Electronic Thesis and Dissertation Repository

8-28-2013 12:00 AM

Pricing and Hedging Index Options with a Dominant Constituent Stock

Helen Cheyne
The University of Western Ontario

Supervisor
Dr. Matt Davison
The University of Western Ontario

Graduate Program in Applied Mathematics
A thesis submitted in partial fulfillment of the requirements for the degree in Master of Science
© Helen Cheyne 2013

PRICING AND HEDGING INDEX OPTIONS WITH A DOMINANT
CONSTITUENT STOCK
(Thesis format: Monograph)

by

Helen Cheyne

Graduate Program in Applied Mathematics

A thesis submitted in partial fulfillment
of the requirements for the degree of
Master of Science

The School of Graduate and Postdoctoral Studies
The University of Western Ontario
London, Ontario, Canada

Abstract

In this paper, we examine the pricing and hedging of an index option where one constituent stock plays an overly dominant role in the index. Under a Geometric Brownian Motion assumption we compare the distribution of the relative value of the index if the dominant stock is modeled separately from the rest of the index (see Eq. 1), or not (see Eq. 2). The former is equivalent to the relative index value being distributed as the sum of two lognormal random variables

$$Y = c_1 e^{Z_1} + c_2 e^{Z_2}, \quad \text{where} \quad Z_i \sim \mathcal{N}(\mu_i, \sigma_i), \quad i = 1, 2 \tag{1}$$

and the latter is distributed as a single lognormal random variable

$$X = c_3 e^{Z_3}, \quad \text{where} \quad Z_i \sim \mathcal{N}(\mu_i, \sigma_i), \quad i = 3. \tag{2}$$

Since $X \neq Y$ in distribution, with Y having fatter tails, we compare the two models. The validity of this theoretical result is verified against empirical stock market data. We look at two main models representing these cases: first, we use numerical methods to solve the two-dimensional problem directly (see Eq. 3); second, we make simplifying assumptions (see Eq. 4) to reduce the two-dimensional Black-Scholes problem to a one-dimensional Black-Scholes problem that can be solved analytically (see Eq. 5). The two-dimensional PDE for V(A,B,t) is

$$V_t + \frac{1}{2}\sigma_A^2 A^2 V_{AA} + \rho \sigma_A \sigma_B ABV_{AB} + \frac{1}{2}\sigma_B^2 B^2 V_{BB} + rAV_A + rBV_B - rV = 0 \tag{3}$$

with the terminal condition $V(A, B, T) = (K - A - B)^+$, where A is the dominant stock and B is the rest of the index. So our simplifying assumption is that

$$V_{AA} = V_{AB} = V_{BB} = V_{II} \quad \text{and} \quad V_A = V_B = V_I \tag{4}$$

from $I = A + B$, so we let $\alpha = \frac{A}{A+B}$ to reduce the full PDE (see Eq. 3) to the one-dimensional PDE for V(I,t)

$$V_t + \frac{1}{2}(\sigma^*)^2 V_{II} + rIV_I - rV = 0 \quad \text{where} \quad (\sigma^*)^2 = \sigma_A^2 \alpha^2 + 2\rho\sigma_A\sigma_B\alpha(1-\alpha) + \sigma_B^2(1-\alpha)^2 \tag{5}$$

with terminal condition $V(I, T) = (K - I)^+$.

Since the terminal conditions are non-smooth the numerical methods are verified by comparison to a Monte Carlo simulated solution.

Attributes of the models that we compare are the relative option price differences and expected hedging profits. We compare the models for various volatilities, dominance levels, correlations and risk free rates.

This work is significant in options trading because when a stock becomes dominant in its index the distribution of the returns changes. Even if the effect is small, given the millions of dollars exposed to index option trades, it has a material impact.

Contents

List of Figures

List of Tables

List of Appendices

Chapter 1

Motivation

Although it suffers from some shortcomings, Geometric Brownian motion (GBM) remains a workhorse for quantitative finance, because of its theoretical elegance and computational tractability. In this model, a stock, S_t, has log returns, $\ln(\frac{S_{t=1}}{S_t})$, that are assumed to be normally distributed. The GBM model[1] gives rise to the famous Black-Scholes model for options pricing[2]; this model is commonly used to price both index[3] options and stock options. However, indices are comprised of stocks, and if stock prices follow GBM, then, as we shall see, Index levels cannot (and vice versa). This forces us, in theory at least, to decide between models.

In this work we examine pricing models and hedging strategies on index options, specifically in the case where one constituent stock makes up a significant proportion of the index. This is motivated by a few reasons:

- The first is the use of Black-Scholes one-dimensional model for pricing index options, with this assumption of the underlying asset following a GBM.

- The second is that there have been instances, over the past few decades in particular, where one stock dominates an index on which there are options written or one industry dominates the general index of an exchange - Nortel in the TSX in 2000, Nokia in the Helsinki Stock Exchange until 2007, the banking and/or the utilities sectors in the TSX now; the financial sector makes up about 35% of the S&P/TSX 60 and the energy sector make up another 23%. Another current Canadian example is the S&P/TSX Global Mining Index, of which approximately 32% is made up of BHP Billiton LTD and BHP

---

[1]The Geometric Brownian Motion, GMB, Model refers to the stock price model where the stock price S_t, is the solution to $dS_t = \mu S_t dt + \sigma S_t dW_T$ where W_t is a standard Brownian Motion. The solution to this stochastic differential equation is $S_{t+1} = S_t e^{(\mu - \frac{\sigma^2}{2})dt + \sigma dW_t}$. See discussion in section 2.1

[2]The Black-Scholes model for option pricing refers to the model where the price of an option at time 0 is given by the solution to $\frac{\partial V}{\partial t} + \frac{\sigma^2}{2} S^2 \frac{\partial^2 V}{\partial S^2} + rS \frac{\partial V}{\partial S} = rV$, given the payout function at maturity, $V(T)$. There are analytic solutions to some simple terminal conditions.

[3]An index is a grouping of stocks based on some common attribute, this could be industry, stock exchange, market share, or other attributes. The index value is a linear combination of the values of the constituent stock values. Stocks can be given different weightings and these weightings can be changed over time. Indices are very popular because they allow a small to moderately sized investor to easily diversify their portfolio, which modern finance suggests is the best way to reduce risk with the smallest decrease in returns. Indices also quickly summarize the behaviour of an entire stock exchange or industry which is useful to all investors.

Billiton PLC. The Korean composite index, KOSPI, is currently dominated by Samsung, which comprises nearly 19% of the index.

- Even though when an individual stock becomes dominant it may hit a cap[4] and on the next rebalancing date, quarterly or possibly twice yearly, it could have its weight reduced to the cap, it may remain dominant enough to merit considering the dominant stock separately from the rest of the index. In the case of a dominant industry in an exchange's multi-sector index it can be the case that none of the individual stocks actually hit the cap.

- The last reason to examine this case is that an asset that follows GBM has individual moves that are lognormally distributed, and unlike a normal random variable, the sum of two correlated lognormally distributed random variables is not also lognormally distributed [5]. It is true that the sum of a sufficiently large number of correlated lognormally distributed random variables is a good approximation to a lognormal distribution, but two is not a sufficiently large number, which is the result of one dominating constituent stock and the rest of the index being fairly evenly weighted. This means that even though the Black-Scholes solution is usually pretty good for pricing index options, in this case there may be more unanticipated risks than usual.

When combined these four reasons motivate us to examine this particular case of pricing index options.

In this work we develop and examine various methods of incorporating the dominant constituent stock into the pricing and hedging of options written on a dominated index. This is not a simple task, because there is no analytic solution to the two-dimensional Black-Scholes PDE, so concessions must be made on the accuracy or the computational complexity of a pricing and hedging method. We compare the hedging profits and their respective risks. Finally we look at a real example of such stock paths and compare our pricing and hedging methods.

---

[4]A cap is a maximum weight that an individual asset may contribute to a capped index. A capped index is an index where rebalancing procedures include ensuring that no constituent asset has surpassed the maximum allowed weight, the cap, for an individual asset. In rebalancing a capped index, if one asset has grown to too large a market share then fewer of that stock are included in the index to reduce its contribution. Some indices are capped, but some are not. Capped indices exist to try to avoid exactly the case that we examine in this thesis.

[5]Note that the sum of a sufficiently large number of correlated lognormal random variables will not approach the normal distribution by the central limit theorem because they are correlated, not independent. If they were independent then they would approach the normal distribution, not the lognormal distribution.

Chapter 2

Examining Combining Assets

In this work we endeavour to show that an index, a linear combination of a related set of stocks, may follow a similar model, GBM, as the stocks themselves under certain conditions. We will then examine one case where these conditions are not met, when one constituent stock dominates the index. In this chapter we start by engaging in an exploration of the distribution of indices made up of GBM stocks and examine an empirical example of an index.

In this and later sections we use stochasic calculus which may also be referred to as Ito calculus; it shares many characteristics with regular calculus but with several differences[1]. A stochastic function, say f, is a function of time and path, that is $f(t, X_t)$. The path, X_t is driven by a standard Brownian motion. To interpret a stochastic function we can either allow time to vary and examine a fixed path, this results in a path that is entirely non-differentable with respect to time. We can also fix time and let the path realizaion vary, in this case the result is a distribution. The differential of a stochastic function is found by Ito's lemma:

$$df(t, X_t) = \frac{\partial f}{\partial t} dt + \frac{\partial f}{\partial X_t} dX_t + \frac{1}{2} \frac{\partial^2 f}{\partial X_t^2} (dX_t)^2$$

It is also worth pointing out that in calculating df terms with a factor of dt^m where $m > 1$ are negligible and that for a standard Brownain motion W_t, $(dW_t)^2 \sim dt$.

2.1 GBM as Lognormal

First we explicitly note that an asset that which follows GBM has relative changes that are lognormally distributed. GBM is defined by:

$$dS_t = \mu S_t dt + \sigma S_t dW_t \tag{2.1}$$

With W_t a standard Brownian motion and where

$$W_t - W_s \sim \mathcal{N}(0, t - s)$$

So

$$W_t - W_s \sim (t - s)\mathcal{N}(0, 1)$$

Note that the notation $Z \sim \mathcal{N}(\mu, \sigma^2)$ simply describes that Z is normally distributed with mean μ and variance of σ^2 or standard deviation of σ. If the time between t_i and t_{i+1} is Δt then we let $W_{t_i} - W_{t_{i-1}} = \Delta t Z$ where $Z \sim \mathcal{N}(0, 1)$ then we can write the solution for equation (2.1) as:

$$S_{t_{i+1}} = S_{t_i} e^{(\mu - \frac{\sigma^2}{2})\Delta t + \sigma \sqrt{\Delta t} Z} \tag{2.2}$$

So we can rearrange (2.2) to get:

$$\frac{S_{t_{i+1}}}{S_{t_i}} = e^{(\mu - \frac{\sigma^2}{2})\Delta t + \sigma \sqrt{\Delta t} Z}$$

If we let $a = (\mu - \frac{\sigma^2}{2})\Delta t$ and $b = \sigma \sqrt{\Delta t}$ then we have:

$$\frac{S_{t_{i+1}}}{S_{t_i}} = e^{a+bZ}$$

Which is the same as writing[1]:

$$\frac{S_{t_{i+1}}}{S_{t_i}} \sim \mathcal{LN}(a, b) \tag{2.3}$$

Where $\mathcal{LN}(a, b)$ denoted the lognormal distribution with parameters a and b, for more detail see Appendix A. So it is easy to see from equation (2.3) that an asset following GBM is the same as that asset having relative changes in value that are lognormally distributed.

2.2   Comparing Lognormal to a Sum of Lognormals

Now that we have demonstrated that examining the lognormal distribution is equivalent to examining an asset that follows GBM we will consider how this is related to a stock or an industry dominating an index.

We first assume that all stocks follow GBM, and for evenly weighted index options we also assume that the index of fairly evenly weighted stocks also follow GBM. This is the same as saying that the sum of a sufficiently large number of lognormal random variables is also lognormally distributed. To see if these are contradictory assumptions we will compare the first four moments of various sets of lognormal random variables. See Appendix B for details of the moment calculations.

Our first case will have n random variables that we can easily calculate analytically: the sum of n lognormal random variables that are all correlated by ρ and are identically distributed $\mathcal{LN}(a, b)$. We will start with the standard lognormal, $\mathcal{LN}(0, 1)$, to see how the moments behave.

In figure 2.1 we found the first four moments for a sum of $n = 1, .., 15$ with correlations from 0 to 1, where all summed variables have the same parameters. This can be seen in the top two plots, that the parameters used are constant for all combinations of ρ and n. In the middle two plots it can be seen that the mean is constant across all values of ρ and n but that the variance is increasing in ρ and decreasing in n but bounded above by $n = 1$ and $\rho = 1$. This makes sense because of diversification, the more assets in a portfolio and the lower the correlation between the assets then the lower the volatility. In the bottom plots it looks as though the skewness

Figure 2.1: The parameters and the first four moments of a sum of n lognormally distributed, ρ correlated, random variables. Parameters are fixed at (c,d)=(0,1) and the variance is allowed to vary.
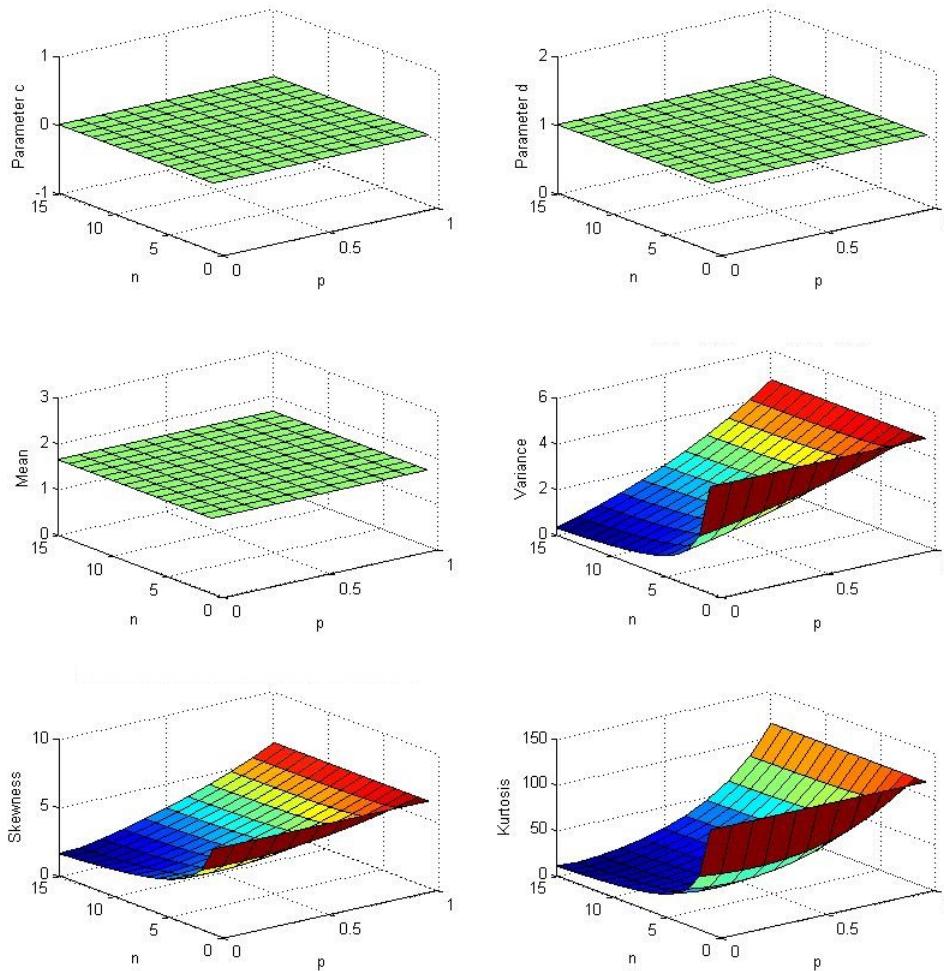
Table 2.1: Initial values to solve equation (2.5)

ρ	equation	formula
0	$x = n\frac{\Upsilon}{M^2} + 1$	——
0.5	$y^2 + (n-1)y - n(\frac{\Upsilon}{M^2} + 1) = 0$	by quadratic formula and $y = \sqrt{x}$
1	$x = \frac{\Upsilon}{M^2} + 1$	——

and the kurtosis also decrease as a function of n and increase as a function of $= \rho$, but this is artificial because the variance follows the same pattern. We should be comparing the tails, kurtosis, between variables with the same variance. To accommodate for the variance we now look at the same plots but we use parameters such that the mean and variance are constant. If we choose a and b to be the parameters for the $n = 1$ case and then find c and d to be the parameters that produce the same mean and variance for any n, ρ pair.

First let $M = \mu_{Y_1(a,b)}$ and $\Upsilon = \sigma^2_{Y_1(a,b)}$, the constant first two moments that we wish to find parameters to match for other values of n and ρ. Then we solve the system of two equations in two unknowns:

$$M = e^{c + \frac{d^2}{2}} \tag{2.4a}$$

$$\Upsilon = \frac{1}{n}(e^{c + \frac{d^2}{2}})^2(e^{d^2} + (n-1)e^{\rho d^2} - n) \tag{2.4b}$$

If we substitute (2.4a) into the equation (2.4b) then we get one equation in one variable:

$$
\begin{aligned}
\Upsilon &= \tfrac{1}{n}M^2(e^{d^2} + (n-1)e^{\rho d^2} - n) \\
\tfrac{n\Upsilon}{M^2} &= e^{d^2} + (n-1)e^{\rho d^2} - n \\
0 &= e^{d^2} + (n-1)e^{\rho d^2} - n(\tfrac{\Upsilon}{M^2} + 1)
\end{aligned} \tag{2.5}
$$

Equation (2.5) can be solved by a root finding method. To do this an initial guess is required (a simple way to do this is to use linear interpolation from the values of ρ for which the equation can be solved analytically). So using the substitution $x = e^{d^2}$ and the initial guesses in Table 2.1 it can be solved by a root finding method, then given d, $c = \ln(M) - \frac{d^2}{2}$.

In figure 2.2 we found the first four moments for a sum of $n = 1, .., 250$ with correlations from 0.2 to 1. In the top two plots the changes required in the parameters to keep the means and variance constant can be seen and in the middle two plots the resulting constant mean and variance can be seen. The bottom two, specifically kurtosis, are the most interesting. It can be seen that for a sum of sufficiently many positively correlated lognormal random variables the third and fourth moments revert to a similar level as that of a single lognormal random variable. It can also be seen that the closer to zero the correlation is the more random variables must be included. The result is that if there are not enough stocks included, or one dominating stock, in an index, especially with low correlation, then pricing and hedging as though the index follows GBM has significantly more tail risk than is accounted for in the model.

