0.0000443.$$

Under an alternative probability measure, where the mean is equal to three and the variance is one, we have

$$dP^*(z(t)) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(z(t)-3)^2} (.01).$$

The corresponding measures at $z(t) = 0$, $z(t) = 1.5$, and $z(t) = 3$ are

$$dP^*(z(t) = 0) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(0-3)^2} (.01) = 0.000443,$$

$$dP^*(z(t) = 1.5) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(1.5-3)^2} (.01) = 0.0012955,$$

$$dP^*(z(t) = 3) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(3-3)^2} (.01) = 0.0039904.$$

We know the correspondence between the two measures $dP(z(t))$ and $dP^*(z(t))$ is $\xi(z(t))$, which is

$$\xi(z(t)) = e^{z(t)\mu - \frac{1}{2}\mu^2}.$$

This function evaluated at $z(t) = 0$, $z(t) = 1.5$, and $z(t) = 3$ is

$$\xi(z(t) = 0) = e^{0(3) - \frac{1}{2}3^2} = 0.01110,$$

$$\xi(z(t) = 1.5) = e^{1.5(3) - \frac{1}{2}3^2} = 1.00000,$$

$$\xi(z(t) = 3) = e^{3(3) - \frac{1}{2}3^2} = 90.01713.$$

Now, these values for $\xi(z(t))$ alter the probability distribution for $z(t)$. That is, if we want the mean of $z(t)$ to be equal to three instead of zero, then we must change the measure of $z(t)$ at each value. Thus, if the mean of $z(t)$ is three the measure for $z(t)$ at $z(t) = 0$ must equal 0.0000443. But the true measure is 0.0039904. To obtain the desired probability measure we multiply the true measure, 0.0039904, by $\xi(z(t))$, 0.01110. The transformation yields a new measure for a normally distributed variable with mean of three and a variance of one. Notice that as $z(t)$ moves further to the right the adjustment factor moves as well. This occurs because the probability measures are changing as well.

From above we know the correspondence between the new probability measure $dP^*(z(t))$ and the original measure, $dP(z(t))$, is the function $\xi(z(t))$. We can give an interpretation to $\xi(z(t))$. Formally,

$$dP^* = \xi(t)dP. \qquad \text{(C57)}$$

We can rewrite this as

$$\frac{dP^*}{dP} = \xi(t), \qquad \text{(C58)}$$

which can be regarded as a derivative. This derivative is called the Radon-Nikodym derivative and reads as, the "derivative" of the measure P^* with respect to P is given by $\xi(t)$. Simply put $\xi(t)$ is a ratio of two probability measures for a given value of $z(t)$. We see that as $z(t)$ changes so will the probability measures. $\xi(t)$ captures this change, hence it is regarded as a derivative of one measure with respect to the other.

Note, in order to write the ratio in (C40) meaningfully, we need the probability mass in the denominator to be different from zero. To perform the inverse transformation, we also need the numerator to be different from zero. Recall the numerator and the denominator are probabilities assigned to infinitesimal intervals *dt*. Hence, in order for the Radon-Nikodym derivative to exist, when P^* assigns a nonzero probability to $dz(t)$, so must P, and vice versa. If this condition is satisfied, then $\xi(t)$ would exist, and we can always go back and forth between the two measures using the relations

$$dP^* = \xi(t)dP \qquad \text{(C59)}$$

and

$$dP = \xi(t)^{-1}dP^*. \qquad \text{(C60)}$$

This means that for all practical purposes, the two measures are equivalent, thereby, they are called equivalent probability measures.

Girsanov theorem

In our example above, we saw how to transform the probability distribution for a finite sequence of random variables. In applications of continuous time finance, the examples provided thus far would be of limited use. Continuous finance deals with continuous or right continuous stochastic processes and we need a method for altering the distribution for such a process. The Girsanov theorem provides the conditions under which the Radon-Nikodym derivative, $\xi(t)$, exists for cases where $z(t)$ is a continuous stochastic process. Transformations of probability measures in continuous finance use this theorem.

The Girsanov theorem states:

If the process $\xi(t) = e^{\int_0^t X(u)dW(u) - \frac{1}{2}\int_0^t X(u)^2 du}$ is a martingale with respect to information sets $I(t)$, and the probability P, then $W^*(t)$, defined by

$$W^*(t) = W(t) - \int_0^t X(u)du \qquad t \in [0,T], \tag{C61}$$

is a Wiener process with respect to $I(t)$ and with respect to the probability measure $P^*(T)$, given by

$$P^*(T, A) = E^*[1_A \xi(T)], \tag{C62}$$

with A being an event determined by $I(T)$ and 1_A being the indicator function of the event.

Consider the following heuristic example. Let $dS(t)$ denote incremental changes in an asset price. Assume that these changes are driven by infinitesimal shocks that have a normal distribution, so that we can represent $S(t)$ using the stochastic differential equation driven by the Wiener process $W(t)$

$$dS(t) = \mu dt + \sigma dW(t) \tag{C63}$$

$W(t)$ is assumed to have the probability distribution P, with

$$dP[W(t)] = \frac{1}{\sqrt{2\pi t}} e^{-\frac{1}{2}(W(t))^2} dW(t). \quad \text{(C64)}$$

Clearly $S(t)$ is not a martingale if the drift term μdt is nonzero. Taking the integral of (C64) from 0 to t we get

$$S(t) = S(0) + \mu \int_0^t ds + \sigma \int_o^t dW(s).$$

Now let $S(0) = 0$ and $W(0) = 0$, the above becomes

$$S(t) = \mu t + \sigma W(t).$$

If $s > t$ then we can write

$$E_t[S(t+s)] = \mu(t+s) + \sigma E_t[W(t+s) - W(t)] + \sigma W(t)$$

$$= \mu(t+s) + \sigma W(t)$$

$$= \mu t + \sigma W(t) + \mu s$$

$$= S(t) + \mu s.$$

It is clear that $S(t)$ is not a martingale. We can however convert $S(t)$ into a martingale by using the Girsanov Theorem to switch to an equivalent measure P^* so the drift of $S(t)$ is zero. To do this we need to come up with a function $\xi(t)$, and multiply it by the original probability measure associated with $S(t)$. The density for $S(t)$ is given by

$$dP[S(t)] = \frac{1}{\sqrt{2\pi\sigma^2 t}} e^{-\frac{1}{2\sigma^2 t}(S(t)-\mu t)^2} dS(t). \quad \text{(C65)}$$