There are a few other points to touch on briefly here that are discussed in more depth in the appendices B and C. First, for insufficient correlation the kurtosis blows up with n, but the

Figure 2.2: The parameters and the first four moments of a sum of n lognormally distributed, ρ correlated, random variables. Parameters vary so that first two moments match the $n = 1$ case with parameters $(a,b)=(0,1)$.
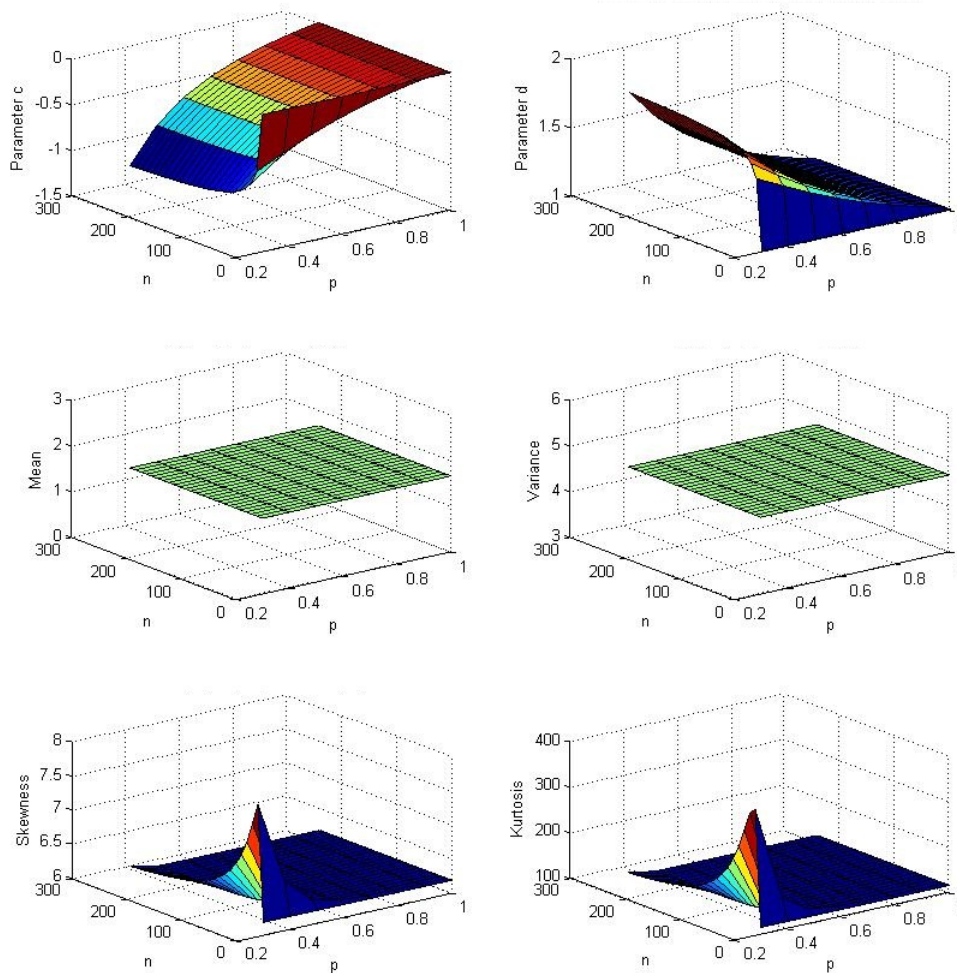
skewness also grows, though linearly, which distorts what the kurtosis tells us about the tails. This must be examined through the use of QQ plots; see appendix A. Second, we see the same behaviour in negatively correlated random variables even in the few points that can be calculated, even though only so many random variables can all have negative correlation to each other[1]. Finally we note that the correlation that we refer to here is between the normal random variables that drive the lognormal random variables, not between the lognormal random variables themselves, they are a transformation on ρ with dependence on n. This is especially prominent when $\rho < 0$.

Since indices are composed for some common attribute, geography, nationality, exchange, or industry, they have common risk factors so their constituent stocks are unlikely to be uncorrelated or negatively correlated. Therefore these cases are not very relevant for our investigation and will not be examined in depth here.

2.3 Empirical Results

Now we will take a quick look at some empirical results. We will look at the iShares S&P TSX Capped Energy Index Fund (XEG)[3,4]. It is, currently, made up of 52 stocks in the Oil Gas & Consumable Fuels sector traded on the TSX at various weightings. We will look at past data for this index and its top ten constituent stocks by weight from January 1st, 2010 to February 28th, 2012.

2.3.1 Log Returns

We will first look at the daily log returns of these assets, not their prices. If we assume that the assets follow a lognormal distribution, $S_{t+1} = S_t e^{a+bZ_t}$, then the log returns follow a normal distribution, $r_t \sim \mathcal{N}(a, b^2)$. Motivated by the discussion in Section 2.1, which showed that the log returns of a GBM model are normally distributed, in this section we investigate the properties of the log returns of real stock data. The statistics we will look at are the first four moments. Note that the first four moments of a normal distribution are $Mean(Z) = \mu$, $Var(Z) = \sigma^2$, $Skew(Z) = 0$, and $Kurt(Z) = 3$. Refer to appendices A and B. Table 2.2 contains the weights of the 10 stocks and the daily moments for the returns of each of the 10 stocks and the index itself.

We can see from table 2.2 that $\mu \approx 0$ and $\sigma \approx 0.3$, which define the drift and volatility of our stocks. We can also see that $skew(S) \approx 0$ and $kurt(S) \approx 4.5$. We can do a student-t test with 10 degrees of freedom to confirm the skewness approximation at the 95%, or $\alpha = 0.05$,

---

[1]See appendix C, section 2 for more detail.

Table 2.2: Weights and moments of returns on XEG stocks, Jan 1, 2010 to Feb 28, 2012, daily frequency

company name	weight	mean	st.dev	skew	kurt
Suncor	17.32	−0.0139	0.3179	−0.0551	4.4681
CNR	11.35	−0.0040	0.3143	0.1159	4.0506
Cenovus	9.05	0.1727	0.3111	0.0287	4.0722
Encana	5.70	−0.2434	0.2912	0.2168	5.5635
Nexen	5.21	−0.0947	0.3285	−0.1308	4.0699
CresPt	4.66	0.1427	0.2149	−0.0031	4.7787
Talisman	4.55	−0.1531	0.3201	−0.2398	4.1782
ImperialOil	4.06	0.0815	0.2295	0.1567	4.4262
CanOilSands	3.67	−0.0683	0.3217	−0.4197	6.1212
Husky	2.8	−0.0182	0.2222	−0.0354	5.0617
Index		0.0021	0.2470	0.1776	6.6401

level[2]:

$$\bar{x} = average(skew(r)) = -0.124$$
$$s = \sqrt{var(skew(r))} = 0.209$$
$$n = 11$$
$$\mu_0 = 0$$

$$H_0 : \mu = \mu_0 \qquad\qquad H_A : \mu \neq \mu_0$$

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$$
$$t = -1.974$$
$$1.974 = |t| < t_{\alpha/2} = 2.228$$

So we do not reject our null hypothesis that our stocks have an average skewness of zero at the 0.05 level. This follows what we would like to see for the log returns to be normally distributed. Now we will do two tests for kurtosis, first for what we were hoping for, $\mu_0 = 3$ and second for $\mu_0 = 4.5$ (a more realistic value for the results we see):

$$\bar{x} = average(kurt(r)) = 4.818$$
$$s = \sqrt{var(kurt(r))} = 0.925$$
$$n = 11$$
$$\mu_0 = 3$$

$$H_0 : \mu = \mu_0 \qquad\qquad H_A : \mu \neq \mu_0$$

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$$
$$t = 6.519$$
$$6.519 = |t| > t_{\alpha/2} = 2.28$$

So we reject our null hypothesis that our returns have the same kurtosis as a normally dis-

Table 2.3: Moments of returns on the top n XEG stocks, unweighted, Jan 1, 2010 to Feb 28, 2012, daily frequency

top n	mean	st.dev	skew	kurt
Top 2	−0.0089	0.2986	0.0782	4.0777
Top 3	0.0455	0.2858	0.0273	4.1394
Top 4	−0.0107	0.2699	−0.0772	4.1910
Top 5	−0.0229	0.2659	−0.0997	4.1530
Top 6	0.0103	0.2426	−0.1254	4.2664
Top 7	−0.0023	0.2425	−0.1492	4.3032
Top 8	0.0120	0.2325	−0.1336	4.2983
Top 9	0.0047	0.2324	−0.1362	4.4038
Top 10	0.0027	0.2267	−0.1340	4.4567

tributed random variable. Now we will test to see how close we are.

$$\mu_0 = 4.5$$

$$H_0 : \mu = \mu_0 \qquad\qquad H_A : \mu \neq \mu_0$$

$$t = \frac{\bar{x} - \mu_0}{s / \sqrt{n}}$$
$$t = 1.141$$
$$1.141 = |t| < t_{\alpha/2} = 2.28$$

So we do not reject the null hypothesis that our returns have a kurtosis a little higher than the normal at 4.5. We can also say that our returns have fatter tails than if they were normally distributed. This means that empirically returns have fatter tails than normal anyway, so when we add two assets together we get even fatter tails.

Next we will look at the moments for the returns of the sum of the top n stocks divided by n (returns of $\frac{1}{n} \sum_{i=1}^{n} S_i$) in table 2.3. This is not weighted as the real index is so it is not entirely realistic, but comparing with the moments of the individual stocks we can infer something. We see that the kurtosis is more affected by the kurtosis of the newly added stock's returns than it is by the number of total stocks included. This means that we are not seeing a reduction in the kurtosis, or the fat tailedness of the returns as more stocks are included. Tails are discussed in appendix A.

Finally, for the sake of being more realistic we will now look at the moments of the returns of the weighted average of the top n stocks as they are weighted in the index in table 2.4. We again see that the kurtosis does not decrease with the number of stocks included, it changes more so with the kurtosis of the newly included stock's kurtosis relative to its weighting.

The last measure of normality of our returns that we will look at is QQ plots. Figure 2.3 shows the QQ plots of each individual stock against a standard normal distribution. In this figure we can see that there is a little discrepancy in our data from the normal distribution between two and three standard deviations from the mean. Overall our data looks fairly normal with slightly fat tails. Next is the QQ plots for the returns of the top n arithmetically averaged

Figure 2.3: QQ plot of the stock returns on XEG's top ten stocks and the whole index against the normal distribution, Jan 1, 2010 to Feb 28, 2012, daily frequency.
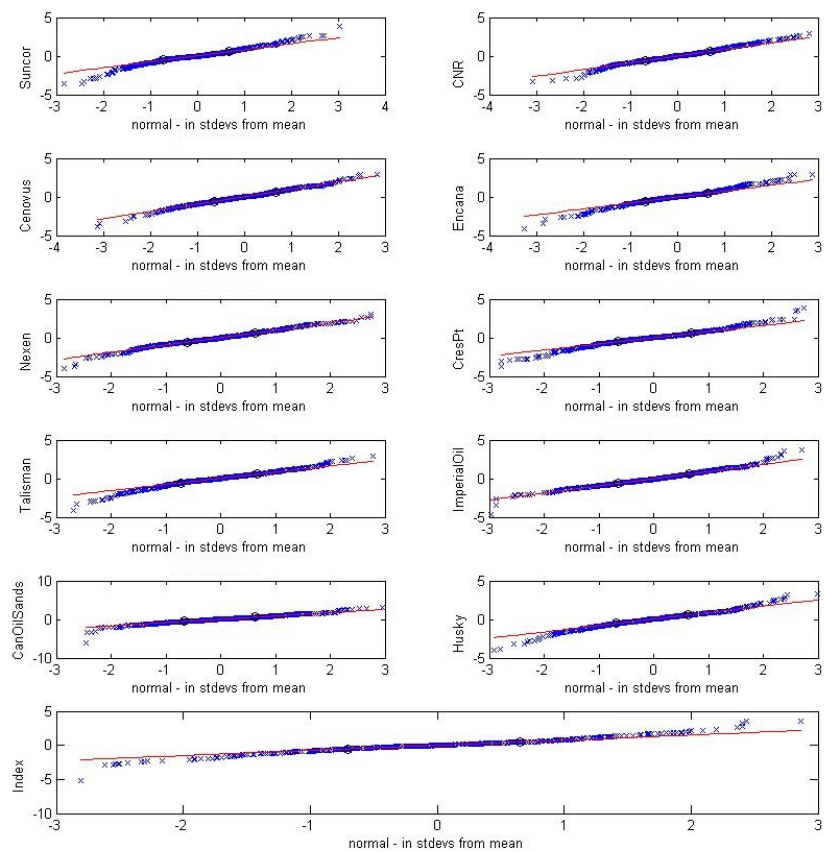
Table 2.4: Moments of returns on the top *n* XEG stocks, weighted, Jan 1, 2010 to Feb 28, 2012, daily frequency

top *n* weighted	*mean*	*st.dev*	*skew*	*kurt*
Top 2	−0.0099	0.2994	0.0577	4.1379
Top 3	0.0290	0.2885	0.0223	4.1605
Top 4	0.0024	0.2792	−0.0359	4.1779
Top 5	−0.0047	0.2760	−0.0483	4.1613
Top 6	0.0102	0.2625	−0.0670	4.1809
Top 7	0.0037	0.2613	−0.0838	4.1875
Top 8	0.0104	0.2543	−0.0855	4.1852
Top 9	0.0070	0.2530	−0.0871	4.2478
Top 10	0.0062	0.2500	−0.0864	4.2739
INDEX	0.0021	0.2470	0.1776	6.6401

stock prices:

In figure 2.4We can see that these are also close to the normal distribution, but not quite. Finally we can look at the returns of a weighted average in the top *n* stocks.

We see the same result again in figure 2.5; the returns of the weighted average in the top *n* stocks is close to normally distributed but with slightly fatter tails.

As a final note on the returns we offer the correlation matrix for the returns between each individual stock in table 2.5. We can see that the returns of individual stocks are highly correlated but the correlation between individual stocks and the whole index are centered around and close to zero. This means that it is possible for a stock to be negatively correlated with the rest of the index so a dominant constituent stock could be negatively correlated with the rest of the index, especially since that case only has two distinct assets so in theory they could be very negatively correlated.

So we can see that even though this index is not evenly weighted, the weights decline steadily, the returns of individual stocks and the returns of the index surprisingly seem to follow a similar distribution. In practice most stock returns have fatter tails than the normal distribution, so indices of such stocks do too, but this means that the GBM model is as applicable to this index as it is to the individual stocks. In the following chapter we will look at a dominated index as if it were made of two distinct, correlated assets: the dominating stock and the rest of the constituent stocks combined into somewhat of a subindex that is sufficiently evenly weighted to be well modeled by GBM.

We have shown that indices are well represented by GBM if they are made up of a large number of positively correlated stocks that are relatively evenly weighted and follow GBM as well. Since in practice this is usually the case, indices are reasonably evenly weighted, much of our work is theoretical, although, if a single asset becomes overly dominant in an index, this picture changes and more tension between the choice of GBM for stocks or for indices exists. Choices must be made about how to model the resulting option pricing problems; next we will

Figure 2.4: QQ plot of the stock returns on the sum of XEG's top $n = 2, 3, ..., 10$ stocks against the normal distribution, Jan 1, 2010 to Feb 28, 2012, daily frequency.
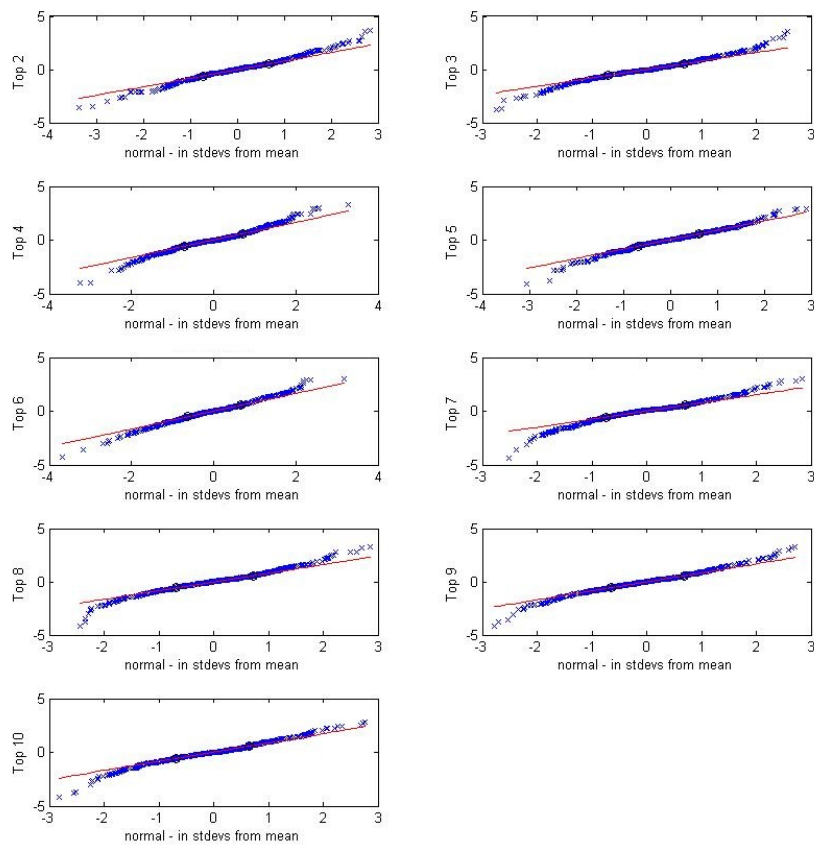
Figure 2.5: QQ plot of the stock returns on the weighted sum of XEG's top $n = 2, 3, ..., 10$ stocks against the normal distribution, Jan 1, 2010 to Feb 28, 2012, daily frequency.
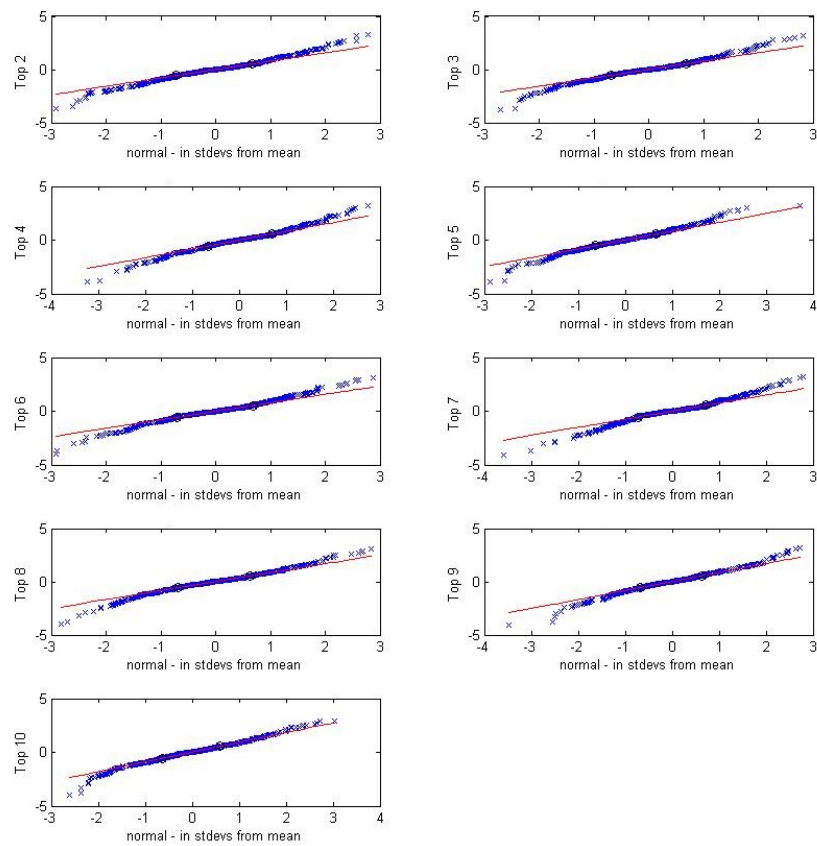
Table 2.5: Correlations between XEG stock returns, Jan 1, 2010 to Feb 28, 2012, daily frequency

	Suncor	CNR	Cenovus	Encana	Nexen	CresPt	
Suncor	1.0000	0.7840	0.7170	0.5860	0.6437	0.5854	...
CNR	0.7840	1.0000	0.7222	0.6233	0.6565	0.6228	...
Cenovus	0.7170	0.7222	1.0000	0.5535	0.5595	0.5804	...
Encana	0.5860	0.6233	0.5535	1.0000	0.5128	0.5031	...
Nexen	0.6437	0.6565	0.5595	0.5128	1.0000	0.4862	...
CresPt	0.5854	0.6228	0.5804	0.5031	0.4862	1.0000	...
Talisman	0.6515	0.6826	0.6002	0.6060	0.5716	0.5239	...
ImperialOil	0.7070	0.6683	0.6631	0.5679	0.5816	0.5512	...
CanOilSands	0.6347	0.6386	0.5826	0.4578	0.4900	0.5270	...
Husky	0.6339	0.6407	0.6025	0.4571	0.5413	0.5352	...
Index	0.0181	−0.0094	−0.0118	0.0099	0.0166	0.0417	...

	Talisman	ImperialOil	CanOilSands	Husky	Index
...	0.6515	0.7070	0.6347	0.6339	0.0181
...	0.6826	0.6683	0.6386	0.6407	−0.0094
...	0.6002	0.6631	0.5826	0.6025	−0.0118
...	0.6060	0.5679	0.4578	0.4571	0.0099
...	0.5716	0.5816	0.4900	0.5413	0.0166
...	0.5239	0.5512	0.5270	0.5352	0.0417
...	1.0000	0.6003	0.5348	0.5301	0.0654
...	0.6003	1.0000	0.5726	0.5955	−0.0416
...	0.5348	0.5726	1.0000	0.5670	−0.0345
...	0.5301	0.5955	0.5670	1.0000	−0.0231
...	0.0654	−0.0416	−0.0345	−0.0231	1.0000

develop methods, that we will later examine, to make this choice.