To switch to a new probability measure P^* such that under P^*, $S(t)$ becomes a martingale, we need to multiply $dP[S(t)]$ by

$$\xi(t) = e^{-\frac{1}{\sigma^2}\left[\mu S(t) - \frac{1}{2}\mu^2 t\right]}. \tag{C66}$$

This yields

$$dP[S(t)]\xi(t) = e^{-\frac{1}{\sigma^2}\left[\mu S(t) - \frac{1}{2}\mu^2 t\right]} \frac{1}{\sqrt{2\pi\sigma^2 t}} e^{-\frac{1}{2\sigma^2 t}(S(t)-\mu t)^2} dS(t). \tag{C67}$$

Rearranging yields

$$dP^*[S(t)] = \frac{1}{\sqrt{2\pi\sigma^2 t}} e^{-\frac{1}{2\sigma^2 t}(S(t))^2} dS(t). \tag{C68}$$

Expression (C68) is an equivalent probability measure for the spot price process. Under this new measure $S(t)$ where is a normally distributed a mean of zero and diffusion of σt. Under the new probability measure we have removed the drift from the spot price without having to estimate it. This is a incredibly convenient for we are not burdened by the unobserved risk premium embedded in the expected return. The dynamics of the spot price may now be expressed in terms of a new driving term $dW^*(t)$.

Recall under the Girsanov theorem $dW^*(t)$ is defined

$$dW^*(t) = dW(t) - X(t)dt, \tag{C69}$$

and for our example $X(t) = \frac{\mu}{\sigma}$. The composition of the last term in expression (C69) is a function of the drift coefficient. That is, $X(t)$ measures how much the original mean will be changed. This is seen by substituting the expression for $X(t)$ above into (C69) and then using (C69) in the process for $S(t)$ to obtain

$$dS(t) = \mu dt + \sigma\left[dW^*(t) - \frac{\mu}{\sigma}dt\right]. \tag{C70}$$

Rearranging yields

$$dS(t) = \sigma dW^*(t). \tag{C71}$$

We see the mean of $S(t)$ has changed and now under P^*, $S(t)$ is a martingale.

In more general terms, the Girsanov theorem states that if we are given a Wiener process $W(t)$, then multiplying the probability distribution of this process by $\xi(t)$, we can obtain a new Wiener process $W^*(t)$ with probability distribution $P^*(t)$. The two processes are related to each other through

$$dW^*(t) = dW(t) - X(t)dt. \tag{C72}$$

That is, $W^*(t)$ is obtained by subtracting an $I(t)$ adapted drift from $W(t)$. If $X(t) = \lambda$ Then $X(t)$ in the Girsanov theorem plays the same role as λ did in the simpler examples above. Again, it measures how much the original mean will be changed.

Consider the Wiener process $W^*(t)$. There is something counter-intuitive about this process. It turns out that both $W^*(t)$ and $W(t)$ are standard Wiener processes. Thus, they do not have any drift. Yet they relate to each other by (C37)

$$dW^*(t) = dW(t) - X(t)dt, \tag{C73}$$

which means that at least one of these processes must have nonzero drift if $X(t)$ is not identical to zero. The point here is $W^*(t)$ has zero drift under P^*, whereas, $W(t)$ has zero drift under P. Hence, $W^*(t)$ can be used to represent unpredictable errors in dynamic systems given that we switch the probability measures from P to P^*. Also, because $W^*(t)$ contains a term $-X(t)dt$, using it as an error term in lieu of $W(t)$ would reduce the drift of the original stochastic differential equation under consideration exactly by $-X(t)dt$. If the $X(t)$ is interpreted as the time dependent risk premium, the transformation would make all risky assets grow at a risk-free rate.

Appendix D

Stochastic Calculus

In this appendix we illustrate the integration techniques used to determine the mean and variance for second order stochastic processes.

The asset pricing models in our analysis presumes that the dynamics of the underlying state variables follow particular stochastic differential equations. The dynamic specification for these state variables has implications for valuing the financial assets under consideration. In particular, the asset pricing models need the mean and variance of the individual state variables to price the financial assets. The specification given to the dynamic increment impacts the distribution of the state variable at time t and thus the its mean and variance. To obtain the mean and variance of the state variables we need to understand the properties of the stochastic differential equations. Consider

$$dX(t) = a(t)dt + f(t)dW(t), \qquad \text{(D1)}$$

where $a(t)$ is the drift coefficient,

$f(t)$ is the diffusion coefficient,

$dZ(t)$ is the increment of a Gauss Wiener process, and

$W(t)$ is $N \sim (0, \sqrt{t})$.

$X(t)$ is a stochastic process. It has an expected drift of $a(t)dt$ with unpredictable movements driven by $dW(t)$. The corresponding integral representation is

$$\int_0^t dX(s) = \int_0^t a(t)ds + \int_0^t f(t)dW(s)$$

$$X(t) - X(0) = a(t)t + \int_0^t f(t)dW(s)$$

$$X(t) = X(0) + a(t)t + \int_0^t f(t)dW(s). \qquad \text{(D2)}$$

If we are to find the mean and the variance of $X(t)$ we must be capable of evaluating the integral $\int_0^t f(t)dW(s)$ in expression (D2). That is, we must evaluate a stochastic integral. To do this we must consider some properties of a Gauss Wiener process.

A Gauss Wiener process is a stochastic process that describes the highly irregular movements of a particle suspended in a liquid. This motion is often referred to as Brownian motion named after the scientist who was first to study this phenomenon. We can describe this motion. Let the location of a particle be described by a Cartesian coordinate system whose origin is the location of the particle at time $t = 0$. Then the three coordinates of the position of the particle vary independently, each according to a stochastic process $W(t)$, $-\infty < t < \infty$, satisfying the following properties:

(i) $W(0) = 0$.

(ii) $W(t) - W(s)$ has a normal distribution with mean 0 and variance $\sigma_w^2(t-s)$ for $s \le t$.

(iii) $W(t_2) - W(t_1), W(t_3) - W(t_2), \cdots, W(t_n) - W(t_{n-1})$ are independent for all $t_1 \le t_2 \le \cdots \le t_n$.

(iv) $E[(W(s) - W(a))(W(t) - W(a))] = \sigma_w^2 \min(s-a, t-a)$.