Chapter 3

Hedging Models for an Index of Two Distinct Assets

In this chapter we will examine hedging methods for an index that fluctuates as if it is made up of two distinct constituent assets: the dominant stock, A, and the rest of the index, B. We will start by deriving a 1+1 dimensional PDE, that is one spatial dimension and one temporal dimension, then we will derive the 2+1 PDE and finally develop a 1+1 PDE approximation to the 2+1 PDE. We will later use the solutions to these PDEs to compare the pricing and hedging merits of each solution.

Say that $I = A + B$ where A and B follow GBM:

$$dA = \mu_A A dt + \sigma_A A dW_t^1 \tag{3.1a}$$

$$dB = \mu_B B dt + \sigma_B B dW_t^2 \tag{3.1b}$$

Where W_t^1 and W_t^2 are standard Brownian Motions and are correlated by ρ, that is $\mathbf{E}[dW_t^1 dW_t^2] = \rho dt$. The problem we want to solve is to price an option on the index, V.

3.1 Hedging with the Index, One-Dimensional Solution

As always the simplest way to price an option is to assume that the underlying asset follows GBM, in our case the index, and price it by the one-dimensional Black-Scholes PDE[5]. This single asset approach ignores the fact that there are two separate underlying assets and therefore only uses σ_I, which depends only on the movements of the index not the individual stocks, and ignores that two separate volatilities may be observable; we will look at that case later. Here we will derive the Black-Scholes one-dimensional PDE. First we assume that

$$dI = \mu_I I dt + \sigma_I I dW_t \tag{3.2}$$

and start with a hedged portfolio containing one option and a countervailing hedge position containing Δ_I units of the underlying asset[5]:

$$\Pi = V - \Delta_I I \tag{3.3}$$

Where Δ_I is the hedging ratio. We want to find a value for Δ_I so that the value of the portfolio changes in a completely predictable way with time: all risk is removed. So with equation (3.3) we calculate $d\Pi$:

$$d\Pi = V_t dt + V_I dI + \frac{1}{2}V_{II}(dI)^2 - \Delta_I dI \tag{3.4a}$$

Notation: V_S denotes the first derivative of V with respect to S, but S_τ denotes $S(\tau)$, where $S = A, B, I$ and τ is any time.

From stochastic calculus and equation (3.2) we know that $(dI)^2 = \sigma_I^2 I^2 dt$ and we let $\Delta_I = V_I$, substitute in $(dI)^2$ and simplify:

$$\begin{aligned}
d\Pi &= V_t dt + \cancel{V_I dI} + \frac{1}{2}V_{II}(dI)^2 - \cancel{\Delta_I dI} \\
d\Pi &= (V_t + \frac{1}{2}V_{II}\sigma_I^2 I^2)dt
\end{aligned} \tag{3.4b}$$

Since this is now a risk free portfolio it must grow at the risk free rate for there to be no arbitrage:

$$\begin{aligned}
d\Pi &= d\Pi \\
r\Pi dt &= (V_t + \frac{1}{2}V_{II}\sigma_I^2 I^2)dt \\
0 &= V_t + \frac{1}{2}V_{II}\sigma_I^2 I^2 + rV_I - rV
\end{aligned} \tag{3.4c}$$

So we get the one-dimensional Black-Scholes PDE

$$V_t + \frac{1}{2}V_{II}\sigma_I^2 I^2 + rV_I - rV = 0 \tag{3.5}$$

With the terminal condition $V(I, T) = G(I_T)$.

3.2　Hedging with Both Constituent Assets

Now we will look at hedging with both stocks; that is, we will derive the two-dimensional Black-Scholes PDE. So we start with the portfolio:

$$\Pi = V - \Delta_A A - \Delta_B B \tag{3.6}$$

Our goal is to find values for both deltas so that the change in the value of the portfolio in equation (3.6) changes only with respect to time, not with respect to either of the Brownian motions. So we find $d\Pi$:

$$d\Pi = dV - \Delta_A dA - \Delta_B dB \tag{3.7a}$$

Where $V = V(I, t) = V(A + B, t) = V(A, B, t)$, so:

$$dV = V_t dt + V_A dA + V_B dB + \frac{1}{2}V_{AA}(dA)^2 + V_{AB}dAdB + \frac{1}{2}V_{BB}(dB)^2 \tag{3.7b}$$

Now substitute dV (3.7b) into $d\Pi$ (3.7a) and set $\Delta_A = V_A$ and $\Delta_B = V_B$ to get the cancellations:

$$
\begin{aligned}
d\Pi &= V_t dt + \cancel{V_A dA} + \cancel{V_B dB} + \tfrac{1}{2}V_{AA}(dA)^2 + V_{AB}dAdB + \tfrac{1}{2}V_{BB}(dB)^2... \\
&\qquad\qquad -\cancel{\Delta_A dA} - \cancel{\Delta_B dB} \\
d\Pi &= V_t dt + \tfrac{1}{2}V_{AA}(dA)^2 + V_{AB}dAdB + \tfrac{1}{2}V_{BB}(dB)^2 \\
d\Pi &= V_t dt + \tfrac{1}{2}V_{AA}\sigma_A^2 A^2 dt + V_{AB}\rho\sigma_A\sigma_B ABdt + \tfrac{1}{2}V_{BB}\sigma_B^2 B^2 dt
\end{aligned}
\tag{3.7c}
$$

Since we now know that $d\Pi$, equation (3.7c), only varies depending on time we know that it is risk free, so we know that it must grow at the risk free rate $d\Pi = r\Pi dt$ for there to be no arbitrage and obviously $d\Pi = d\Pi$ so:

$$
r\Pi dt = V_t dt + \frac{1}{2}V_{AA}\sigma_A^2 A^2 dt + V_{AB}\rho\sigma_A\sigma_B ABdt + \frac{1}{2}V_{BB}\sigma_B^2 B^2 dt
\tag{3.7d}
$$

Which gives us the PDE:

$$
V_t + \frac{1}{2}\sigma_A^2 A^2 V_{AA} + \rho\sigma_A\sigma_B AB V_{AB} + \frac{1}{2}\sigma_B^2 B^2 V_{BB} + rAV_A + rBV_B - rV = 0
\tag{3.8}
$$

With $V(A, B, T) = F(A_T, B_T) = G(A_T + B_T) = G(I_T)$

Since $V(A, B, t) = V(I, t) = V(A + B, t)$ it is easy to think, incorrectly, that no matter what the payoff function of V is $V_A = V_B$ so $\Delta_A = \Delta_B$ which would mean that there is only one, not two, distinct delta values. Maybe we do not need to hedge with both assets individually. Maybe we only need to hedge with the index. We look at this next.

3.3 Hedging with the Index, Two-Dimensional Approximation

If we use both the poor assumption that $\Delta_A = \Delta_B = \Delta_I$ and let $A = \gamma I$ so $B = (1 - \gamma)I$ to reduce our two-dimensional PDE, equation (3.8), becomes

$$
\begin{aligned}
&V_t + \tfrac{1}{2}\sigma_A^2\gamma^2 I^2 V_{II} + \rho\sigma_A\sigma_B\gamma(1-\gamma)I^2 V_{II} \quad ... \\
&+\tfrac{1}{2}\sigma_B^2(1-\gamma)^2 I^2 V_{II} + r\gamma IV_I + r(1-\gamma)IV_I \quad -rV = 0
\end{aligned}
$$

$$
V_t + \frac{1}{2}(\sigma^\star)^2 I^2 V_{II} + rIV_I - rV = 0
\tag{3.9}
$$

With $V(A, B, T) = F(A_T, B_T) = G(A_T + B_T) = G(I_T)$ where

$$
(\sigma^\star)^2 = \sigma_A^2\gamma^2 + 2\rho\sigma_A\sigma_B\gamma(1-\gamma) + \sigma_B^2(1-\gamma)^2
$$

Equation (3.9) is a one-dimensional approximation to the two-dimensional PDE that describes the true solution to our problem. This approximation has its advantages and shortcomings. The first and most obvious advantage is that it can be solved analytically for V and its deltas. A second advantage is that it has a good value for $(\sigma^\star)^2$ because varies depending on how A and B make up I, and how σ_A and σ_B contribute to the index's volatility. This is better

than just pricing the option on the whole index and using σ_I^2, the variance of I. Finally γ can vary through time as the composition of I changes. The shortcomings are in the assumptions. First, even though the functions that are V_A and V_B are the same, if $\sigma_A \neq \sigma_B$ then their values are not necessarily the same. These are not only used in the PDE but also to hedge, so the effect is compounded. Second, the curvature of the pricing surface is lost when we go from V_{AA}, V_{AB}, and V_{BB} to only V_{II}. Finally, dI does not actually follow GBM if both dA and dB do.

Putting the pros and cons aside we will now ensure that equation (3.9) is the PDE that we would get from hedging arguments. We have the same option as above but we will only use the index to hedge, so we have the portfolio:

$$\Pi = V - \Delta_I I \tag{3.10}$$

and $V = V(I, t)$ only with $I = A + B$. With this we calculate $d\Pi$:

$$d\Pi = V_t dt + V_I dI + \frac{1}{2} V_{II}(dI)^2 - \Delta_I dI \tag{3.11a}$$

But dI can not be written in the form $dS = \mu S\, dt + \sigma S\, dW_t$ so it is not a GBM on its own, it is a function of GBMs so we use Ito's lemma:

$$
\begin{aligned}
I &= A + B \\
dI &= 1dA + 1dB + 0 \\
dI &= dA + dB
\end{aligned}
\tag{3.11b}
$$

Before we substitute in we also need $(dI)^2$

$$
\begin{aligned}
(dI)^2 &= (dA + dB)^2 \\
(dI)^2 &= (dA)^2 + 2dAdB + (dB)^2 \\
(dI)^2 &= \sigma_A^2 A^2 dt + 2\rho\sigma_A\sigma_B ABdt + \sigma_B^2 B^2 dt
\end{aligned}
\tag{3.11c}
$$

Now we let $\Delta_I = V_I$, substitute in $(dI)^2$ and simplify:

$$
\begin{aligned}
d\Pi &= V_t dt + \cancel{V_I dI} + \frac{1}{2} V_{II}(dI)^2 - \cancel{\Delta_I dI} \\
d\Pi &= V_t dt + \frac{1}{2} V_{II}(\sigma_A^2 A^2 dt + 2\rho\sigma_A\sigma_B ABdt + \sigma_B^2 B^2 dt) \\
d\Pi &= (V_t + \frac{1}{2}\sigma_A^2 A^2 V_{II} + \rho\sigma_A\sigma_B AB V_{II} + \frac{1}{2}\sigma_B^2 B^2 V_{II})dt
\end{aligned}
\tag{3.11d}
$$

Again, since this is now a risk free portfolio it must grow at the risk free rate for there to be no arbitrage:

$$
\begin{aligned}
d\Pi &= d\Pi \\
r\Pi dt &= (V_t + \frac{1}{2}\sigma_A^2 A^2 V_{II} + \rho\sigma_A\sigma_B AB V_{II} + \frac{1}{2}\sigma_B^2 B^2 V_{II})dt \\
0 &= V_t + \frac{1}{2}\sigma_A^2 A^2 V_{II} + \rho\sigma_A\sigma_B AB V_{II} + \frac{1}{2}\sigma_B^2 B^2 V_{II} + rV_I - rV
\end{aligned}
\tag{3.11e}
$$

With the substitutions $A = \gamma I$ so $B = (1 - \gamma)I$ this is the same as the equation (3.9), the one-dimensional PDE that we derived by reducing the 2-dimensional PDE:

$$V_t + \frac{1}{2}(\sigma^\star)^2 I^2 V_{II} + rIV_I - rV = 0$$

where
$$(\sigma^\star)^2 = \sigma_A^2 \gamma^2 + 2\rho\sigma_A\sigma_B\gamma(1-\gamma) + \sigma_B^2(1-\gamma)^2$$

With $V(A, B, T) = F(A_T, B_T) = G(A_T + B_T) = G(I_T)$. We will refer to this one-dimensional approximation model as the "gamma" model, the "gamma" approximation, or the "gamma" solution going forward.

With these assumptions it appears that the GBM model can be perfectly hedged with either the underlying stocks or the underlying index alone.

So now we have a two-dimensional pricing model and a one-dimensional approximation as well as the simplest one-dimensional Black-Scholes pricing models that we need to compare. To do this we need solutions to the PDEs. For Black-Scholes, and by extension the one-dimensional approximation, we have analytic solutions, but we need to solve the two-dimensional PDE numerically. In the following chapter we will look at two methods to acheive this.

Chapter 4

Solving of the Two-Dimensional PDE

Now that we have derived the two-dimensional PDE (3.8) we need to find a solution for it. In this chapter we develop two separate methods to solve the two-dimensional PDE. The first is a finite difference method to solve the PDE and second, as a benchmark, a Monte Carlo Method. The numerical method we use is the Alternating Direction Implicit (ADI) method[6], a finite difference scheme; we start by simplifying equation (3.8) through a change of variables, then derive the ADI method and discuss its implementation. Finally we will develop a Monte Carlo simulation to solve the PDE.

4.1   ADI Method

This finite difference method, the Alternating Direction Implicit Method[6], is derived from Crank-Nicholson and treats each spatial direction separately.

4.1.1   Reducing the PDE to a heat equation

To simplify our numerical solution we first reduce our PDE to a heat equation by some substitutions. Here we present the short version of this transformation but the long, more intuitive transformations are presented in appendix D.2.

We start with the partial differential equation, the two-dimensional Black-Scholes PDE, equation (3.8). As an example, we show the case of a put option on an index with a dominant constituent stock:

$$V_t + \frac{1}{2}\sigma_A^2 A^2 V_{AA} + \rho\sigma_A\sigma_B ABV_{AB} + \frac{1}{2}\sigma_B^2 B^2 V_{BB} + rAV_A + rBV_B - rV = 0 \qquad (4.1a)$$

with the terminal condition:

$$V(A, B, T) = f(A, B) = (K - A - B)^+ \qquad (4.1b)$$

where $(x)^+ = \max(x, 0)$

Table 4.1: Substitutions to Reduce 2D Black-Scholes PDE to Heat Equation

$$
\begin{array}{llllll}
V(A, B, t) & = & Ke^{-r\tau}v(x, y, \tau) & v(x, y, \tau) & = & \frac{1}{K}e^{r(T-t)}V(A, B, t) \\[2mm]
\hline
t & = & T - \tau & \tau & = & T - t \\[2mm]
A & = & Ke^x e^{-(r-\frac{\sigma_A^2}{2})\tau} & x & = & \ln(\frac{A}{K}) + (r - \frac{\sigma_A^2}{2})\tau \\[2mm]
B & = & Ke^y e^{-(r-\frac{\sigma_B^2}{2})\tau} & y & = & \ln(\frac{B}{K}) + (r - \frac{\sigma_B^2}{2})\tau
\end{array}
$$

and the boundary conditions:

$$
\begin{aligned}
V(0, B, t) & = & g_0(B, t) & = & Ke^{-r(T-t)}\mathcal{N}(-d_2(B)) - B\mathcal{N}(-d_1(B)) \\
V(A, 0, t) & = & h_0(A, t) & = & Ke^{-r(T-t)}\mathcal{N}(-d_2(A)) - A\mathcal{N}(-d_1(A)) \\
V(\infty, B, t) & = & g_\infty(B, t) & = & 0 \\
V(A, \infty, t) & = & h_\infty(A, t) & = & 0
\end{aligned}
\tag{4.1c}
$$

where

$$
\begin{aligned}
d_1(*) & = & \frac{\ln(*/K) + (r + \frac{\sigma_*^2}{2})(T-t)}{\sigma_* \sqrt{T-t}} \\
d_2(*) & = & d_1(*) - \sigma_* \sqrt{T - t} \\
* & = & A, B
\end{aligned}
\tag{4.1d}
$$

To simplify the problem we will use the substitutions in Table 4.1. Here is how the terms of the PDE will change. First, the constant term:

$$
V = Ke^{-r\tau}v
\tag{4.2a}
$$

Second, the derivative with respect to time. For this one we will need the chain rule because $v = v(x(\tau), y(\tau), \tau)$ and later the product rule:

$$
\begin{aligned}
\frac{\partial}{\partial t} & = & \frac{\partial \tau}{\partial t}\frac{\partial}{\partial \tau} + \frac{\partial \tau}{\partial t}\frac{\partial x}{\partial \tau}\frac{\partial}{\partial x} + \frac{\partial \tau}{\partial t}\frac{\partial y}{\partial \tau}\frac{\partial}{\partial y} \\
& = & -\frac{\partial}{\partial \tau} - (r - \frac{\sigma_A^2}{2})\frac{\partial}{\partial x} - (r - \frac{\sigma_B^2}{2})\frac{\partial}{\partial y} \\
\frac{\partial V}{\partial t} & = & \frac{\partial}{\partial t}(Ke^{-r\tau}v) \\
& = & rKe^{-r\tau}v - Ke^{-r\tau}\frac{\partial v}{\partial t} \\
-\frac{\partial V}{\partial t} & = & -rKe^{-r\tau}v + Ke^{-r\tau}\frac{\partial v}{\partial \tau} + Ke^{-r\tau}(r - \frac{\sigma_A^2}{2})\frac{\partial v}{\partial x} + Ke^{-r\tau}(r - \frac{\sigma_B^2}{2})\frac{\partial v}{\partial y}
\end{aligned}
\tag{4.2b}
$$

Next, the first and second spatial derivatives:

$$
\begin{aligned}
\frac{\partial}{\partial A} & = & \frac{\partial x}{\partial A}\frac{\partial}{\partial x} \\
& = & \frac{1}{A}\frac{\partial}{\partial x} \\
\frac{\partial V}{\partial A} & = & Ke^{-r\tau}\frac{1}{A}\frac{\partial v}{\partial x}
\end{aligned}
\tag{4.2c}
$$

$$
\begin{aligned}
\frac{\partial^2}{\partial A^2} &= \frac{\partial}{\partial A}\left(\frac{1}{A}\frac{\partial}{\partial x}\right) \\
&= \frac{1}{A}\frac{\partial}{\partial A}\left(\frac{\partial}{\partial x}\right) - \frac{1}{A^2}\frac{\partial}{\partial x} \\
&= \frac{1}{A^2}\frac{\partial^2}{\partial x^2} - \frac{1}{A^2}\frac{\partial}{\partial x} \\
\frac{\partial^2 V}{\partial A^2} &= Ke^{-r\tau}\left(\frac{1}{A^2}\frac{\partial^2 v}{\partial x^2} - \frac{1}{A^2}\frac{\partial v}{\partial x}\right)
\end{aligned}
\tag{4.2d}
$$

$$
\begin{aligned}
\frac{\partial}{\partial B} &= \frac{\partial y}{\partial B}\frac{\partial}{\partial y} \\
&= \frac{1}{B}\frac{\partial}{\partial y} \\
\frac{\partial V}{\partial B} &= Ke^{-r\tau}\frac{1}{B}\frac{\partial v}{\partial y}
\end{aligned}
\tag{4.2e}
$$

$$
\begin{aligned}
\frac{\partial^2}{\partial B^2} &= \frac{\partial}{\partial B}\left(\frac{1}{B}\frac{\partial}{\partial y}\right) \\
&= \frac{1}{B}\frac{\partial}{\partial B}\left(\frac{\partial}{\partial y}\right) - \frac{1}{B^2}\frac{\partial}{\partial y} \\
&= \frac{1}{B^2}\frac{\partial^2}{\partial y^2} - \frac{1}{B^2}\frac{\partial}{\partial y} \\
\frac{\partial^2 V}{\partial B^2} &= Ke^{-r\tau}\left(\frac{1}{B^2}\frac{\partial^2 v}{\partial y^2} - \frac{1}{B^2}\frac{\partial v}{\partial y}\right)
\end{aligned}
\tag{4.2f}
$$

And the mixed derivative:

$$
\begin{aligned}
\frac{\partial^2}{\partial A \partial B} &= \frac{\partial}{\partial A}\left(\frac{1}{B}\frac{\partial}{\partial y}\right) \\
&= \frac{1}{A}\frac{\partial}{\partial x}\frac{1}{B}\frac{\partial}{\partial y} \\
&= \frac{1}{AB}\frac{\partial^2}{\partial x \partial y} \\
\frac{\partial^2 V}{\partial A \partial B} &= Ke^{-r\tau}\frac{1}{AB}\frac{\partial^2 v}{\partial x \partial y}
\end{aligned}
\tag{4.2g}
$$