Here σ_w^2 is some positive constant. Given its irregular motion the Wiener process is a continuous, nowhere differentiable function. The integral

$$\int_a^b f(t)dW(t) \qquad \text{(D3)}$$

does not exist in the usual sense. Nevertheless, it is possible to give meaning to this integral. One way of doing so is to define the integral as

$$\lim_{\varepsilon\to 0}\int_a^b f(t)\left(\frac{W(t+\varepsilon)+W(t)}{\varepsilon}\right)dt, \tag{D4}$$

provided the indicated limit exists. To see that this limit does indeed exist and to evaluate it explicitly, we observe that

$$\int_a^b f(t)\left(\frac{W(t+\varepsilon)+W(t)}{\varepsilon}\right)dt = \int_a^b f(t)\left(\frac{1}{\varepsilon}\int_t^{t+\varepsilon} dW(s)ds\right)dt,$$

$$\int_a^b f(t)\left(\frac{W(t+\varepsilon)+W(t)}{\varepsilon}\right)dt = \int_a^b f(t)\left(\frac{1}{\varepsilon}\int_t^{t+\varepsilon} W'(s)ds\right)dt.$$

Now

$$\frac{d}{dt}W(t) = W'(t).$$

Therefore

$$\int_a^b f(t)\left(\frac{W(t+\varepsilon)+W(t)}{\varepsilon}\right)dt = \int_a^b f(t)\left(\frac{1}{\varepsilon}\int_t^{t+\varepsilon} \frac{d}{dt}W(s)ds\right)dt,$$

or

$$\int_a^b f(t)\left(\frac{W(t+\varepsilon)+W(t)}{\varepsilon}\right)dt = \int_a^b f(t)\frac{d}{dt}\left(\frac{1}{\varepsilon}\int_t^{t+\varepsilon} W(s)ds\right)dt. \tag{D5}$$

Integrating the right-hand side by parts

$$uv\Big|_a^b - \int_a^b u\,dv, \tag{D6}$$

where $u = \frac{1}{\varepsilon}\int_t^{t+\varepsilon} W(s)ds$, $v = f(t)$ and $dv = f'(t)dt$.

We have

$$\int_a^b f(t)\left(\frac{W(t+\varepsilon)+W(t)}{\varepsilon}\right)dt = f(t)\left(\frac{1}{\varepsilon}\int_t^{t+\varepsilon} W(s)ds\right)\Bigg|_a^b - \int_a^b f(t)\left(\frac{1}{\varepsilon}\int_t^{t+\varepsilon} W(s)ds\right)dt. \tag{D7}$$

Since the Wiener process has continuous sample functions, it follows that the right-hand side of (D7) converges to

$$f(t)W(t)\Big|_a^b - \int_a^b f'(t)W(t)dt. \qquad \text{(D8)}$$

Thus we are led to define

$$\int_a^b f(t)dW(t)$$

as the limit of the right-hand side of (D7) as $\varepsilon \to 0$, that is, by the formula

$$\int_a^b f(t)W(t) = f(b)W(b) - f(a)W(a) - \int_a^b f'(t)W(t)dt. \qquad \text{(D9)}$$

Note that the right side of the above expression is well defined and that it agrees with the usual integration by part formula.

Since the Wiener process is a Gaussian process, it follows from property (i) and property (ii) that

$$\int_a^b f(t)dW(t)$$

is normally distributed with a mean of zero. Taking the expectation and noting that we can interchange the expected value and integration, we obtain

$$E\left[\int_a^b f(t)dW(t)\right] = \int_a^b f(t)E[dW(t)] = 0. \qquad \text{(D10)}$$

To compute the variance for *W(t)* we need to show that if $a \le b$ and *c(t)* is another continuously differentiable sample function of *W(t)* on [a,b], then

$$E\left[\int_a^b f(t)dW(t)\int_a^b g(t)dW(t)\right] = \sigma_w^2\int_a^b f(t)g(t)dt. \qquad \text{(D11)}$$

The above result is called Ito's isometry. To start the proof we rewrite (D11). We need to add and subtract

$$\int_a^b f'(t)W(a)dt \tag{D12}$$

from (D9). This yields

$$\int_a^b f(t)W(t) = f(b)W(b) - f(a)W(a) - \int_a^b f'(t)W(t)dt + \int_a^b f'(t)W(a)dt - \int_a^b f'(t)W(a)dt . \tag{D13}$$

Rearranging the above gives

$$\int_a^b f(t)W(t) = f(b)W(b) - f(a)W(a) - [f(b)W(a) - f(a)W(a)] - \int_a^b f'(t)(W(t) - W(a))dt$$

or

$$\int_a^b f(t)W(t) = f(b)(W(b) - W(a)) - \int_a^b f'(t)(W(t) - W(a))dt . \tag{D14}$$

Given the result above we have

$$\int_a^b g(t)W(t) = g(b)(W(b) - W(a)) - \int_a^b g'(t)(W(t) - W(a))dt . \tag{D15}$$

Now we need want to find

$$E\left[\int_a^b f(t)dW(t)\int_a^b g(t)dW(t)\right] = E\left\{\left(f(b)(W(b) - W(a)) - \int_a^b f'(t)(W(t) - W(a))dt\right)\right.$$

$$\left. \mathrm{x}\left(g(b)(W(b) - W(a)) - \int_a^b g'(t)(W(t) - W(a))dt\right)\right\}.$$

Or

$$E\left[\int_a^b f(t)dW(t)\int_a^b g(t)dW(t)\right] = E\left[\int_a^b f'(t)(W(t) - W(a))dt\int_a^b g'(t)(W(t) - W(a))dt\right]$$

$$- E\left[f(b)(W(b) - W(a))\int_a^b g'(t)(W(t) - W(a))dt\right]$$

$$-E\left[g(b)(W(b)-W(a))\int_a^b f'(t)(W(t)-W(a))dt\right]$$

$$+E[g(b)(W(b)-W(a))f(b)(W(b)-W(a))]. \qquad \text{(D16)}$$

We will evaluate the right side of (D16) by breaking it up into four separate terms. We begin with

$$E\left[\int_a^b f'(t)(W(t)-W(a))dt\int_a^b g'(t)(W(t)-W(a))dt\right]. \qquad \text{(D17)}$$

The above expression can be rewritten as

$$E\left[\int_a^b f'(t)(W(t)-W(a))dt\int_a^b g'(t)(W(t)-W(a))dt\right]$$

$$=\int_a^b f'(t)\int_a^b g'(t)E[(W(t)-W(a))(W(t)-W(a))]dtdt. \qquad \text{(D18)}$$