Now we substitute the results from equations (4.2) into the original PDE, equation (4.1):

$$
\begin{aligned}
&-rKe^{-r\tau}v + Ke^{-r\tau}\frac{\partial v}{\partial \tau} + Ke^{-r\tau}\left(r - \frac{\sigma_A^2}{2}\right)\frac{\partial v}{\partial x} + Ke^{-r\tau}\left(r - \frac{\sigma_B^2}{2}\right)\frac{\partial v}{\partial y} \\
&= \frac{1}{2}\sigma_A^2 A^2 Ke^{-r\tau}\left(\frac{1}{A^2}\frac{\partial^2 v}{\partial x^2} - \frac{1}{A^2}\frac{\partial v}{\partial x}\right) + \rho\sigma_A\sigma_B ABKe^{-r\tau}\frac{1}{AB}\frac{\partial^2 v}{\partial x \partial y} \\
&+ \frac{1}{2}\sigma_B^2 B^2 Ke^{-r\tau}\left(\frac{1}{B^2}\frac{\partial^2 v}{\partial y^2} - \frac{1}{B^2}\frac{\partial v}{\partial y}\right) + rAKe^{-r\tau}\frac{1}{A}\frac{\partial v}{\partial x} + rBKe^{-r\tau}\frac{1}{B}\frac{\partial v}{\partial y} - rKe^{-r\tau}v
\end{aligned}
$$

Cancel out $Ke^{-r\tau}$ from every term and then cancel $-rv$ from both sides:

$$
\begin{aligned}
&\frac{\partial v}{\partial \tau} + \left(r - \frac{\sigma_A^2}{2}\right)\frac{\partial v}{\partial x} + \left(r - \frac{\sigma_B^2}{2}\right)\frac{\partial v}{\partial y} \\
&= \frac{1}{2}\sigma_A^2 A^2\left(\frac{1}{A^2}\frac{\partial^2 v}{\partial x^2} - \frac{1}{A^2}\frac{\partial v}{\partial x}\right) + \rho\sigma_A\sigma_B AB\frac{1}{AB}\frac{\partial^2 v}{\partial x \partial y} \\
&+ \frac{1}{2}\sigma_B^2 B^2\left(\frac{1}{B^2}\frac{\partial^2 v}{\partial y^2} - \frac{1}{B^2}\frac{\partial v}{\partial y}\right) + rA\frac{1}{A}\frac{\partial v}{\partial x} + rB\frac{1}{B}\frac{\partial v}{\partial y}
\end{aligned}
$$

Simplify within each term:

$$
\begin{aligned}
&\frac{\partial v}{\partial \tau} + \left(r - \frac{\sigma_A^2}{2}\right)\frac{\partial v}{\partial x} + \left(r - \frac{\sigma_B^2}{2}\right)\frac{\partial v}{\partial y} \\
&= \frac{1}{2}\sigma_A^2\left(\frac{\partial^2 v}{\partial x^2} - \frac{\partial v}{\partial x}\right) + \rho\sigma_A\sigma_B\frac{\partial^2 v}{\partial x \partial y} + \frac{1}{2}\sigma_B^2\left(\frac{\partial^2 v}{\partial y^2} - \frac{\partial v}{\partial y}\right) + r\frac{\partial v}{\partial x} + r\frac{\partial v}{\partial y}
\end{aligned}
$$

Rearrange:

$$
\begin{aligned}
&\frac{\partial v}{\partial \tau} + \left(r - \frac{\sigma_A^2}{2}\right)\frac{\partial v}{\partial x} + \left(r - \frac{\sigma_B^2}{2}\right)\frac{\partial v}{\partial y} \\
&= \frac{1}{2}\sigma_A^2\frac{\partial^2 v}{\partial x^2} + \rho\sigma_A\sigma_B\frac{\partial^2 v}{\partial x \partial y} + \frac{1}{2}\sigma_B^2\frac{\partial^2 v}{\partial y^2} + \left(r - \frac{1}{2}\sigma_A^2\right)\frac{\partial v}{\partial x} + \left(r - \frac{1}{2}\sigma_B^2\right)\frac{\partial v}{\partial y}
\end{aligned}
$$

Finally, cancel out the first derivative terms to get the simplified problem, the two-dimensional heat equation:

$$\frac{\partial v}{\partial \tau} = \frac{1}{2}\sigma_A^2 \frac{\partial^2 v}{\partial x^2} + \rho\sigma_A\sigma_B \frac{\partial^2 v}{\partial x \partial y} + \frac{1}{2}\sigma_B^2 \frac{\partial^2 v}{\partial y^2}$$

Therefore solving our original problem is equivalent to solving the two-dimensional heat equation:

$$\frac{\partial v}{\partial \tau} = \frac{1}{2}\sigma_A^2 \frac{\partial^2 v}{\partial x^2} + \rho\sigma_A\sigma_B \frac{\partial^2 v}{\partial x \partial y} + \frac{1}{2}\sigma_B^2 \frac{\partial^2 v}{\partial y^2} \tag{4.3a}$$

with the initial condition:

$$v(x, y, 0) = \frac{1}{K}f(A, B) = (1 - e^x - e^y)^+ \tag{4.3b}$$

and the boundary conditions:

$$
\begin{aligned}
v(-\infty, y, \tau) &= \frac{1}{K}e^{r\tau}\bar{g}_0(Ke^{y-(r-\frac{\sigma_B^2}{2})\tau}, \tau) = \mathcal{N}(-d_2^\star(y)) - e^{y+\frac{\sigma_B^2}{2}\tau}\mathcal{N}(-d_1^\star(y)) \\
v(x, -\infty, \tau) &= \frac{1}{K}e^{r\tau}\bar{h}_0(Ke^{x-(r-\frac{\sigma_A^2}{2})\tau}, \tau) = \mathcal{N}(-d_2^\star(x)) - e^{x+\frac{\sigma_A^2}{2}\tau}e^{r\tau}\mathcal{N}(-d_1^\star(x)) \\
v(\infty, y, \tau) &= \frac{1}{K}e^{r\tau}\bar{g}_\infty(Ke^{y-(r-\frac{\sigma_B^2}{2})\tau}, \tau) = 0 \\
v(x, \infty, \tau) &= \frac{1}{K}e^{r\tau}\bar{h}_\infty(Ke^{x-(r-\frac{\sigma_A^2}{2})\tau}, \tau) = 0
\end{aligned}
\tag{4.3c}
$$

where

$$
\begin{aligned}
d_1^\star(\bullet) &= \frac{\bullet + \sigma_*^2\tau}{\sigma_* \sqrt{\tau}} \\
d_2^\star(\bullet) &= d_1^\star(\bullet) - \sigma_* \sqrt{\tau} = \frac{\bullet}{\sigma_* \sqrt{\tau}} \\
\bullet &= x, y \text{ and } * = A, B
\end{aligned}
\tag{4.3d}
$$

Now we have a two-dimensional PDE that has constant parameters, no first derivative terms and no constant terms. These were all possible sources for numerical error in a numerical method that will no longer contribute to our solution.

4.1.2 Developing the ADI Method

Here we will develop an ADI scheme to solve equation (4.3) numerically. Again the PDE is:

$$\frac{\partial v}{\partial \tau} = \frac{1}{2}\sigma_A^2 \frac{\partial^2 v}{\partial x^2} + \rho\sigma_A\sigma_B \frac{\partial^2 v}{\partial x \partial y} + \frac{1}{2}\sigma_B^2 \frac{\partial^2 v}{\partial y^2}$$

Starting with the original stock prices and τ as parameters we will discretize as shown in Table 4.2. This gives us the discretization for x and y:

$$
\begin{aligned}
x &\rightarrow X(p, n) &= \ln(\frac{p \times f}{K}) + (r - \frac{\sigma_A^2}{2})n \times k \\
y &\rightarrow Y(q, n) &= \ln(\frac{q \times g}{K}) + (r - \frac{\sigma_B^2}{2})n \times k
\end{aligned}
$$

Table 4.2: Discretization to Solve Equation (4.3) Numerically

cont.	$\rightarrow$	descr.	counter	step size	bounds
A	$\rightarrow$	$p \times f$	$p = 0...P$	f	$P = Amax/f$
B	$\rightarrow$	$q \times g$	$q = 0...Q$	g	$Q = Bmax/g$
τ	$\rightarrow$	$n \times k$	$n = 0...N$	h	$L = T/h$

The notation used is:

$$v(x, y, \tau) = v(x(A, \tau), y(B, \tau), \tau) \rightarrow u(X(p, n), Y(q, n), n) = u_{p,q}^n$$

We use the Peaceman-Rachford algorithm[6,7], which starts like Crank-Nicolson[6,8], by centering the difference scheme around $\tau = (n+\frac{1}{2})h$. For the temporal derivative we use the central difference:

$$\left.\frac{\partial u}{\partial \tau}\right|_{\tau=(n+\frac{1}{2})h} = \delta_t u^{n+\frac{1}{2}} = \frac{u^{n+1} - u^n}{h} + O(h^2) \tag{4.4a}$$

For the spatial derivatives we will use the central difference operator again, but first we will use the average at $\tau = nh$ and $\tau = (n + 1)h$ to get $\tau = (n + \frac{1}{2})h$:

$$\left.\frac{\partial^2 u}{\partial x^2}\right|_{\tau=(n+\frac{1}{2})h} = \frac{1}{2}\left(\left.\frac{\partial^2 u}{\partial x^2}\right|_{\tau=(n+1)h} + \left.\frac{\partial^2 u}{\partial x^2}\right|_{\tau=nh}\right) = \frac{1}{2}\delta_x^2(u^{n+1} + u^n) \tag{4.4b}$$

$$\left.\frac{\partial^2 u}{\partial x \partial y}\right|_{\tau=(n+\frac{1}{2})h} = \frac{1}{2}\left(\left.\frac{\partial^2 u}{\partial x \partial y}\right|_{\tau=(n+1)h} + \left.\frac{\partial^2 u}{\partial x \partial y}\right|_{\tau=nh}\right) = \frac{1}{2}\delta_x\delta_y(u^{n+1} + u^n) \tag{4.4c}$$

$$\left.\frac{\partial^2 u}{\partial y^2}\right|_{\tau=(n+\frac{1}{2})h} = \frac{1}{2}\left(\left.\frac{\partial^2 u}{\partial y^2}\right|_{\tau=(n+1)h} + \left.\frac{\partial^2 u}{\partial y^2}\right|_{\tau=nh}\right) = \frac{1}{2}\delta_y^2(u^{n+1} + u^n) \tag{4.4d}$$

So our PDE is now:

$$\frac{u^{n+1} - u^n}{h} = \left(\frac{\sigma_A^2}{4}\delta_x^2 + \frac{1}{2}\rho\sigma_A\sigma_B\delta_x\delta_y + \frac{\sigma_B^2}{4}\delta_y^2\right)(u^{n+1} + u^n) + O(h^2) \tag{4.5a}$$

Then we bring all u^{n+1} terms to the left and all the u^n to the right (except for the mixed derivative, that all stays on the right):

$$\left(1 - \frac{h\sigma_A^2}{4}\delta_x^2 - \frac{h\sigma_B^2}{4}\delta_y^2\right)u^{n+1} = \left(1 + \frac{h\sigma_A^2}{4}\delta_x^2 + \frac{h\sigma_B^2}{4}\delta_y^2\right)u^n$$
$$+ \frac{h}{2}\rho\sigma_A\sigma_B\delta_x\delta_y(u^{n+1} + u^n) + O(h^3) \tag{4.5b}$$

Now we will define operators for our equation:

$$
\begin{aligned}
D_1 &= \frac{\sigma_A^2}{2}\delta_x^2 \\
D_2 &= \frac{\sigma_B^2}{2}\delta_y^2
\end{aligned}
\tag{4.5c}
$$

And our equation becomes:

$$
\begin{aligned}
\left(1 - \tfrac{h}{2}D_1 - \tfrac{h}{2}D_2\right)u^{n+1} &= \left(1 + \tfrac{h}{2}D_1 + \tfrac{h}{2}D_2\right)u^n \\
&+ \tfrac{h}{2}\rho\sigma_A\sigma_B\delta_x\delta_y(u^{n+1} + u^n) + O(h^3)
\end{aligned}
\tag{4.5d}
$$

To be able to factor we add $\frac{h^2}{4}D_1 D_2 u^{n+1}$ to both sides and on the right side add and subtract $\frac{h^2}{4}D_1 D_2 u^n$. The factorization we use is:

$$
1 \pm d_1 \pm d_2 + d_1 d_2 = (1 \pm d_1)(1 \pm d_2)
$$

So we get:

$$
\begin{aligned}
\left(1 - \tfrac{h}{2}D_1\right)\left(1 - \tfrac{h}{2}D_2\right)u^{n+1} &= \left(1 + \tfrac{h}{2}D_1\right)\left(1 + \tfrac{h}{2}D_2\right)u^n \\
+ \tfrac{h^2}{4}D_1 D_2(u^{n+1} - u^n) &+ \tfrac{h}{2}\rho\sigma_A\sigma_B\delta_x\delta_y(u^{n+1} + u^n) + O(h^3)
\end{aligned}
\tag{4.5e}
$$

But $u^{n+1} - u^n = O(h)$ so $h^2(u^{n+1} - u^n) = O(h^3)$ and that term can be absorbed into the existing $O(h^3)$ term.

$$
\begin{aligned}
\left(1 - \tfrac{h}{2}D_1\right)\left(1 - \tfrac{h}{2}D_2\right)u^{n+1} &= \left(1 + \tfrac{h}{2}D_1\right)\left(1 + \tfrac{h}{2}D_2\right)u^n \\
&+ \tfrac{h}{2}\rho\sigma_A\sigma_B\delta_x\delta_y(u^{n+1} + u^n) + O(h^3)
\end{aligned}
\tag{4.6a}
$$

This can be split to give the Alternating Direction Implicit method for our PDE:

$$
\boxed{
\begin{aligned}
\left(1 - \tfrac{h}{2}D_1\right)\tilde{u} &= \left(1 + \tfrac{h}{2}D_2\right)u^n &+& \tfrac{h}{2}\rho\sigma_A\sigma_B\delta_x\delta_y u^n \\
\left(1 - \tfrac{h}{2}D_2\right)u^{n+1} &= \left(1 + \tfrac{h}{2}D_1\right)\tilde{u} &+& \tfrac{h}{2}\rho\sigma_A\sigma_B\delta_x\delta_y \tilde{u}
\end{aligned}
}
\tag{4.6b}
$$

Which, as before, is $O(h^3)$. When we define our spatial differential operators then we will get error bounds with respect to our spatial discretizations.

Note that in equation (4.6b) at time $n + 1$ everything on the right-hand side of the first step is known and the left side is unknown; after the first step everything on the right-hand side is known and the left side is unknown; after the second step u^{n+1} is now known.

Equation (4.6b) is a single step scheme, so we must only have an initial condition at the first time step, and no priming is needed to start using our algorithm. $u_{p,q}^0$ is known for all p, q.

But before we get too confident about the high accuracy of our scheme we should check that our original scheme, equation (4.6a), and our two part ADI scheme, equation (4.6b), are equivalent. We start with the second part of our ADI scheme and operate on it with $\left(1 - \tfrac{h}{2}D_1\right)$

$$
\left(1 - \frac{h}{2}D_1\right)\left(1 - \frac{h}{2}D_2\right)u^{n+1} = \left(1 + \frac{h}{2}D_1\right)\left(1 - \frac{h}{2}D_1\right)\tilde{u}
$$

$$+\frac{h}{2}\rho\sigma_A\sigma_B\delta_x\delta_y\left(1-\frac{h}{2}D_1\right)\tilde{u} + O(h^3)$$

substituting in the first part of our ADI scheme we get:

$$\left(1-\frac{h}{2}D_1\right)\left(1-\frac{h}{2}D_2\right)u^{n+1} = \left(1+\frac{h}{2}D_1\right)\left(\left(1+\frac{h}{2}D_2\right)u^n + \frac{h}{2}\rho\sigma_A\sigma_B\delta_x\delta_y u^n\right)$$

$$+\frac{h}{2}\rho\sigma_A\sigma_B\delta_x\delta_y\left(\left(1+\frac{h}{2}D_2\right)u^n + \frac{h}{2}\rho\sigma_A\sigma_B\delta_x\delta_y u^n\right) + O(h^3)$$

which simplifies to:

$$\left(1-\frac{h}{2}D_1\right)\left(1-\frac{h}{2}D_2\right)u^{n+1} = \left(1+\frac{h}{2}D_1\right)\left(1+\frac{h}{2}D_2\right)u^n$$

$$+\frac{h}{2}\rho\sigma_A\sigma_B\delta_x\delta_y\left[\left(1+\frac{h}{2}D_1\right) + \frac{h}{2}\rho\sigma_A\sigma_B\delta_x\delta_y + \left(1+\frac{h}{2}D_2\right)\right]u^n + O(h^3)$$

which only matches our original scheme if we allow our error term to absorb the other h terms. So our scheme has an order of accuracy $O(h)$ not $O(h^3)$.

If we allow our new error term $O(h)$ to absorb the h terms before we start this re-arrangement we see that the two schemes do match with first order accuracy. Our original scheme is now only $O(h)$,

$$\left(1-\frac{h}{2}D_1\right)\left(1-\frac{h}{2}D_2\right)u^{n+1} = \left(1+\frac{h}{2}D_1\right)\left(1+\frac{h}{2}D_2\right)u^n + O(h) \qquad (4.7a)$$

and the ADI scheme is also only $O(h)$,

$$\begin{array}{|rcccc|}\hline \left(1-\frac{h}{2}D_1\right)\tilde{u} & = & \left(1+\frac{h}{2}D_2\right)u^n & + & O(h) \\ \left(1-\frac{h}{2}D_2\right)u^{n+1} & = & \left(1+\frac{h}{2}D_1\right)\tilde{u} & + & O(h) \\ \hline \end{array} \qquad (4.7b)$$

To show that equations (4.7a) and (4.7b) are equivalent we start with the second part of our ADI scheme and operate on it with $\left(1-\frac{h}{2}D_1\right)$

$$\left(1-\frac{h}{2}D_1\right)\left(1-\frac{h}{2}D_2\right)u^{n+1} = \left(1+\frac{h}{2}D_1\right)\left(1-\frac{h}{2}D_1\right)\tilde{u} + O(h)$$

and substitute in the first part of our ADI scheme to get

$$\left(1-\frac{h}{2}D_1\right)\left(1-\frac{h}{2}D_2\right)u^{n+1} = \left(1+\frac{h}{2}D_1\right)\left(1+\frac{h}{2}D_2\right)u^n + O(h)$$

which is our original scheme. Therefore the ADI scheme in equation (4.7b), with the mixed derivative, is only first order accurate in time.