The right-hand side of the above can be written as

$$=\int_a^b [(f(t)-W(a))(g(t)-W(a))]dt. \qquad \text{(D19)}$$

To show this result, consider the following. We know from property (iv)

$$E[(W(s)-W(a))(W(t)-W(a))]=\sigma_w^2\min(s-a,t-a).$$

Now writing the right-hand side of (D18) as

$$=\int_a^b f'(s)\int_a^b g'(t)E[(W(s)-W(a))(W(t)-W(a))]dtds, \qquad \text{(D20)}$$

and using the property for the Wiener process we get

$$\int_a^b f'(s)\int_a^b g'(t)E[(W(s)-W(a))(W(t)-W(a))]dtds=$$

$$\sigma_w^2\int_a^b f'(s)\int_a^b g'(t)\min(s-a,t-a)dtds. \qquad \text{(D21)}$$

Now rewrite the inner integral as

$$= \int_a^s (t-a) g'(t) dt + (s-a) \int_s^b g'(t) dt . \tag{D22}$$

Now integrating the first integral by parts yields

$$u = g(t) ,\ v = (t-a) \text{, and } dv = dt$$

$$= (t-a) g(t) \Big|_a^s - \int_a^s g(t) dt + (s-a) \int_s^b g'(t) dt . \tag{D23}$$

Integrating the above we get

$$= (t-a) g(t) \Big|_a^s - \int_a^s g(t) dt + (s-a)(g(b) - g(s)),$$

$$= (s-a) g(s) - (a-a) g(a) - \int_a^s g(t) dt + (s-a)(g(b) - g(s)),$$

$$= (s-a) g(b) - \int_a^s g(t) dt ,$$

$$= \int_a^s (g(b) - g(t)) dt . \tag{D24}$$

Thus the left-hand side of (D18) equals

$$= \sigma_w^2 \int_a^b f'(s) \int_a^s (g(b) - g(t)) dt ds . \tag{D25}$$

Interchanging the order of integration gives

$$= \sigma_w^2 \int_a^b (g(b) - g(t)) \int_t^b f'(s) ds dt ,$$

$$= \sigma_w^2 \int_a^b (g(b) - g(t))(f(b) - f(t)) dt . \tag{D26}$$

This is also equal to

$$= \sigma_w^2 \int_a^b (g(t) - g(b))(f(t) - f(b)) dt . \tag{D27}$$

Therefore we can write (D18) as

$$E\left[\int_a^b f'(t)(W(t)-W(a))dt\int_a^b g'(t)(W(t)-W(a))dt\right]=\sigma_w^2\int_a^b (g(t)-g(b))(f(t)-f(b))dt \text{ .(D28)}$$

Now observe

$$-E\left[f(b)(W(b)-W(a))\int_a^b g'(t)(W(t)-W(a))dt\right]. \tag{D29}$$

Interchanging the order of integration and the expectation the above expression is equal to

$$=-f(b)\int_a^b g'(t)E[(W(b)-W(a))(W(t)-W(a))]dt. \tag{D30}$$

Using the property $E[(W(s)-W(a))(W(t)-W(a))]=\sigma_w^2 \min(s-a,t-a)$ the above can be written as

$$=-\sigma_w^2 f(b)\int_a^b g'(t)(t-a)dt. \tag{D31}$$

Now we use integration by parts to obtain

$$=-\sigma_w^2 f(b)\left[(t-a)g(t)\big|_a^b-\int_a^b g(t)dt\right],$$

$$=-\sigma_w^2 f(b)\left[(b-a)g(b)-\int_a^b g(t)dt\right],$$

$$=-\sigma_w^2 f(b)\int_a^b g(b)-g(t)dt. \tag{D32}$$

Consequently

$$-E\left[f(b)(W(b)-W(a))\int_a^b g'(t)(W(t)-W(a))dt\right]=\sigma_w^2\int_a^b f(b)(g(t)-g(b))dt. \tag{D33}$$

Similarly we find that

$$-E\left[g(b)(W(b)-W(a))\int_a^b f'(t)(W(t)-W(a))dt\right]=\sigma_w^2\int_a^b g(b)(f(t)-f(b))dt\,. \quad \text{(D34)}$$

Lastly we note

$$E[g(b)(W(b)-W(a))f(b)(W(b)-W(a))]=g(b)f(b)E[(W(b)-W(a))(W(b)-W(a))]. \quad \text{(D35)}$$

This is equal to

$$g(b)f(b)E[(W(b)-W(a))(W(b)-W(a))]=g(b)f(b)\sigma_w^2(b-a)\,, \quad \text{(D36)}$$

or

$$g(b)f(b)E[(W(b)-W(a))(W(b)-W(a))]=\sigma_w^2\int_a^b g(b)f(b)dt\,. \quad \text{(D37)}$$

Now plug (D28), (D33), (D34), and (D37) into (D16) we get

$$E\left[\int_a^b f(t)dW(t)\int_a^b g(t)dW(t)\right]=\sigma_w^2\int_a^b (g(t)-g(b))(f(t)-f(b))dt+\sigma_w^2\int_a^b f(b)(g(t)-g(b))dt$$

$$+\sigma_w^2\int_a^b g(b)(f(t)-f(b))dt+\sigma_w^2\int_a^b g(b)f(b)dt\,. \quad \text{(D38)}$$

The above reduces to

$$E\left[\int_a^b f(t)dW(t)\int_a^b g(t)dW(t)\right]=\sigma_w^2\int_a^b f(t)g(t)dt\,. \quad \text{(D39)}$$

This is what we wanted to show. The above is called Ito's isometry and is very powerful result. To see this, note that the term on the left-hand side of (D38) is composed of two stochastic integrals. That is the integrals contain random variables. In fact, as of time t-1 the term

$$W(t)-W(t-1)$$

is a random variable, and the sum

$$\sum_{t=0}^{n} f(t)\big(W(t) - W(t-1)\big)$$

is an integral with respect to a random variable. The integral itself is thus a random variable and cannot be evaluated using rules of deterministic calculus. Evaluating these integrals is simply not easy, but we do not have to evaluate these integrals. Ito's isometry shows that the multiplication of two stochastic integrals is an integral over the well defined continuous diffusion coefficients with respect to time. The right-hand side of (D39) is deterministic and can be evaluated using standard Riemann sums.

Appendix E

A Discussion of the Feynman-Kac Formula

In financial economics, contingent claim pricing models often use equivalent martingale measures to price a financial asset. These probability measures are useful in that they provide an arbitrage-free conditional forecast for the terminal spot price of an underlying asset. Given this forecast, economists can use this analytical solution to price any derivative security. This method of solution for the price of a derivative security is motivated by the Feynman-Kac formula. The Feynman-Kac formula shows there is a correspondence between a class of conditional expectations and a set partial differential equation. This is useful in that if a known conditional expectation exists for the partial differential equation, then we may use this forecast to value the contingent claim instead of solving the partial differential equation. That is, if the forecast is easy to calculate then it is beneficial to value the asset via the conditional expectation as opposed to dealing with a partial differential equation.

The purpose of this appendix is to illustrate Feynman-Kac by example. That is we will show how the conditional expectation is linked to a particular partial differential equation by mechanically deriving the partial differential equation from a conditional expectation. Three examples are considered. The first example is for a model with deterministic discount rates and a random cash flow. The second example deals with a stochastic discount factor and a known cash flow. The final example and more meaningful one considers both a stochastic discount factor and a random cash flow.

Example 1 Deterministic discount rate and random cash flow

Consider the function $F(x(t))$ of a random process $x(t) \in [0,\infty)$ defined by the conditional expectation

$$F(x(0)) = E_0\left[\int_0^\infty e^{-\beta s} g(x(s))ds\right], \qquad \text{(E1)}$$

where $\beta > 0$ represents a constant instantaneous discount rate, $g(x(s))$ is some continuous payout that depends on the value assumed by the random process $x(t)$. $E_t[\ ]$ is the expectation conditional on the information set $I(t)$. The process $x(t)$ obeys the stochastic differential equation

$$dx(t) = \mu dt + \sigma dZ(t), \qquad \text{(E2)}$$

where μ and σ are known constants. $F(x(t))$ is interpreted as the expected value of some discounted future cash flow $g(x(s))$ that depends on the random variable $x(s)$. For now we focus on the case of a deterministic discount factor and show how we obtain a corresponding partial differential equation for the conditional expectation in equation (E1). Once we have this result, we will turn our focus to random discount factors.

We obtain a partial differential equation that corresponds to the expectation in (E1) in several steps. We proceed in a mechanical way, to illustrate the derivation. First, consider a small time interval $0 < \Delta$ and split the period $[0,\infty)$ into two. One being the immediate future, represented by the interval $[0,\Delta]$, and the other represented by $[\Delta,\infty)$. The integral inside the expectation in equation (E1) can be rewritten as

$$\int_0^\infty e^{-\beta s} g(x(s))ds = \int_0^\Delta e^{-\beta s} g(x(s))ds + \int_\Delta^\infty e^{-\beta s} g(x(s))ds.$$

Thus, expression (E1) can be alternatively written as

$$F(x(0)) = E_t\left[\int_0^{\Delta} e^{-\beta s} g(x(s))ds + \int_{\Delta}^{\infty} e^{-\beta s} g(x(s))ds\right]. \qquad \text{(E3)}$$

Expression (E3) offers some insight to the relationship between the conditional expectation and a related partial differential equation. Both the conditional expectation and partial differential equations yield forecasts for the state variable. We are interested in finding the correspondence between the two forecasts, which yield the same result. Looking at equation (E1), we see the forecast is for the entire investment horizon. Expression (E3) breaks this forecast into two intervals; the immediate future and the remaining time horizon. Intuitively, this is similar to the makeup of a stochastic differential equation. That is, the dynamics of the state variable are given by an instantaneous drift, indicating the immediate expected movement, and a diffusion term, dictating unexpected movements. Comparing the characteristics of a stochastic differential equation to expression (E3), we may say in very general terms that the integrals in expression (E3) is an approximation of the dynamics of $g(x(t))$.

Continuing our analysis, the next step involves some elementary transformations that are intended to introduce the future value of $F(x(t))$ to the right-hand side of expression (E3). Multiply and divide the second term in expression (E3) by $e^{-\beta(\Delta)}$. This yields

$$F(x(0)) = E_0\left[\int_0^{\Delta} e^{-\beta s} g(x(s))ds + \int_{\Delta}^{\infty} e^{-\beta s}\left(e^{-\beta(\Delta)}e^{\beta(\Delta)}\right)g(x(s))ds\right],$$

$$F(x(0)) = E_0\left[\int_0^{\Delta} e^{-\beta s} g(x(s))ds + e^{-\beta(\Delta)}\int_{\Delta}^{\infty} e^{-\beta(s-\Delta)} g(x(s))ds\right]. \qquad \text{(E4)}$$

Recall the recursive property of conditional expectations. When conditional expectations are nested, it is the expectation with respect to the smaller information set that is relevant. Thus, if we have $I(t) \subseteq I(s)$, we can write

$$E_t[E_s(Y(t))] = E_t[Y(t)].$$

This permits us to rewrite expression (E4) as

$$F(x(0)) = E_0\left[\int_0^{\Delta} e^{-\beta s} g(x(s))ds + e^{-\beta(\Delta)} E_{\Delta}\left[\int_{\Delta}^{\infty} e^{-\beta(s-\Delta)} g(x(s))ds\right]\right]. \tag{E5}$$

Recognizing the second term inside the inner brackets on the right-hand side of expression (E5) as $F(x(\Delta))$, we rewrite (E5) as

$$F(x(0)) = E_t\left[\int_0^{\Delta} e^{-\beta s} g(x(s))ds + e^{-\beta(\Delta)} F(x(\Delta))\right]. \tag{E6}$$

Now, grouping all terms on the right-hand side of equation (E6) yields

$$E_0\left[\int_0^{\Delta} e^{-\beta s} g(x(s))ds + e^{-\beta(\Delta)} F(x(\Delta))\right] - F(x(0)) = 0.$$

Note, $F(x(t))$ is a known value of $F(\ )$ at time 0. Since the expectation is taken with respect to $I(t)$, we know that

$$F(x(t)) = E_t[F(x(t))].$$

Therefore we may move $F(x(0))$ in expression (E6) inside the expectation operator. This yields

$$E_0\left[\int_0^{\Delta} e^{-\beta s} g(x(s))ds + e^{-\beta(\Delta)} F(x(\Delta)) - F(x(0))\right] = 0. \tag{E7}$$