Boundary conditions: On the boundaries we know u^n for any n, that is we know $u^n_{0,q}$, $u^n_{p,0}$, $u^n_{P,q}$, and $u^n_{p,Q}$, but we don't yet have any boundary values of $\tilde{u}$. From our scheme we can get $\tilde{u}^n_{0,q}$

and $\tilde{u}^n_{P,q}$, but not $\tilde{u}^n_{p,0}$ and $\tilde{u}^n_{p,Q}$. Without the cross derivative $\tilde{u}^n_{p,0}$ and $\tilde{u}^n_{p,Q}$ are no longer required. To find $\tilde{u}^n_{0,q}$ and $\tilde{u}^n_{P,q}$ we add the first step to the opposite side of the second step as follows:

$$
\begin{aligned}
\left(1 - \tfrac{h}{2}D_1\right)\tilde{u} & = \left(1 + \tfrac{h}{2}D_2\right)u^n + \tfrac{h}{2}\rho\sigma_A\sigma_B\delta_x\delta_y u^n \\
+ \quad \left(1 + \tfrac{h}{2}D_1\right)\tilde{u} + \tfrac{h}{2}\rho\sigma_A\sigma_B\delta_x\delta_y\tilde{u} & = \left(1 - \tfrac{h}{2}D_2\right)u^{n+1} \\
\hline
\left(2 + \tfrac{h}{2}\rho\sigma_A\sigma_B\delta_x\delta_y\right)\tilde{u} & = \left(1 - \tfrac{h}{2}D_2\right)u^{n+1} + \left(1 + \tfrac{h}{2}D_2\right)u^n \\
& \quad + \tfrac{h}{2}\rho\sigma_A\sigma_B\delta_x\delta_y u^n + O(h)
\end{aligned}
$$

$$
2\tilde{u} = \left(1 - \frac{h}{2}D_2\right)u^{n+1} + \left(1 + \frac{h}{2}D_2\right)u^n + \frac{h}{2}\rho\sigma_A\sigma_B\delta_x\delta_y(u^n - \tilde{u}) + O(h)
$$

$$
\tilde{u} = \frac{1}{2}\left(1 - \frac{h}{2}D_2\right)u^{n+1} + \frac{1}{2}\left(1 + \frac{h}{2}D_2\right)u^n + \frac{h}{4}\rho\sigma_A\sigma_B\delta_x\delta_y(u^n - \tilde{u}) + O(h) \tag{4.8}
$$

In equation (4.8) the mixed term is of order $O(h)$ and it must be canceled out and added to the error to get the boundary conditions, this reinforces that the order of this scheme is only $O(h)$.

$$
\tilde{u} = \frac{1}{2}\left(1 - \frac{h}{2}D_2\right)u^{n+1} + \frac{1}{2}\left(1 + \frac{h}{2}D_2\right)u^n + O(h)
$$

With a slight modification to the above scheme we can get back up to second order in time but it requires involving a third time step. This requires priming the grid to get a second level of initial condition, which is not necessarily simple. To maintain our second order accuracy the priming step must also be second order accurate in time. Here is the modification:

$$
\boxed{
\begin{aligned}
\left(1 - \tfrac{h}{2}D_1\right)\tilde{u} & = \left(1 + \tfrac{h}{2}D_2\right)u^n & + & \tfrac{h}{2}\rho\sigma_A\sigma_B\delta_x\delta_y\bar{u} + O(h^2) \\
\left(1 - \tfrac{h}{2}D_2\right)u^{n+1} & = \left(1 + \tfrac{h}{2}D_1\right)\tilde{u} & + & \tfrac{h}{2}\rho\sigma_A\sigma_B\delta_x\delta_y\bar{u} + O(h^2)
\end{aligned}
} \tag{4.9}
$$

where $\bar{u} = \frac{3}{2}u^n - \frac{1}{2}u^{n-1} + O(h^2)$

This is from the second order Taylor expansion with the backward difference, first order, derivative approximation:

$$
\begin{aligned}
\bar{u} = u^{n-\frac{1}{2}} & = u^n + \tfrac{h}{2}\delta_\tau u^n + O(h^2) \\
\bar{u} & = u^n + \tfrac{h}{2}\left(\tfrac{u^n - u^{n-1}}{h} + O(h)\right) + O(h^2) \\
\bar{u} & = \tfrac{3}{2}u^n - \tfrac{1}{2}u^{n-1} + O(h^2)
\end{aligned}
$$

so the boundary condition works out as follows:

$$
\begin{aligned}
\left(1 - \tfrac{h}{2}D_1\right)\tilde{u} & = \left(1 + \tfrac{h}{2}D_2\right)u^n + \tfrac{h}{2}\rho\sigma_A\sigma_B\delta_x\delta_y\bar{u} \\
+ \quad \left(1 + \tfrac{h}{2}D_1\right)\tilde{u} + \tfrac{h}{2}\rho\sigma_A\sigma_B\delta_x\delta_y\bar{u} & = \left(1 - \tfrac{h}{2}D_2\right)u^{n+1} \\
\hline
2\tilde{u} + \tfrac{h}{2}\rho\sigma_A\sigma_B\delta_x\delta_y\bar{u} & = \left(1 - \tfrac{h}{2}D_2\right)u^{n+1} + \left(1 + \tfrac{h}{2}D_2\right)u^n \\
& \quad + \tfrac{h}{2}\rho\sigma_A\sigma_B\delta_x\delta_y\bar{u} + O(h^2)
\end{aligned}
$$

$$\tilde{u} = \frac{1}{2}\left(1 - \frac{h}{2}D_2\right)u^{n+1} + \frac{1}{2}\left(1 + \frac{h}{2}D_2\right)u^n + O(h^2) \tag{4.10}$$

Which maintains our second order accuracy.

For the priming step we will use the second order accurate forward time central space scheme. We start with the discretization described in Table 4.2 and our heat equation, equation (4.3) :

$$\frac{\partial v}{\partial \tau} = \frac{1}{2}\sigma_A^2\frac{\partial^2 v}{\partial x^2} + \rho\sigma_A\sigma_B\frac{\partial^2 v}{\partial x\partial y} + \frac{1}{2}\sigma_B^2\frac{\partial^2 v}{\partial y^2}$$

Then we use the first order forward difference approximation for the temporal derivative:

$$\left.\frac{\partial u}{\partial \tau}\right|_{\tau=nh} = \delta_t u^n = \frac{u_{p,q}^{n+1} - u_{p,q}^n}{h} + O(h) \tag{4.11a}$$

for the second order spatial derivatives we will use the second order central differences:

$$\left.\frac{\partial^2 u}{\partial x^2}\right|_{x=pf} = \delta_x^2 u_{p,q}^n = \left(\frac{u_{p+1,q}^n - 2u_{p,q}^n + u_{p-1,q}^n}{f^2}\right) + O(f^2) \tag{4.11b}$$

$$\left.\frac{\partial^2 u}{\partial y^2}\right|_{y=qg} = \delta_y^2 u_{p,q}^n = \left(\frac{u_{p,q+1}^n - 2u_{p,q}^n + u_{p,q-1}^n}{g^2}\right) + O(g^2) \tag{4.11c}$$

$$\left.\frac{\partial^2 u}{\partial x\partial y}\right|_{x=pf,y=qg} = \delta_x\delta_y u_{p,q}^n = \left(\frac{u_{p+1,q+1}^n - u_{p+1,q-1}^n - u_{p-1,q+1}^n + u_{p-1,q-1}^n}{2fg}\right) \\ + O(f^2 + g^2) \tag{4.11d}$$

So our PDE becomes:

$$\frac{u_{p,q}^{n+1} - u_{p,q}^n}{h} + O(h) = \frac{1}{2}\sigma_A^2\left(\frac{u_{p+1,q}^n - 2u_{p,q}^n + u_{p-1,q}^n}{f^2}\right)\rho\sigma_A\sigma_B\left(\frac{u_{p+1,q+1}^n - u_{p+1,q-1}^n - u_{p-1,q+1}^n + u_{p-1,q-1}^n}{2fg}\right) \\ + \frac{1}{2}\sigma_B^2\left(\frac{u_{p,q+1}^n - 2u_{p,q}^n + u_{p,q-1}^n}{g^2}\right) + O(f^2 + g^2)$$

So solving for $u_{p,q}^{n+1}$ we get:

$$u_{p,q}^{n+1} = u_{p,q}^n\left(1 - \frac{h\sigma_A^2}{f^2} - \frac{h\sigma_B^2}{g^2}\right) + \frac{h\sigma_A^2}{2f^2}\left(u_{p+1,q}^n + u_{p-1,q}^n\right) + \frac{h\sigma_B^2}{2g^2}\left(u_{p,q+1}^n + u_{p,q-1}^n\right) \\ + \frac{h\rho\sigma_A\sigma_B}{2fg}\left(u_{p+1,q+1}^n - u_{p+1,q-1}^n - u_{p-1,q+1}^n + u_{p-1,q-1}^n\right) + O(hf^2 + hg^2 + h^2) \tag{4.12}$$

for $p = 1...P - 1$, $q = 1...Q - 1$ and $n = 1$. At $n = 0$ we use our initial condition and for $p = 0, P$ and $q = 0, Q$ we use the boundary conditions. Now we can solve for our whole pricing surface at the first time step and use our ADI scheme to solve for the subsequent time steps. This scheme is second order accurate in time and space so our ADI scheme's second order accuracy remains intact after priming.

Note that we do not use the forward time central space scheme for all the time steps because it is only conditionally stable, at best, and the ADI scheme is unconditionally stable because it is derived from the Crank-Nicholson scheme.

Table 4.3: Parameters Required as Input to ADI Method

r	–	risk free interest rate
K	–	strike price of the put
T	–	time to maturity
σ_A	–	standard deviation on the returns for A
σ_B	–	standard deviation on the returns for B
ρ	–	correlation between the returns of A and B
f	–	spatial step size, for A
g	–	spatial step size, for B
h	–	temporal step size

4.1.3   Implementing the ADI Method

To implement our scheme we need to be given the parameters in table 4.3 and calculate suffi-
ciently large values for maximum values for *A* and *B* to create a false upper boundary. For the
option on a dominated index this is:

$$
\begin{aligned}
Amax &= \max(\lceil Ke^{-(r-\frac{\sigma_A^2}{2})T+\star\sigma_A\sqrt{T}}\rceil, K+1) \\
Bmax &= \max(\lceil Ke^{-(r-\frac{\sigma_B^2}{2})T+\star\sigma_B\sqrt{T}}\rceil, K+1)
\end{aligned}
\tag{4.13}
$$

where $\star$ is the number of standard deviations beyond *K* that *A* or *B* need to be by maturity
(the number of standard deviations implies the confidence that *A* or *B* will remain within this
bound in the relevant time frame).

Next we need to decide what order of accuracy we wish to use for our spatial differences.
We will use second order accurate approximations which are sufficient for our purpose and
easy to derive for second order derivatives and for the mixed derivative:

$$
D_1 u_{p,q}^n = \frac{\sigma_A^2}{2}\delta_x^2 u_{p,q}^n = \frac{\sigma_A^2}{2}\left(\frac{u_{p+1,q}^n - 2u_{p,q}^n + u_{p-1,q}^n}{f^2}\right) + O(f^2)
\tag{4.14a}
$$

$$
D_2 u_{p,q}^n = \frac{\sigma_B^2}{2}\delta_y^2 u_{p,q}^n = \frac{\sigma_B^2}{2}\left(\frac{u_{p,q+1}^n - 2u_{p,q}^n + u_{p,q-1}^n}{g^2}\right) + O(g^2)
\tag{4.14b}
$$

$$
\rho\sigma_A\sigma_B\delta_x\delta_y u_{p,q}^n = \rho\sigma_A\sigma_B\left(\frac{u_{p+1,q+1}^n - u_{p+1,q-1}^n - u_{p-1,q+1}^n + u_{p-1,q-1}^n}{4fg}\right) + O(f^2 + g^2)
\tag{4.14c}
$$

So the first step is:

$$
\tilde{u}_{p,q} - \frac{h}{2}\frac{\sigma_A^2}{2}\left(\frac{\tilde{u}_{p+1,q} - 2\tilde{u}_{p,q} + \tilde{u}_{p-1,q}}{f^2}\right)
$$

$$= u_{p,q}^n + \frac{h}{2}\frac{\sigma_B^2}{2}\left(\frac{u_{p,q+1}^n - 2u_{p,q}^n + u_{p,q-1}^n}{g^2}\right)$$

$$+\frac{h}{2}\rho\sigma_A\sigma_B\left(\frac{\bar{u}_{p+1,q+1} - \bar{u}_{p+1,q-1} - \bar{u}_{p-1,q+1} + \bar{u}_{p-1,q-1}}{4fg}\right)$$

Simplified to:

$$-\frac{h\sigma_A^2}{4f^2}\tilde{u}_{p+1,q} + \left(1 + \frac{h\sigma_A^2}{2f^2}\right)\tilde{u}_{p,q} - \frac{h\sigma_A^2}{4f^2}\tilde{u}_{p-1,q}$$

$$= \frac{h\sigma_B^2}{4g^2}u_{p,q+1}^n + \left(1 - \frac{h\sigma_B^2}{2g^2}\right)u_{p,q}^n + \frac{h\sigma_B^2}{4g^2}u_{p,q-1}^n \qquad (4.15)$$

$$+\frac{h\rho\sigma_A\sigma_B}{8fg}\left(\bar{u}_{p+1,q+1} - \bar{u}_{p+1,q-1} - \bar{u}_{p-1,q+1} + \bar{u}_{p-1,q-1}\right)$$

The second step is:

$$u_{p,q}^{n+1} - \frac{h}{2}\frac{\sigma_B^2}{2}\left(\frac{u_{p,q+1}^{n+1} - 2u_{p,q}^{n+1} + u_{p,q-1}^{n+1}}{g^2}\right)$$

$$= \tilde{u}_{p,q} + \frac{h}{2}\frac{\sigma_A^2}{2}\left(\frac{\tilde{u}_{p+1,q} - 2\tilde{u}_{p,q} + \tilde{u}_{p-1,q}}{f^2}\right)$$

$$+\frac{h}{2}\rho\sigma_A\sigma_B\left(\frac{\bar{u}_{p+1,q+1} - \bar{u}_{p+1,q-1} - \bar{u}_{p-1,q+1} + \bar{u}_{p-1,q-1}}{4fg}\right)$$

Simplified to:

$$-\frac{h\sigma_B^2}{4g^2}u_{p,q+1}^{n+1} + \left(1 + \frac{h\sigma_B^2}{2g^2}\right)u_{p,q}^{n+1} - \frac{h\sigma_B^2}{4g^2}u_{p,q-1}^{n+1}$$

$$= \frac{h\sigma_A^2}{4f^2}\tilde{u}_{p+1,q} + \left(1 - \frac{h\sigma_A^2}{2f^2}\right)\tilde{u}_{p,q} + \frac{h\sigma_A^2}{4f^2}\tilde{u}_{p-1,q} \qquad (4.16)$$

$$+\frac{h\rho\sigma_A\sigma_B}{8fg}\left(\bar{u}_{p+1,q+1} - \bar{u}_{p+1,q-1} - \bar{u}_{p-1,q+1} + \bar{u}_{p-1,q-1}\right)$$

These can both be solved for their respective unknowns in matrix form at every time step given initial and boundary conditions. For the first ADI step from the form $Ax = b$ we have for $q = 1 : Q - 1$:

$$A = \begin{bmatrix}
\left(1 + \frac{h\sigma_A^2}{2f^2}\right) & \frac{-h\sigma_A^2}{4f^2} & 0 & \cdots & \cdots & 0 \\
\frac{-h\sigma_A^2}{4f^2} & \left(1 + \frac{h\sigma_A^2}{2f^2}\right) & \frac{-h\sigma_A^2}{4f^2} & 0 & \cdots & 0 \\
0 & \ddots & \ddots & \ddots & 0 & \vdots \\
\vdots & 0 & \ddots & \ddots & \ddots & 0 \\
\vdots & \vdots & 0 & \frac{-h\sigma_A^2}{4f^2} & \left(1 + \frac{h\sigma_A^2}{2f^2}\right) & \frac{-h\sigma_A^2}{4f^2} \\
0 & 0 & \cdots & 0 & \frac{-h\sigma_A^2}{4f^2} & \left(1 + \frac{h\sigma_A^2}{2f^2}\right)
\end{bmatrix}$$

$$x = \begin{bmatrix}
\tilde{u}_{1,q} \\
\tilde{u}_{2,q} \\
\vdots \\
\vdots \\
\tilde{u}_{P-2,q} \\
\tilde{u}_{P-1,q}
\end{bmatrix}$$

$b =$

$$
\begin{bmatrix}
\frac{h\sigma_B^2}{4g^2} u_{1,q+1}^n + \left(1 - \frac{h\sigma_B^2}{2g^2}\right) u_{1,q}^n + \frac{h\sigma_B^2}{4g^2} u_{1,q-1}^n + \frac{h\rho\sigma_A\sigma_B}{8fg} \left(\bar{u}_{1,q+1} \right. \\
\left. \quad \dots\, -\bar{u}_{1,q-1} - \bar{u}_{0,q+1} + \bar{u}_{0,q-1}\right) + \frac{h\sigma_A^2}{4f^2} \tilde{u}_{0,q} \\[1em]
\frac{h\sigma_B^2}{4g^2} u_{2,q+1}^n + \left(1 - \frac{h\sigma_B^2}{2g^2}\right) u_{2,q}^n + \frac{h\sigma_B^2}{4g^2} u_{2,q-1}^n + \frac{h\rho\sigma_A\sigma_B}{8fg} \left(\bar{u}_{3,q+1} \right. \\
\left. \quad \dots\, -\bar{u}_{3,q-1} - \bar{u}_{1,q+1} + \bar{u}_{1,q-1}\right) \\[1em]
\vdots \\
\vdots \\[1em]
\frac{h\sigma_B^2}{4g^2} u_{P-2,q+1}^n + \left(1 - \frac{h\sigma_B^2}{2g^2}\right) u_{P-2,q}^n + \frac{h\sigma_B^2}{4g^2} u_{P-2,q-1}^n + \frac{h\rho\sigma_A\sigma_B}{8fg} \left(\bar{u}_{P-1,q+1} \right. \\
\left. \quad \dots\, -\bar{u}_{P-1,q-1} - \bar{u}_{P-3,q+1} + \bar{u}_{P-3,q-1}\right) \\[1em]
\frac{h\sigma_B^2}{4g^2} u_{P-1,q+1}^n + \left(1 - \frac{h\sigma_B^2}{2g^2}\right) u_{P-1,q}^n + \frac{h\sigma_B^2}{4g^2} u_{P-1,q-1}^n + \frac{h\rho\sigma_A\sigma_B}{8fg} \left(\bar{u}_{P,q+1}\bar{u} \right. \\
\left. \quad \dots\, -\bar{u}_{P,q-1} - \bar{u}_{P-2,q+1} + \bar{u}_{P-2,q-1}\right) + \frac{h\sigma_A^2}{4f^2} \tilde{u}_{P,q}
\end{bmatrix}
$$

Everything in b is known.

For the second ADI step from the form $Cy = d$ we have for $p = 1 : P - 1$:

$$
C =
\begin{bmatrix}
\left(1 + \frac{h\sigma_B^2}{2g^2}\right) & \frac{-h\sigma_B^2}{4g^2} & 0 & \cdots & \cdots & 0 \\
\frac{-h\sigma_B^2}{4g^2} & \left(1 + \frac{h\sigma_B^2}{2g^2}\right) & \frac{-h\sigma_B^2}{4g^2} & 0 & \cdots & 0 \\
0 & \ddots & \ddots & \ddots & 0 & \vdots \\
\vdots & 0 & \ddots & \ddots & \ddots & 0 \\
\vdots & \vdots & 0 & \frac{-h\sigma_B^2}{4g^2} & \left(1 + \frac{h\sigma_B^2}{2g^2}\right) & \frac{-h\sigma_B^2}{4g^2} \\
0 & 0 & \cdots & 0 & \frac{-h\sigma_B^2}{4g^2} & \left(1 + \frac{h\sigma_B^2}{2g^2}\right)
\end{bmatrix}
$$

$$
y =
\begin{bmatrix}
u_{p,1}^{n+1} \\
u_{p,2}^{n+1} \\
\vdots \\
\vdots \\
u_{p,Q-2}^{n+1} \\
u_{p,Q-1}^{n+1}
\end{bmatrix}
$$

$d =$

$$
\begin{bmatrix}
\frac{h\sigma_A^2}{4f^2}\tilde{u}_{p+1,1} + \left(1 - \frac{h\sigma_A^2}{2f^2}\right)\tilde{u}_{p,1} + \frac{h\sigma_A^2}{4f^2}\tilde{u}_{p-1,1} + \frac{h\rho\sigma_A\sigma_B}{8fg}\left(\bar{u}_{p+1,2}\right. \\
\left. \ldots - \bar{u}_{p+1,0} - \bar{u}_{p-1,3} + \bar{u}_{p-1,0}\right) + \frac{h\sigma_B^2}{4g^2}u_{p,0}^{n+1} \\[2em]
\frac{h\sigma_A^2}{4f^2}\tilde{u}_{p+1,2} + \left(1 - \frac{h\sigma_A^2}{2f^2}\right)\tilde{u}_{p,2} + \frac{h\sigma_A^2}{4f^2}\tilde{u}_{p-1,2} + \frac{h\rho\sigma_A\sigma_B}{8fg}\left(\bar{u}_{p+1,3}\right. \\
\left. \ldots - \bar{u}_{p+1,1} - \bar{u}_{p-1,3} + \bar{u}_{p-1,1}\right) \\[2em]
\vdots \\
\vdots \\[1em]
\frac{h\sigma_A^2}{4f^2}\tilde{u}_{p+1,Q-2} + \left(1 - \frac{h\sigma_A^2}{2f^2}\right)\tilde{u}_{p,Q-2} + \frac{h\sigma_A^2}{4f^2}\tilde{u}_{p-1,Q-2} + \frac{h\rho\sigma_A\sigma_B}{8fg}\left(\bar{u}_{p+1,Q-1}\right. \\
\left. \ldots - \bar{u}_{p+1,Q-3} - \bar{u}_{p-1,Q-1} + \bar{u}_{p-1,Q-3}\right) \\[2em]
\frac{h\sigma_A^2}{4f^2}\tilde{u}_{p+1,Q-1} + \left(1 - \frac{h\sigma_A^2}{2f^2}\right)\tilde{u}_{p,Q-1} + \frac{h\sigma_A^2}{4f^2}\tilde{u}_{p-1,Q-1} + \frac{h\rho\sigma_A\sigma_B}{8fg}\left(\bar{u}_{p+1,Q}\right. \\
\left. \ldots - \bar{u}_{p+1,Q-2} - \bar{u}_{p-1,Q} + \bar{u}_{p-1,Q-2}\right) + \frac{h\sigma_B^2}{4g^2}u_{p,Q}^{n+1}
\end{bmatrix}
$$

Everything in d is known. When solving these matrix equations it is important to note the sparsity of the square matrices, making use of this can improve the computational complexity from $O(N \times \min(P, Q) \times \max(P, Q)^3)$ to $O(N \times P \times Q)$, a significant improvement for fine grids or large strike prices. See the appendix for a detailed description of a linear complexity tridiagonal matrix equation solver, the Thomas Algorithm[6,9].