Now, add and subtract $F(x(\Delta))$ to the left-hand side of expression (E7). This yields

$$E_0\left[\int_0^{\Delta} e^{-\beta s} g(x(s))ds + e^{-\beta(\Delta)} F(x(\Delta)) - F(x(0)) + F(x(\Delta)) - F(x(\Delta))\right] = 0.$$

Rearranging the above yields

$$E_0\left[\int_0^{\Delta} e^{-\beta s} g(x(s))ds + \left(e^{-\beta(\Delta)} - 1\right)F(x(\Delta)) + F(x(\Delta)) - F(x(0))\right] = 0. \qquad \text{(E8)}$$

The term inside the brackets in equation (E8) describes the movement of the stochastic process. We are interested in understanding the movement for an infinitesimal interval, therefore we take the derivative of (E8) with respect to Δ. That is,

$$\lim_{\Delta\to 0}\frac{1}{\Delta}E_t\left[\int_0^{\Delta} e^{-\beta s} g(x(s))ds + \left(e^{-\beta(\Delta)} - 1\right)F(x(\Delta)) + F(x(\Delta)) - F(x(0))\right] = 0 \qquad \text{(E9)}$$

The first term is the derivative with respect to the upper limit of a Riemann integral. That is

$$\lim_{\Delta\to 0}\frac{1}{\Delta}\int_0^{\Delta} e^{-\beta s} g(x(s))ds = \frac{\partial}{\partial\Delta}\int_0^{\Delta} e^{-\beta s} g(x(s))ds,$$

$$\frac{\partial}{\partial\Delta}\int_0^{\Delta} e^{-\beta s} g(x(s))ds = (1)e^{-\beta\Delta} g(x(\Delta))\Big|_{\Delta=0}$$

$$= e^{-\beta 0} g(x(0)),$$

$$= g(x(0)) \qquad \text{E(10)}$$

The second term, is in fact, a standard derivative of $e^{\beta x}$ evaluated at $x = 0$

$$\lim_{\Delta\to 0}\frac{1}{\Delta}\left(e^{-\beta(\Delta)} - 1\right) = \lim_{\Delta\to 0}\frac{\left(e^{-\beta(\Delta)} - e^{-\beta(0)}\right)}{\Delta},$$

$$\lim_{\Delta\to 0}\frac{\left(e^{-\beta(\Delta)} - e^{-\beta(0)}\right)}{\Delta} = \frac{\partial e^{-\beta(\Delta)}}{\partial\Delta}$$

$$\frac{\partial e^{-\beta(\Delta)}}{\partial\Delta} = -\beta e^{-\beta(\Delta)}\Big|_{\Delta=0}.$$

$$= -\beta e^{-\beta(0)}$$

$$= -\beta. \qquad \text{(E11)}$$

The last term, involves the expectation of a stochastic differential and hence requires the application of Ito's lemma. That is,

$$F(x(\Delta)) - F(x(0)) \cong F_x dx + \frac{1}{2} F_{xx} dx^2 + F_t \Delta ,$$

$$F(x(\Delta)) - F(x(0)) \cong F_x [\mu\Delta + \sigma dZ(\Delta)] + \frac{1}{2} F_{xx} \sigma^2 \Delta + F_t \Delta ,$$

$$F(x(\Delta)) - F(x(0)) \cong F_x [\mu\Delta + \sigma dZ(\Delta)] + \frac{1}{2} F_{xx} \sigma^2 \Delta + F_t \Delta ,$$

$$F(x(\Delta)) - F(x(0)) \cong \left[F_x \mu + \frac{1}{2} F_{xx} \sigma^2 + F_t \right] \Delta + \sigma dZ(\Delta)$$

Thus,

$$\lim_{\Delta \to 0} \frac{1}{\Delta} [F(x(\Delta)) - F(x(0))] = \lim_{\Delta \to 0} \frac{1}{\Delta} \left(\left[F_x \mu + \frac{1}{2} F_{xx} \sigma^2 + F_t \right] \Delta + \sigma dZ(\Delta) \right),$$

$$\lim_{\Delta \to 0} \frac{1}{\Delta} \left(\left[F_x \mu + \frac{1}{2} F_{xx} \sigma^2 + F_t \right] \Delta + \sigma dZ(\Delta) \right) = F_x \mu + \frac{1}{2} F_{xx} \sigma^2 + F_t + \lim_{\Delta \to 0} \frac{1}{\Delta} (\sigma dZ(\Delta)). \quad \text{(E12)}$$

Substituting expressions (E10), (E11), and (E12) into (E9) yields

$$\lim_{\Delta \to 0} \frac{1}{\Delta} E_0 \left[\int_0^\Delta e^{-\beta s} g(x(s)) ds + \left(e^{-\beta(\Delta)} - 1\right) F(x(\Delta)) + F(x(\Delta)) - F(x(0)) \right]$$

$$= E_0 \left[g(x(0)) - \beta F(x(0)) + F_x \mu + \frac{1}{2} F_{xx} \sigma^2 + F_t + \lim_{\Delta \to 0} \frac{1}{\Delta} (\sigma dZ(\Delta)) \right] = 0 \quad \text{(E13)}$$

Distributing the expectation operator in expression (E13) yields

$$g(x(0)) - \beta F(x(0)) + F_x \mu + \frac{1}{2} F_{xx} \sigma^2 + F_t + \lim_{\Delta \to 0} \frac{1}{\Delta} [\sigma E_0 (dZ(\Delta))] = 0,$$

$$g(x(0)) - \beta F(x(0)) + F_x \mu + \frac{1}{2} F_{xx} \sigma^2 + F_t = 0,$$

$$F_x\mu + \frac{1}{2}F_{xx}\sigma^2 + F_t - \beta F(x(0)) + g(x(0)) = 0. \qquad \text{(E14)}$$

Expression (E14) is the desired partial differential equation we wanted to reach. This partial differential equation corresponds to the expectation of the present value of a cash flow stream $g(x(s))$. If this present value is given by the conditional expectation shown above, then it cannot be an arbitrary function of $x(t)$. That is, its behavior over time must satisfy some constraints due to the expected future behavior of $x(t)$. These constraints lead us to the partial differential equation.