4.1.4   ADI Deltas

For the ADI solution we find the values for Δ_A and Δ_B numerically with second order approximations and linear interpolations between points on the discretized grid as needed for hedging. For points on the pricing grid we have:

$$
\begin{aligned}
\Delta_A(a_i, b_i, \tau) &= \Delta_A(p, q, n) = \frac{u_{p+1,q}^n - u_{p-1,q}^n}{f^2} \\
\Delta_B(a_i, b_i, \tau) &= \Delta_B(p, q, n) = \frac{u_{p,q+1}^n - u_{p,q-1}^n}{g^2}
\end{aligned}
\tag{4.17a}
$$

and for values not on the discretized grid two-dimensional linear interpolation is used. If we want the slope at a point (A, B) where $A \in (a_1, a_2)$ and $B \in (b_1, b_2)$ and we know the values of Δ_A and Δ_B at (a_1, b_1), (a_1, b_2), (a_2, b_1), and (a_2, b_2) then

$$
\begin{aligned}
\Delta_A(A, b_1, \tau) &= \Delta_A(a_1, b_1, \tau)\frac{a_2 - A}{a_2 - a_1} + \Delta_A(a_2, b_1, \tau)\frac{A - a_1}{a_2 - a_1} \\
\Delta_A(A, b_2, \tau) &= \Delta_A(a_1, b_2, \tau)\frac{a_2 - A}{a_2 - a_1} + \Delta_A(a_2, b_2, \tau)\frac{A - a_1}{a_2 - a_1}
\end{aligned}
\tag{4.17b}
$$

and

$$
\Delta_A(A, B, \tau) = \Delta_A(A, b_1, \tau)\frac{b_2 - B}{b_2 - b_1} + \Delta_A(A, b_2, \tau)\frac{B - b_1}{b_2 - b_1}
\tag{4.17c}
$$

Figure 4.1: Visual Representation of Two-Dimensional Interpolation
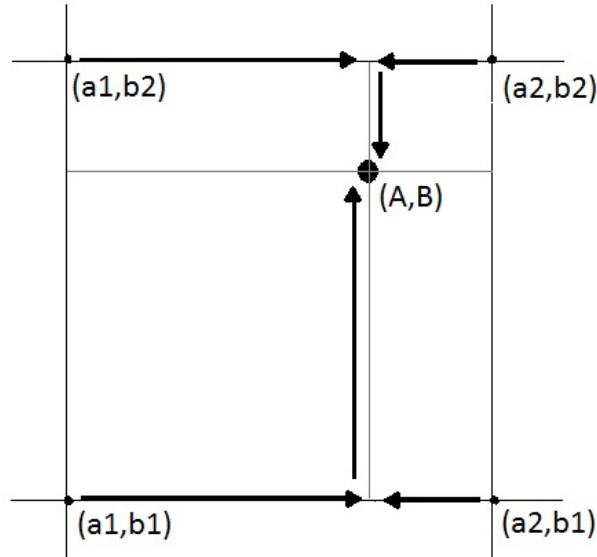
See figure 4.1 for a visual representation of this process. First we did the horizontal interpolation then the vertical. The interpolation is similar for Δ_B. This also works for hedging to find prices that are not on the grid.

4.2 Analytic Approximation - "gamma" Solution

Now we will look at the solution that was developed in section 3.3 where we use a one-dimensional approximation to the true two-dimensional PDE while still tracking the proportion of the index that is made up by each asset. That is $A = \gamma I$ and $B = (1 - \gamma)I$; since A, B and I all vary with time γ also varies over time. This is the PDE, equation (3.9) restated:

$$V_t + \frac{1}{2}(\sigma^\star)^2 I^2 V_{II} + rIV_I - rV = 0$$

with

$$(\sigma^\star)^2 = \sigma_A^2 \gamma^2 + 2\rho\sigma_A\sigma_B\gamma(1 - \gamma) + \sigma_B^2(1 - \gamma)^2$$

With $V(A, B, T) = (K - A - B)^+$ as the terminal condition in the case of a put. The solution to this is

$$V = Ke^{-r\tau}\Phi(-d_2) - I\Phi(-d_1) \tag{4.18}$$

Where Φ is the standard normal CDF and

$$d_1 = \frac{\ln(\frac{I}{K}) + (r + \frac{1}{2}(\sigma^\star)^2)\tau}{\sigma^\star\sqrt{\tau}}$$

$$d_1 = d_1 - \sigma^\star \sqrt{\tau}$$

The values of Δ_I can be found analytically,

$$\Delta_I = -\Phi(-d_1) \tag{4.19}$$

but we are more interested in Δ_A and Δ_B which need to be found numerically as they were for the ADI method:

$$\Delta_A(p, q, n) = \frac{u^n_{p+1,q} - u^n_{p-1,q}}{f^2} \tag{4.20a}$$

$$\Delta_B(p, q, n) = \frac{u^n_{p,q+1} - u^n_{p,q-1}}{g^2} \tag{4.20b}$$

But here $\Delta_A(p, q, n)$ should equal $\Delta_B(p, q, n)$ because of our assumptions that created this model.

4.3   Monte Carlo Solution

The third method that we will use will serve as a benchmark and is a simple Monte Carlo method.

$$V(A_t, B_t, t) = e^{-r(T-t)} \mathbf{E}[(K - A_T - B_T)^+ | \mathcal{F}_t] \tag{4.21a}$$

Where

$$A_T = A_t e^{(r - \frac{\sigma_A^2}{2})(T-t) + \sigma_A \sqrt{T-t}x} \tag{4.21b}$$

$$B_T = B_t e^{(r - \frac{\sigma_B^2}{2})(T-t) + \sigma_B \sqrt{T-t}y} \tag{4.21c}$$

and

$$(x, y) \sim \mathcal{N}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix} \right) \tag{4.21d}$$

The Monte Carlo simulation at any, or every, point is done by taking a sample of size M from the standard bivariate normal distribution and

$$V(A_t, B_t, t) = e^{-r(T-t)} \frac{1}{M} \sum_{i=1}^{M} (K - A_T(i) - B_T(i))^+ \tag{4.22a}$$

Where

$$A_T(i) = A_t e^{(r - \frac{\sigma_A^2}{2})(T-t) + \sigma_A \sqrt{T-t}x_i} \tag{4.22b}$$

$$B_T(i) = B_t e^{(r-\frac{\sigma_B^2}{2})(T-t)+\sigma_B \sqrt{T-t}y_i} \tag{4.22c}$$

For the deltas we follow a similar process but first we must set up the expectation for which we will perform a Monte Carlo simulation. Start with the definition of delta:

$$
\begin{aligned}
\Delta_A(t) &= \frac{\partial}{\partial A_t} V(t) \\
&= \frac{\partial}{\partial A_t} \left(e^{-r(T-t)} \mathbf{E}[(K - A_T - B_T)^+ | \mathcal{F}_t] \right) \\
&= e^{-r(T-t)} \mathbf{E}[\frac{\partial}{\partial A_t} ((K - A_T - B_T)^+) | \mathcal{F}_t] \\
&= e^{-r(T-t)} \mathbf{E}[-\frac{\partial A_T}{\partial A_t} \mathbf{1}_{(A_T+B_T<K)} | \mathcal{F}_t]
\end{aligned}
\tag{4.23a}
$$

Where $\mathbf{1}_{(A_T+B_T<K)}$ is an indicator function. Similarly,

$$
\begin{aligned}
\Delta_B(t) &= \frac{\partial}{\partial B_t} V(t) \\
&= \frac{\partial}{\partial B_t} \left(e^{-r(T-t)} \mathbf{E}[(K - A_T - B_T)^+ | \mathcal{F}_t] \right) \\
&= e^{-r(T-t)} \mathbf{E}[\frac{\partial}{\partial B_t} ((K - A_T - B_T)^+) | \mathcal{F}_t] \\
&= e^{-r(T-t)} \mathbf{E}[-\frac{\partial B_T}{\partial B_t} \mathbf{1}_{(A_T+B_T<K)} | \mathcal{F}_t]
\end{aligned}
\tag{4.23b}
$$

We can also easily find that:

$$\frac{\partial A_T}{\partial A_t} = e^{(r-\frac{\sigma_A^2}{2})(T-t)+\sigma_A \sqrt{T-t}x} \tag{4.23c}$$

$$\frac{\partial B_T}{\partial B_t} = e^{(r-\frac{\sigma_B^2}{2})(T-t)+\sigma_B \sqrt{T-t}y} \tag{4.23d}$$

So:

$$\Delta_A(A_t, B_t, t) = e^{-r(T-t)} \frac{1}{M} \sum_{i=1}^{M} -\frac{\partial A_T(i)}{\partial A_t} \mathbf{1}_{(A_T(i)+B_T(i)<K)} \tag{4.23e}$$

$$\Delta_B(A_t, B_t, t) = e^{-r(T-t)} \frac{1}{M} \sum_{i=1}^{M} -\frac{\partial B_T(i)}{\partial B_t} \mathbf{1}_{(A_T(i)+B_T(i)<K)} \tag{4.23f}$$

Where

$$A_T(i) = A_t e^{(r-\frac{\sigma_A^2}{2})(T-t)+\sigma_A \sqrt{T-t}x_i} \tag{4.23g}$$

$$B_T(i) = B_t e^{(r-\frac{\sigma_B^2}{2})(T-t)+\sigma_B \sqrt{T-t}y_i} \tag{4.23h}$$

We have developed a finite difference scheme solution and a simulation approach to solve equation (3.8) for V and to determine the delta values so that we can hedge the option. Now that we have two methods to solve the two-dimensional PDE, we need to compare their pricing capabilities. Next we will compare these methods to the analytic solutions from the one-dimensional "gamma" approximation that we developed in section 3.3, then later we will look at their hedging profits.

Chapter 5

Comparing Pricing Methods

So far we have derived three distinct PDEs that define the pricing behaviour of an index option with a dominant constituent asset. The first of these, equation (3.5), is truely one-dimensional, it only depends on the index as a whole. The second PDE, equation (3.8), is truely two-dimensional, it varies with the dominant asset and the rest of the index, as well as their derivatives. The third PDE, equation (3.9), is a one-dimensional approximation to the two-dimensional PDE; the volatility varies with changes in the level of dominance, though the derivatives are only with respect to the index as a whole. We have developed two numerical methods to solve the two-dimensional PDE: the ADI method, which is a finite difference scheme, and a Monte Carlo method. We will now compare the pricing surfaces and the delta, the first spatial derivative, surfaces that our three two-dimensional methods, ADI, Monte Carlo and the "gamma" approximation produce. Later we will include the solutions from the one-dimensional model in our hedging comparison.

All the figures in this chapter are produced with the parameters in table 5.1. The parameter r describes the market, K and T describe the option, σ_A, σ_B and ρ describe the underlying assets, and f, g and h describe the grid size used. These parameters were chosen to make the attributes of the models visible while not being too unrealistic.

Table 5.1: Parameters Used to Compare Pricing and Delta Surfaces

r	=	10%
K	=	\$5
T	=	1 year
σ_A	=	0.25
σ_B	=	0.10
ρ	=	0.3
f	=	5¢
g	=	5¢
h	=	1 business day = 1/252 years

5.1 ADI versus Monte Carlo

First we will compare the ADI solution for the option prices and deltas to the Monte Carol simulations. Figure 5.1 is the surface for the price of the option at inception using the ADI method. One of the issues with this solution can be seen here, on the $B = 0$ boundary near the strike price of the 3D plot, it can be seen that along the strike the option is undervalued. Compared to the boundary condition, which we know to be correct, the price should not drop that drastically. In contrast to the $A = 0$ boundary we can see that this issue is exaggerated by higher values of σ. In the 2D plots this can also be seen at $A = 0$ and $B = 0$. We can also see in the cross section line plots that the option is over valued by ADI when the option is in the money by the $A, B = 2.5$ line curving back down at the vertical axis.

Figure 5.2 shows the option pride from Monte Carlo; this is our benchmark option price surface. To better see the over and under valuing of the option by the ADI method we look at the difference between the ADI and Monte Carlo option prices in figure 5.3. Here we can see an exaggerated detailing of what we saw before, that there is underpricing in the ADI solution at the money, which is inflated by the higher value of σ_A and that the ADI solution over prices the option when it is in the money.

Next we will look the delta surfaces. First, figures 5.4 and 5.5, are the ADI deltas, followed by the Monte Carlo deltas in figures 5.6 and 5.7.The issues that were visible in the ADI pricing surface are even more evident in the surfaces for the ADI deltas. In the plot of the ADI Δ_A values, figure 5.4, the error at the $A = 0$ boundary is emphasized, we can see that the pricing surface increases away from the boundary before taking on the correct, though overpriced, shape. It is hard to see major issues with the $B = 0$ boundary here because we are taking the derivative parallel to it.In the plot of the ADI Δ_B values, figure 5.5, the error at the $B = 0$ boundary is emphasized, we can see that the pricing surface increases away from the boundary again in the money and under priced near the strike, before taking on the correct, though overpriced, shape. It is hard to see major issues with the $A = 0$ boundary here because we are taking the derivative parallel to it. Figures 5.6 and 5.7 show what we expect the deltas to look like, from the Monte Carlo simulation.

To see the error along the strike price we look at the difference between the solutions' deltas in figures 5.14 and 5.15. in figure 5.14 we again see the error on the $A = 0$ boundary but we also see error along the strike price. The price changes too quickly at the strike, the slope is too low, then too high, and then slightly low again as it passes the strike price. We will see similar results from Δ_B in firure 5.15.

5.2 ADI versus "gamma"

Next we will look compare the ADI solution for the option prices and deltas to the "gamma" analytic approximations. We have already seen the surface for the price of the option and its deltas at inception using the ADI method, so the pricing surface at inception by Monte Carlo simulation is shown in figure 5.10, and the deltas by Monte Carlo simulation are in figures 5.11 and 5.12. These surfaces all look as expected but lack the curvature implied from the true solution to the problem.

Next we compare the ADI pricing surface to the one from the "gamma" analytic approx-