Example 2 Stochastic interest rate and a known cash flow

Consider the price of a pure discount default free bond, $P(t,T)$, in a no-arbitrage setting. Assume that the instantaneous spot rate $r(t)$ is a Markov process and write the price of the bond with a par value of $1, using the familiar formula

$$P(t,T) = E_t^*\left[1e^{-\int_t^T r(s)ds}\right], \qquad \text{(E15)}$$

with

$$P(T,T) = 1. \qquad \text{(E16)}$$

Here the expectation is taken with respect to an equivalent probability measure, which is conditioned on an information set available at time *t*, namely $I(t)$. This is assumed to include the current observation on the spot rate $r(t)$. If $r(t)$ is a Markov process $P(t,T)$ will depend on the latest observation of $r(t)$. Because we are in the ***risk-neutral*** world, as dictated by the use of the assumed probability measure, the $r(t)$ process will follow the dynamics given by the stochastic differential equation

$$dr(t) = [a(r(t),t) - \lambda_t b(r(t),t)]dt + b(r(t),t)dW^*(t), \qquad \text{(E17)}$$

where $W^*(t)$ is a Wiener process under the risk neutral measure. The λ_t is the market price of interest rate risk defined by

$$\lambda_t = \frac{\mu - r(t)}{\sigma} \qquad \text{(E18)}$$

with μ and σ being short-hand notation for the drift and diffusion components of the bond price dynamics

$$dP = \mu P dt + \sigma P dW^*(t). \qquad \text{(E19)}$$

Note the Weiner term in expression (E19) is the same as the Weiner term in expression (E17).

We again have a conditional expectation and a process that is driving it, just as in the previous case. This means that we can apply the same steps used there and obtain a partial differential equation that corresponds to *P(t,T)*. Yet, in the present case, this partial differential equation may also have some practical use in pricing bonds. It can be solved numerically, or if a closed-form solution exists, analytically.

The same steps will be applied in a mechanical way. First, split the interval $[t,T]$ into two parts to write

$$P(t,T) = E_t^*\left[\left(e^{-\int_t^{t+\Delta} r(s)ds}\right)\left(e^{-\int_{t+\Delta}^{T} r(s)ds}\right)\right]. \qquad \text{(E20)}$$

Second, try to introduce the future price of the bond, $P(t+\Delta,T)$, in this expression. In fact, the second exponential on the right-hand side can easily be recognized as $P(t+\Delta,T)$ once we use the recursive property of conditional expectations. Using

$$E_t[E_{t+\Delta}(\cdot)] = E_t[\cdot],$$

we can write

$$P(t,T) = E_t^*\left[\left(e^{-\int_t^{t+\Delta} r(s)ds}\right)\left(e^{-\int_{t+\Delta}^{T} r(s)ds}\right)\right]$$

as

$$P(t,T) = E_t^*\left[\left(e^{-\int_t^{t+\Delta} r(s)ds}\right)E_{t+\Delta}^*\left(e^{-\int_{t+\Delta}^{T} r(s)ds}\right)\right],$$

$$P(t,T) = E_t^*\left[\left(e^{-\int_t^{t+\Delta} r(s)ds}\right)P(t+\Delta,T)\right]. \qquad \text{(E21)}$$

In the third step, group all terms inside the expectation sign on the right-hand side of expression (E21)

$$E_t^*\left[\left(e^{-\int_t^{t+\Delta} r(s)ds}\right)P(t+\Delta,T) - P(t,T)\right] = 0. \qquad \text{(E22)}$$

Now, add and subtract $P(t+\Delta,T)$ from the right-hand side of expression (E22)

$$E_t^*\left[\left(e^{-\int_t^{t+\Delta} r(s)ds}\right)P(t+\Delta,T) - P(t,T) - P(t+\Delta,T) + P(t+\Delta,T)\right] = 0,$$

$$E_t^*\left[\left(e^{-\int_t^{t+\Delta} r(s)ds} - 1\right)P(t+\Delta,T) + P(t+\Delta,T) - P(t,T)\right] = 0. \qquad \text{(E23)}$$

To find the instantaneous change we take the derivative of equation (E23) with respect to Δ. That is

$$\lim_{\Delta\to 0}\frac{1}{\Delta}E_t^*\left[\left(e^{-\int_t^{t+\Delta} r(s)ds}-1\right)P(t+\Delta,T)+P(t+\Delta,T)-P(t,T)\right]=0. \qquad \text{(E24)}$$

The first expression in equation (E24)

$$\lim_{\Delta\to 0}\frac{1}{\Delta}E_t^*\left[\left(e^{-\int_t^{t+\Delta} r(s)ds}-1\right)P(t+\Delta,T)\right]$$

is the derivative with respect to the upper limit of a Riemann integral. Alternatively, the expression above can be written as

$$\frac{\partial}{\partial\Delta}E_t^*\left[\left(e^{-\int_t^{t+\Delta} r(s)ds}-1\right)P(t+\Delta,T)\right]=E_t^*\left[\frac{\partial}{\partial\Delta}\left(e^{-\int_t^{t+\Delta} r(s)ds}-1\right)P(t+\Delta,T)\right],$$

$$E_t^*\left[\frac{\partial}{\partial\Delta}\left(e^{-\int_t^{t+\Delta} r(s)ds}-1\right)P(t+\Delta,T)\right]=E_t^*\left[-r(t+\Delta)e^{-\int_t^{t+\Delta} r(s)ds}P(t+\Delta,T)\Big|_{\Delta=0}\right],$$

$$=E_t^*\left[-r(t)e^{-\int_t^{t} r(s)ds}P(t,T)\right],$$

$$=E_t^*\left[-r(t)\left(e^{-(r(t)-r(t))}\right)P(t,T)\right]$$

$$=E_t^*\left[-r(t)\left(e^{0}\right)P(t,T)\right]$$

$$=E_t^*\left[-r(t)P(t,T)\right] \qquad \text{(E25)}$$

Next, we apply Ito's lemma to the second term of equation (24) and take the expectation

$$\lim_{\Delta\to 0}\frac{1}{\Delta}E_t^*\left[P(t+\Delta,T)-P(t,T)\right]\cong\lim_{\Delta\to 0}\frac{1}{\Delta}E_t^*\left[P_t\Delta+P_r dr+\frac{1}{2}P_{rr}dr^2\right],$$

$$=\lim_{\Delta\to 0}\frac{1}{\Delta}E_t^*\left[P_t\Delta+P_r\left[a(r(t),t)-\lambda_t b(r(t),t)\right]\Delta+P_r b(r(t),t)dW^*(t)+\frac{1}{2}P_{rr}b(r(t),t)^2\Delta\right],$$

$$= \lim_{\Delta \to 0} \frac{1}{\Delta} E_t^* \left[\left[P_t + P_r\left(a(r(t),t) - \lambda_t b(r(t),t)\right) + \frac{1}{2} P_{rr} b(r(t),t)^2 \right] \Delta \right],$$

$$= E_t^* \left[P_t + P_r\left(a(r(t),t) - \lambda_t b(r(t),t)\right) + \frac{1}{2} P_{rr} b(r(t),t)^2 \right]. \qquad \text{(E26)}$$