Figure 5.1: Pricing Surface from ADI Numerical Solution, Parameters as in Table 5.1
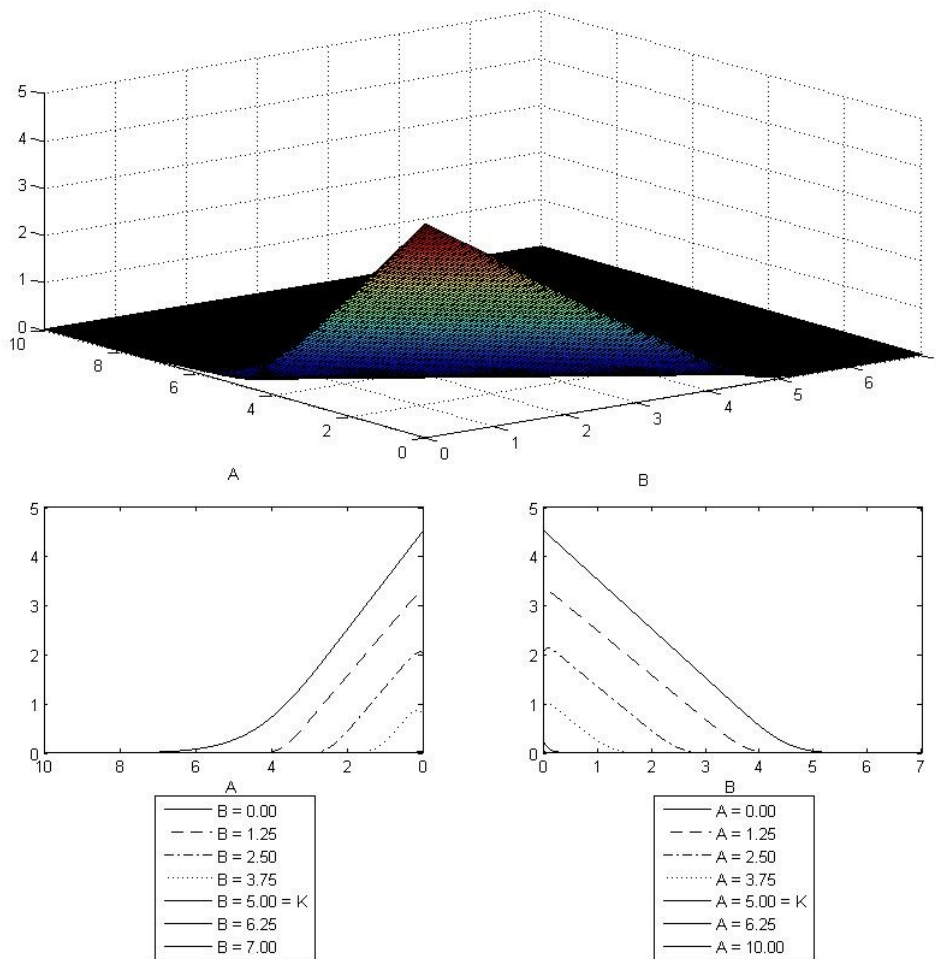
Figure 5.2: Pricing Surface from Monte Carlo , Parameters as in Table 5.1
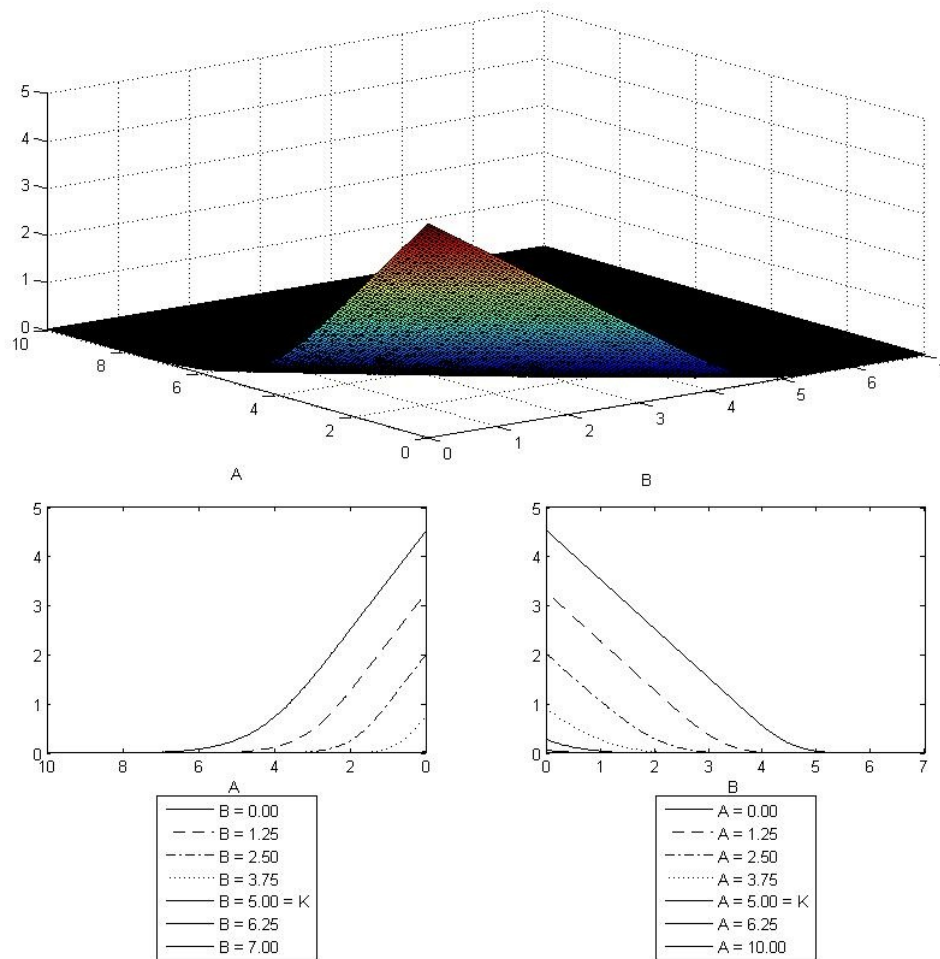
Figure 5.3: Difference between ADI and Monte Carlo Pricing Surfaces, Parameters as in Table 5.1
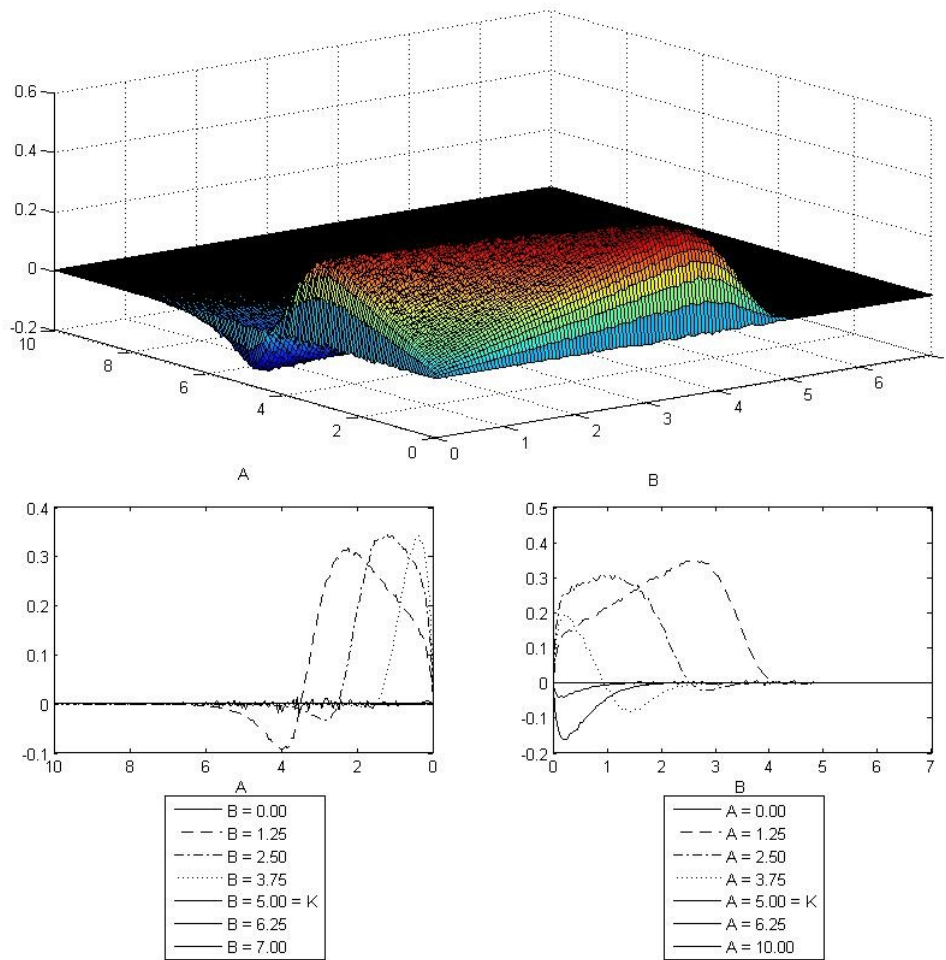
Figure 5.4: Δ_A Surface from ADI Numerical Solution, Parameters as in Table 5.1
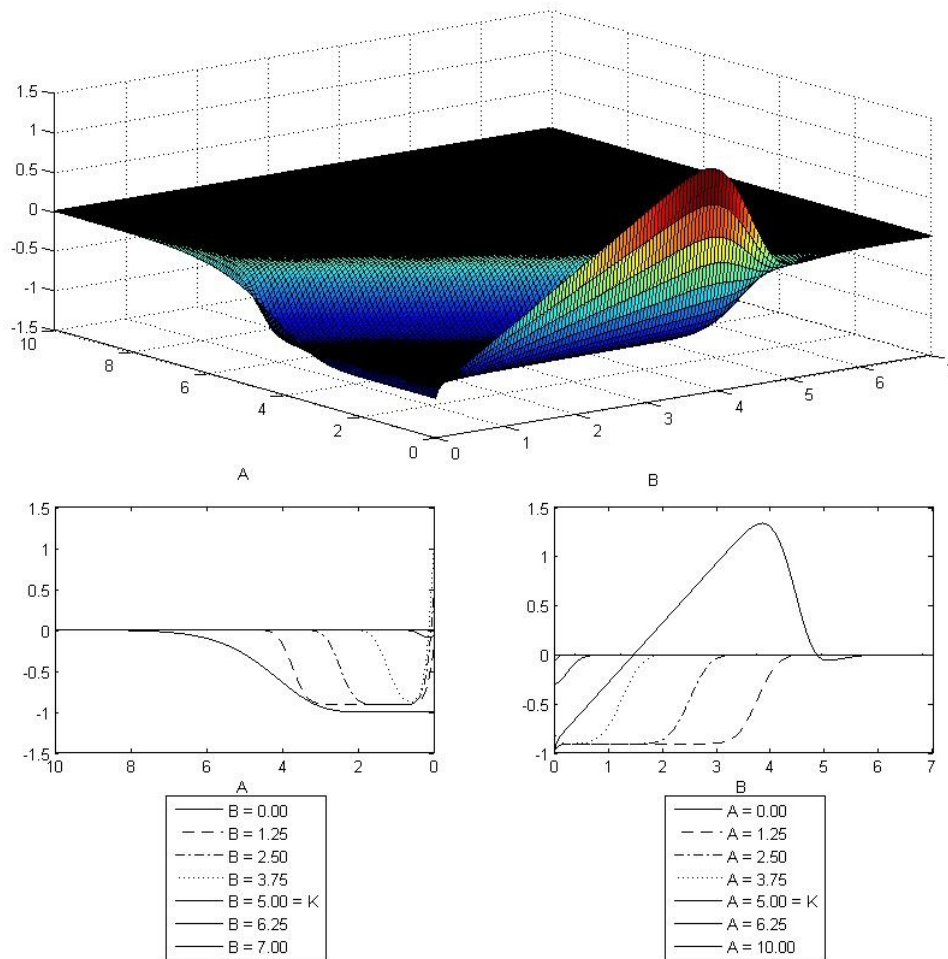
Figure 5.5: Δ_B Surface from ADI Numerical Solution, Parameters as in Table 5.1
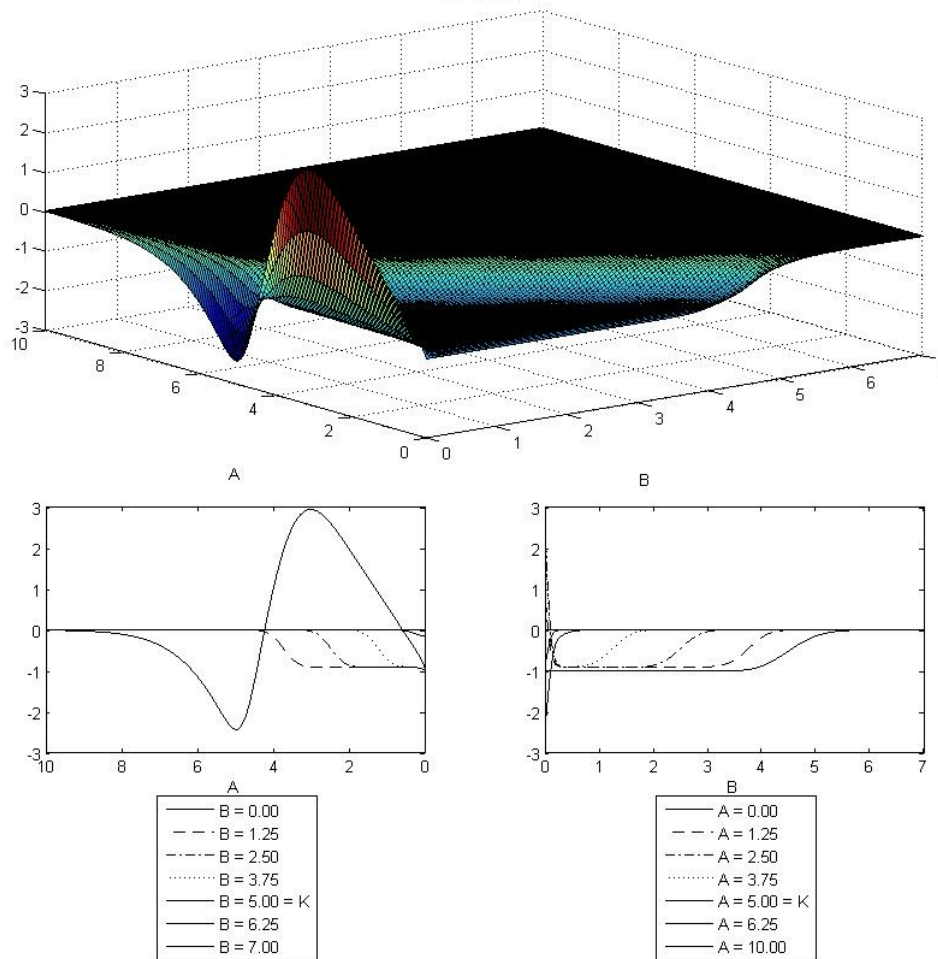
Figure 5.6: Δ_A Surface from Monte Carlo, Parameters as in Table 5.1
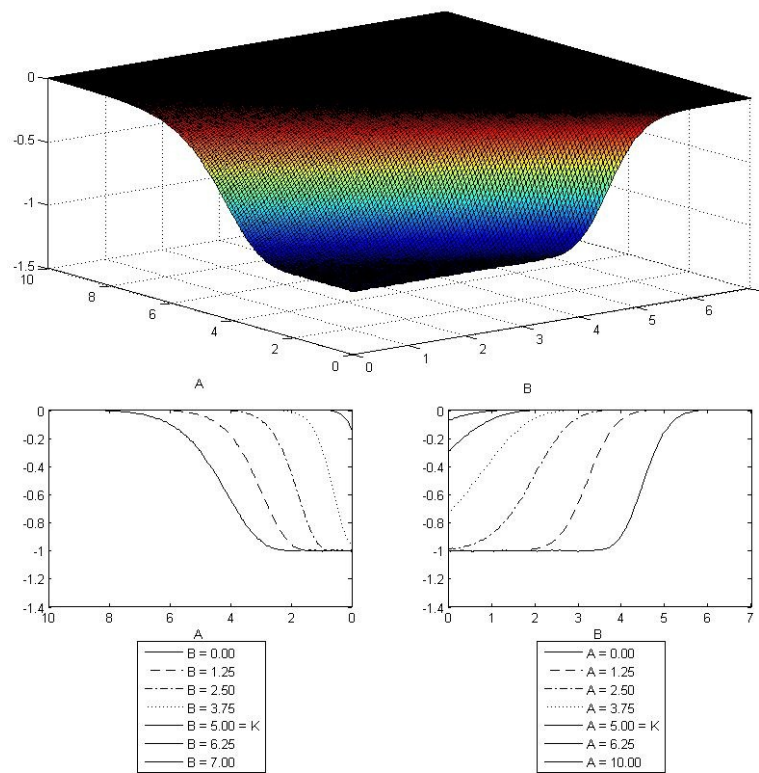
Figure 5.7: Δ_B Surface from Monte Carlo , Parameters as in Table 5.1
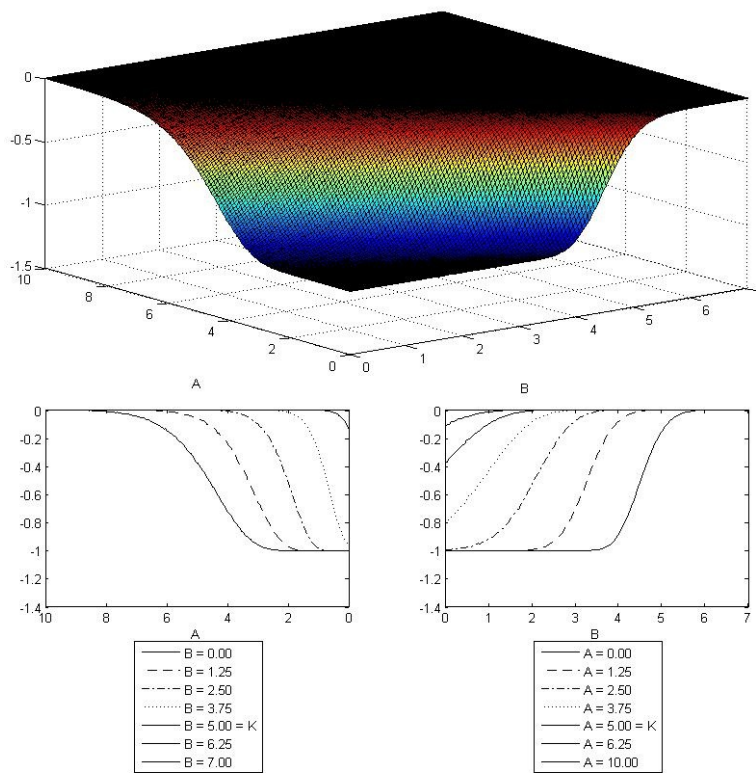
Figure 5.8: Difference between ADI and Monte Carlo Δ_A Surfaces, Parameters as in Table 5.1
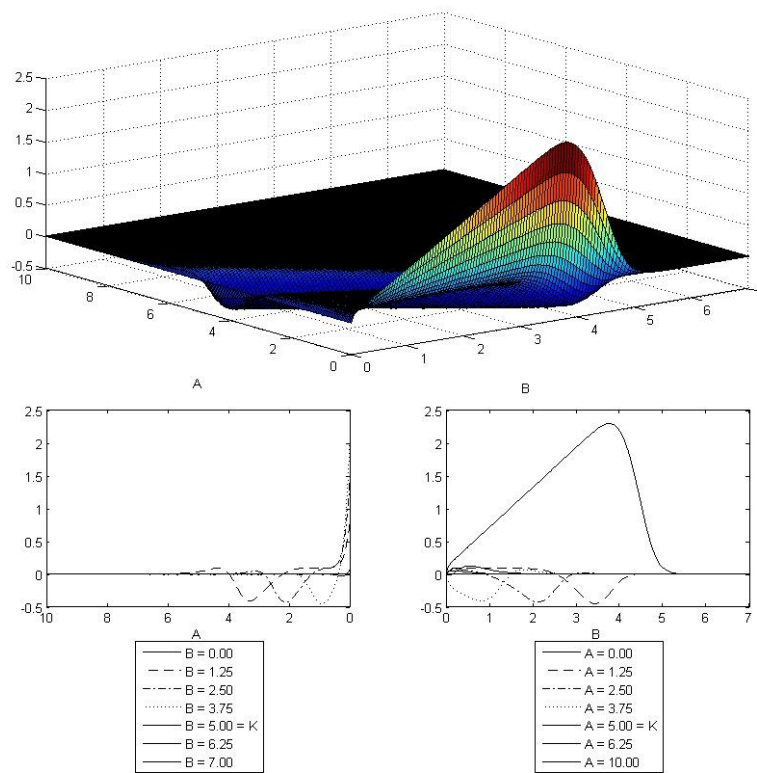
Figure 5.9: Difference between ADI and Monte Carlo Δ_B Surfaces, Parameters as in Table 5.1
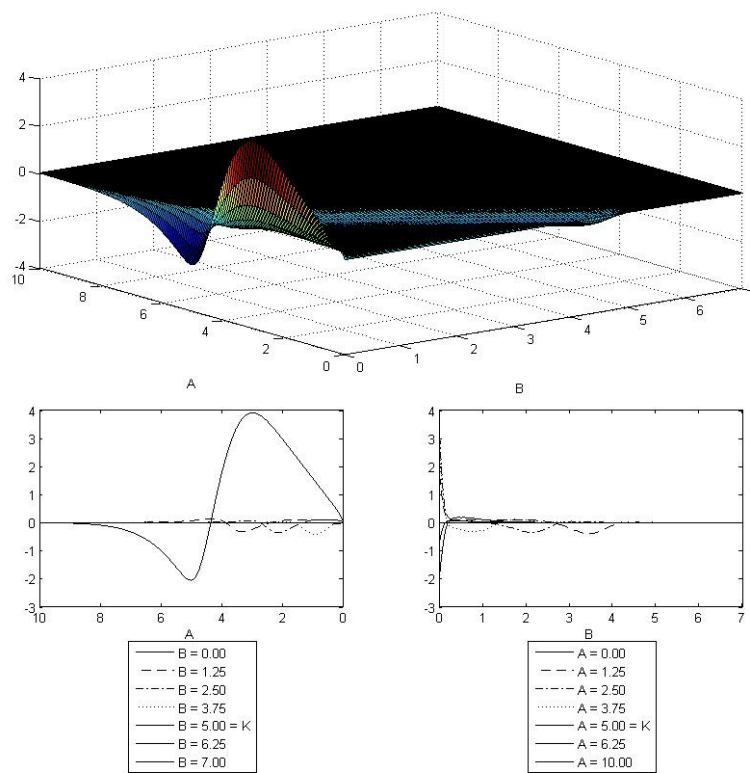
Figure 5.10: Pricing Surface from "gamma" Approximation, Parameters as in Table 5.1
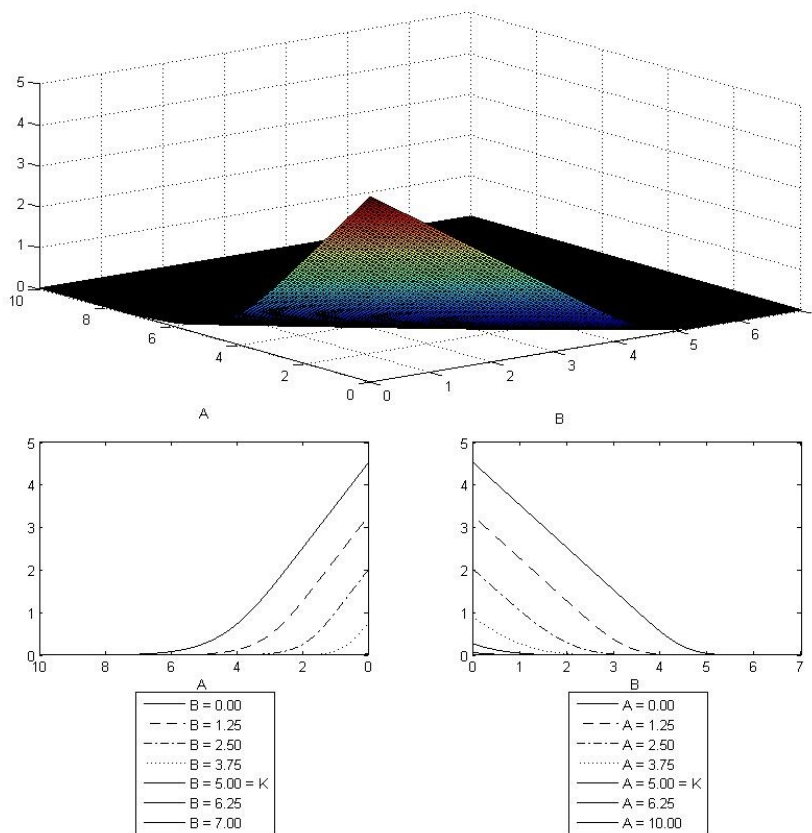
Figure 5.11: Δ_A Surface from "gamma" Approximation, Parameters as in Table 5.1
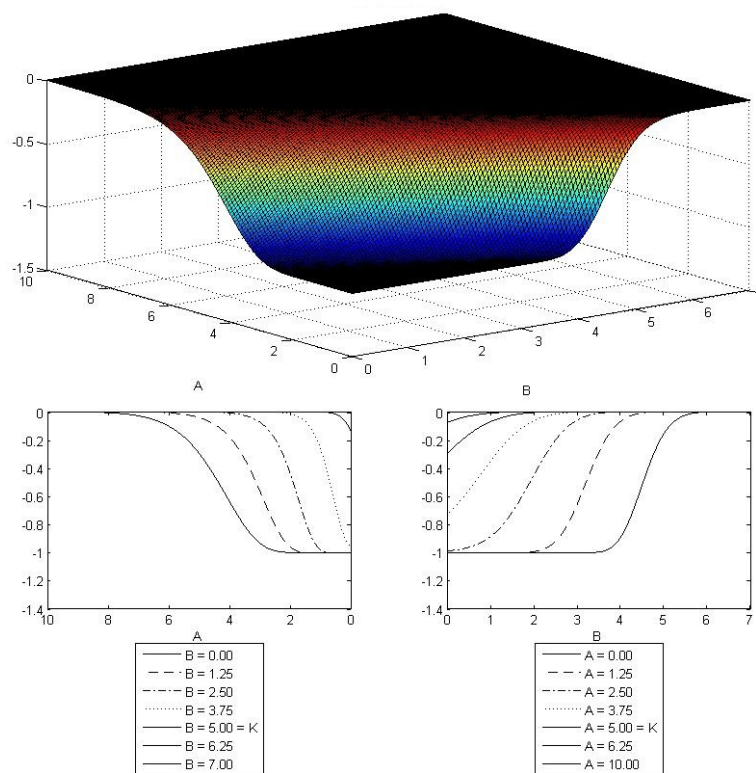
Figure 5.12: Δ_B Surface from "gamma" Approximation, Parameters as in Table 5.1
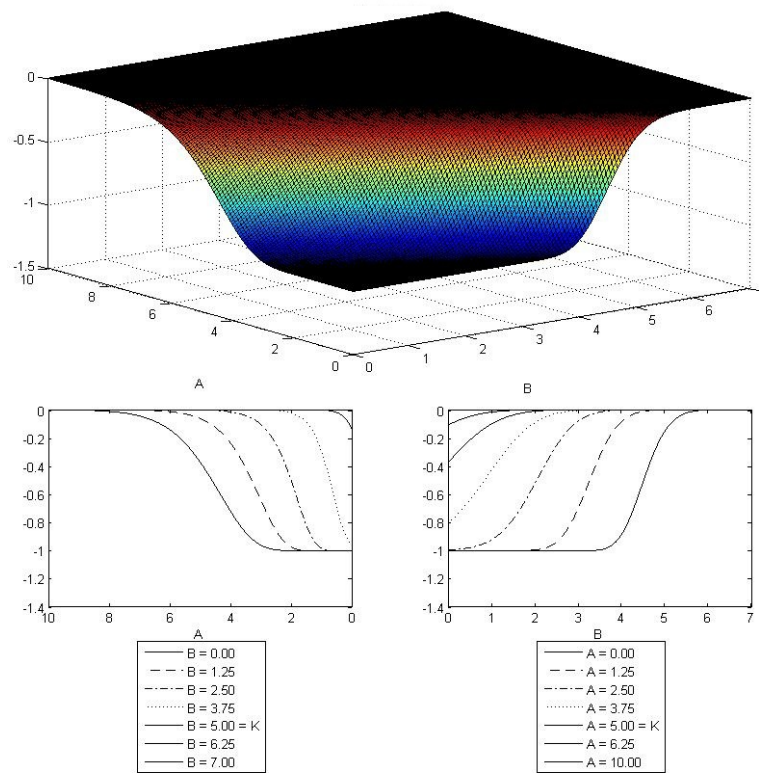
Figure 5.13: Difference between ADI and "gamma" Pricing Surfaces, Parameters as in Table 5.1
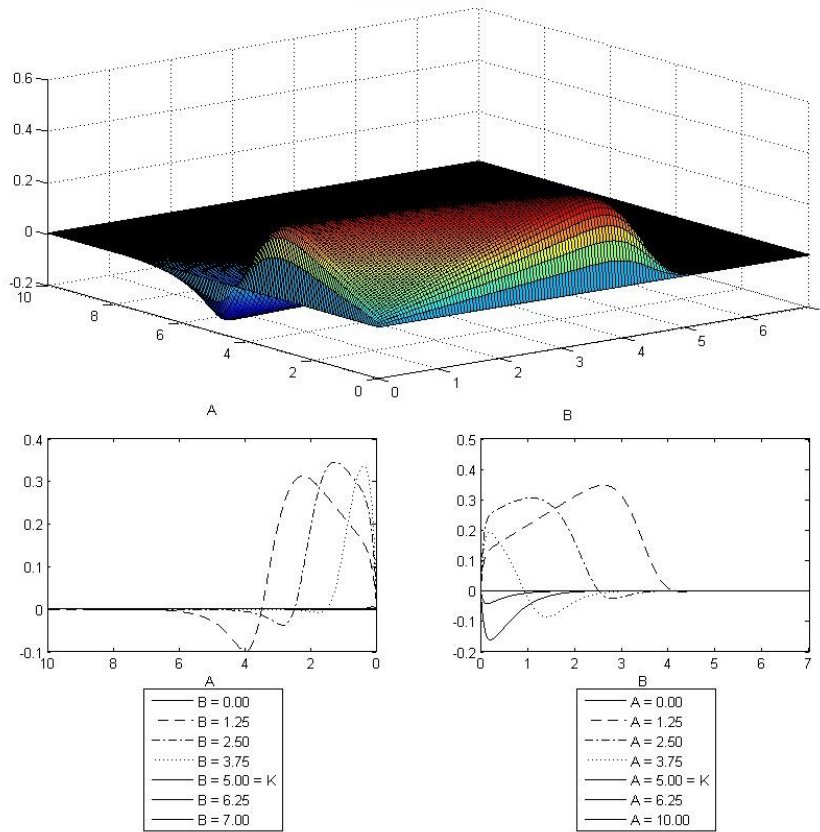
imations in figure 5.13. This looks the same as the comparison of ADI to the Monte Carlo simulation. Now we compare the deltas in figures 5.14 and 5.15. These appear to be the same as the comparison to the Monte Carlo simulation.