Substitute equations (E25) and (E26) into (E24) yields

$$E_t^* \left[\frac{1}{2} P_{rr} b(r(t),t)^2 + P_r\left(a(r(t),t) - \lambda_t b(r(t),t)\right) - r(t)P(t,T) + P_t \right] = 0.$$

Distributing the expectation operator yields

$$\frac{1}{2} P_{rr} b(r(t),t)^2 + P_r\left(a(r(t),t) - \lambda_t b(r(t),t)\right) - r(t)P(t,T) + P_t = 0 \qquad \text{(E27)}$$

Equation (E27) is the partial differential equation for an arbitrage-free valuation model for pure discount bonds with no default risk, where the only source of uncertainty is with a stochastic interest rate process. Here the cash payout at maturity is a known amount, the par value of the bond. We cannot use this model if the future cash flow cash flow was uncertain. We can, however, combine the first two examples. That is, we can develop the relationship between the conditional expectation for a random interest rate and stochastic cash flow and the corresponding partial differential equation

Example 3 Stochastic discount factor with a stochastic cash flow

We have two cases where the existence of a certain type of conditional expectation led to a corresponding partial differential equation. In the first, case there was a random cash flow stream depending on an underlying process $x(t)$ but the discount rate was constant. In the second case, the instrument paid a single, fixed cash flow at maturity, yet the discount factor was random. Our purpose now is to combine the above two cases and derive a partial differential equation that corresponds to the conditional expectation

$$F(x(t),r(t),t)=E_t^*\left[\int_t^T\left(e^{-\int_t^u r(s)ds}\right)g(x(u))du\right]. \tag{E28}$$

Expression (E28) represents the price of an instrument that makes variable cash flow payments at times $u \in [t,T]$, and hence needs to be evaluated using the random discount factor

$$e^{-\int_t^u r(s)ds}, \tag{E29}$$

at each u. Here we cannot directly apply the expectation operator in (E28) because the interest process is correlated with $g(x(u))$. If the two components are independent then the expectation could be evaluated. To illustrate this problem consider the following example.

Consider to random variables X and Y. Each random variable has two observations, which we denote as $[x_1,x_2]$, and $[y_1,y_2]$, with marginal probabilities denoted by $f(x_i)$ and $f(y_i)$; where $i=1,2$. If X and Y are independent then

$$E(XY)=E(X)E(Y). \tag{180}$$

To see this we know that

$$E(XY)=\sum_y\sum_y xyf(x,y), \tag{181}$$

where $f(x,y)$ is the joint probability density function for X and Y. We may rewrite the double sum as

$$=x_1y_1f(x_1,y_1)+x_2y_1f(x_2,y_1)+x_1y_2f(x_1,y_2)+x_2y_2f(x_2,y_2). \tag{182}$$

Since X and Y are independent the joint density function can be rewritten as the product of the marginal density functions. This is

$$= x_1 f(x_1) y_1 f(y_1) + x_2 f(x_2) y_1 f(y_1) + x_1 y_2 f(x_1) f(y_2) + x_2 f(x_2) y_2 f(y_2). \quad (183)$$

Simplifying expression (183) yields

$$= [x_1 f(x_1) + x_2 f(x_2)] y_1 f(y_1) + [x_1 f(x_1) + x_2 f(x_2)] y_2 f(y_2).$$

$$= E[X] y_1 f(y_1) + E[X] y_2 f(y_2). \quad (184)$$

Factoring out the expected value of X yields

$$= (E[X])[y_1 f(y_1) + y_2 f(y_2)]$$

$$= E[X]E[Y]. \quad (185)$$

We see that if X and Y are independent then the expectation of the multiplicative interaction of X and Y is equal to the multiplicative of the individual expectations. If X is analogous to $e^{\int_t^u -r(s)ds}$ and Y to $g(x(u))$, then we may use the above result to evaluate the expectation.

The result above is good only if the two variables are independent, and for our most financial assets this is simply unrealistic. Therefore, we need to consider the expectation for two correlated variables. Let X and Y be two dependent variables. What is $E(XY)$ equal to? By definition

$$E(XY) = \sum_y \sum_y xyf(x,y). \quad (186)$$

We may rewrite the double sum as

$$= x_1 y_1 f(x_1, y_1) + x_2 y_1 f(x_2, y_1) + x_1 y_2 f(x_1, y_2) + x_2 y_2 f(x_2, y_2). \quad (187)$$

We may further rewrite the above by rewriting the joint densities as follows

$$= x_1 y_1 f(x_1) f(y_1 \mid x_1) + x_2 y_1 f(x_2) f(y_1 \mid x_2)$$

$$+ x_1 y_2 f(x_1) f(y_2 \mid x_1) + x_2 y_2 f(x_2) f(y_2 \mid x_2). \quad (188)$$

Simplifying the expression above yields

$$= x_1 y_1 f(x_1) f(y_1 \mid x_1) + x_1 y_2 f(x_1) f(y_2 \mid x_1)$$

$$+ x_2 y_1 f(x_2) f(y_1 \mid x_2) + x_2 y_2 f(x_2) f(y_2 \mid x_2). \tag{189}$$

Factoring yields

$$= x_1 f(x_1)[y_1 f(y_1 \mid x_1) + y_2 f(y_2 \mid x_1)] + x_2 f(x_2)[y_1 f(y_1 \mid x_2) + y_2 f(y_2 \mid x_2)].$$

$$= x_1 f(x_1) E[Y \mid X = x_1] + x_2 f(x_2) E[Y \mid X = x_2]. \tag{190}$$

The expression above cannot be simplified any further. The expectation for the multiplicative interaction between two correlated random variable X and Y does not simplify to simple expression. This is why we may not find a closed for expression for the Black-Scholes model when the interests is stochastic.