5.3 "gamma" versus Monte Carlo

Finally we compare the "gamma" analytic approximation to the Monte Carlo simulation to see how much curvature is lost in our approximation. First the difference between the pricing surfaces is shown in figure 5.16. From this plot it is very obvious that the Monte Carlo simulation's pricing surface is not smooth and this makes it very hard to see the trends in the difference. However, we need to do so to see the shortfalls of the "gamma" analytic approximation; in the 2D plots it can be seen that there are some definite trends in the difference but at very small magnitudes.

Figure 5.14: Difference between ADI and "gamma" Δ_A Surfaces, Parameters as in Table 5.1
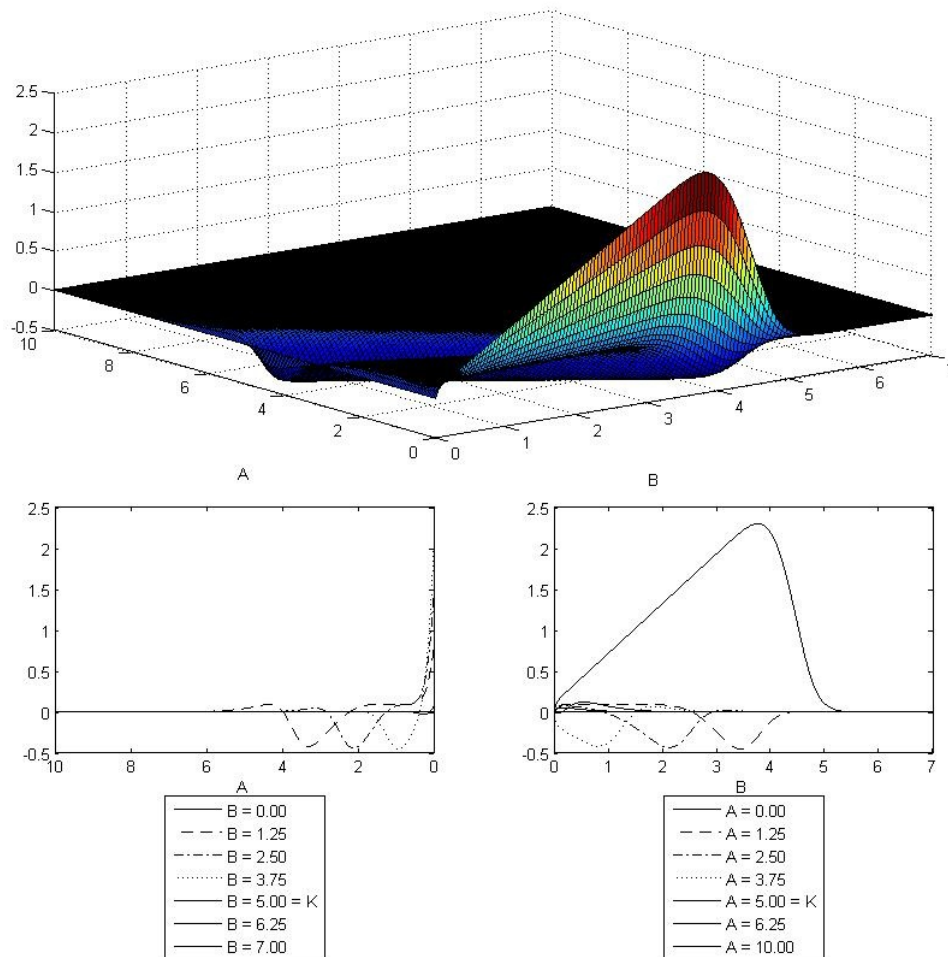
Figure 5.15: Difference between ADI and "gamma" Δ_B Surfaces, Parameters as in Table 5.1
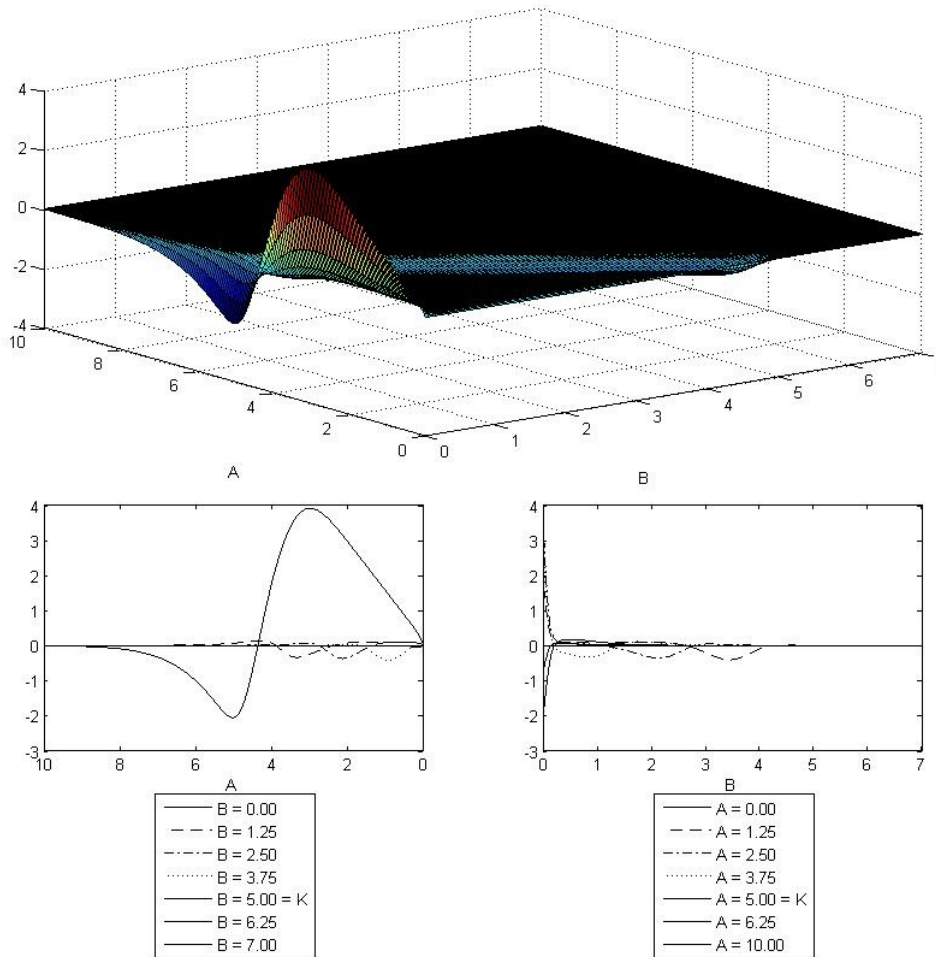
After thorough examination of the 3D plot from various angles we can see that the prices from the "gamma" analytic approximation are high when the option is just out of the money and low when the option is just in the money away from the boundaries. This trend is strongest in the $\gamma = \frac{1}{2}$ area and persists for more values to the side where the underlying asset with higher volatility varies more rapidly. For hedging this is the most important area at which to have accurate prices.

To get more insight into this issue we now look at the comparison of the deltas in figures 5.17 and 5.18. In these plots we see that the Monte Carlo simulation's Δ_A is also non-smooth, but we can also start to see some trends.

These are again difficult to read from only one angle but after much examination we can see that at the money Δ_A from the "gamma" analytic approximation is higher than it should be, or the pricing surface is flatter than it should be in the A direction. Since at either extreme there is no difference in the deltas to make up for this flatness on either side of the strike price Δ_A from the "gamma" analytic approximation is high, it is more drastic when the option is in the money than out of the money but it is present on both sides. As with the over and under pricing these trends are more persistent toward the boundary where the more volatile underlying asset changes rapidly.

The trends present for Δ_B are similar to those in Δ_A but more persistent and the differences come much closer to, and stretch along, the $B = 0$ boundary than they did for Δ_A. All of this confirms that the pricing surface from the "gamma" analytic approximation is steeper and underpriced just in the money, flatter at the money, and steeper and overpriced just out of the money compared to our benchmark prices from the Monte Carlo simulation.

Even though figures 5.16 to 5.18 are very volatile by comparing relative difference plots it can be seen that the trends in the difference follow definite trends, and these trends are on a larger scale than the Monte Carlo errors. Figure 5.19 is a two-dimensional depiction of how the "gamma" pricing surface compares to the one from Monte Carlo simulation which is much easier to read.

Now we have a feel for the pricing and delta surfaces from our two-dimensional models. Next we will compare the hedging strategies from both the one and two-dimensional methods for some theoretical parameters and later we will perform the same exercise for some empirical data.

Figure 5.16: Difference between "gamma" and Monte Carlo Pricing Surfaces, Parameters as in Table 5.1
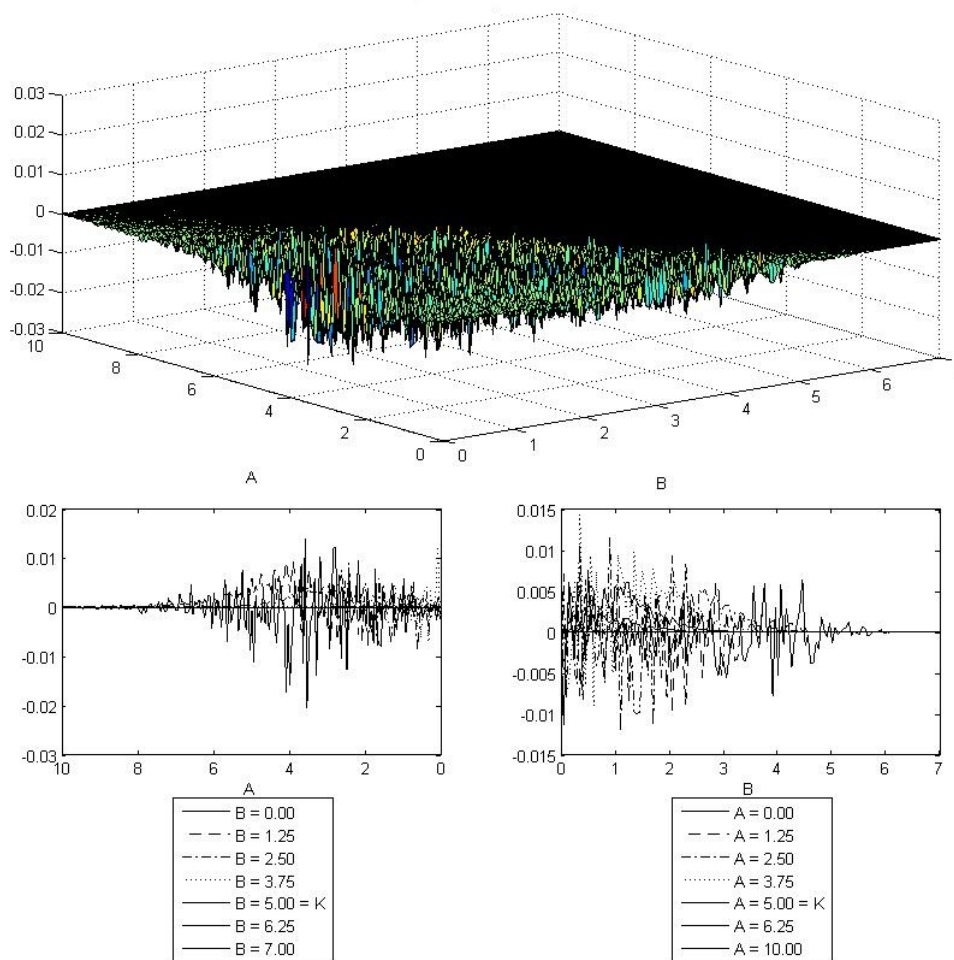
Figure 5.17: Difference between "gamma" and Monte Carlo Δ_A Surfaces, Parameters as in Table 5.1
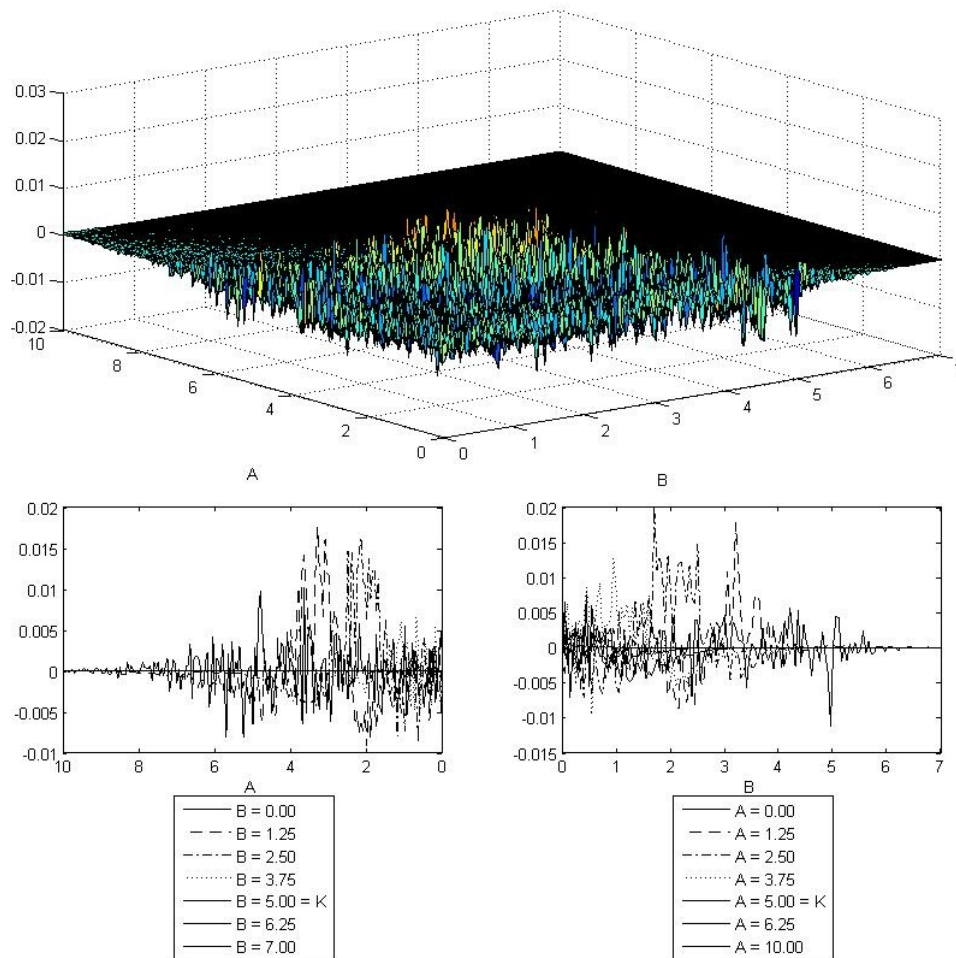
Figure 5.18: Difference between "gamma" and Monte Carlo Δ_B Surfaces, Parameters as in Table 5.1
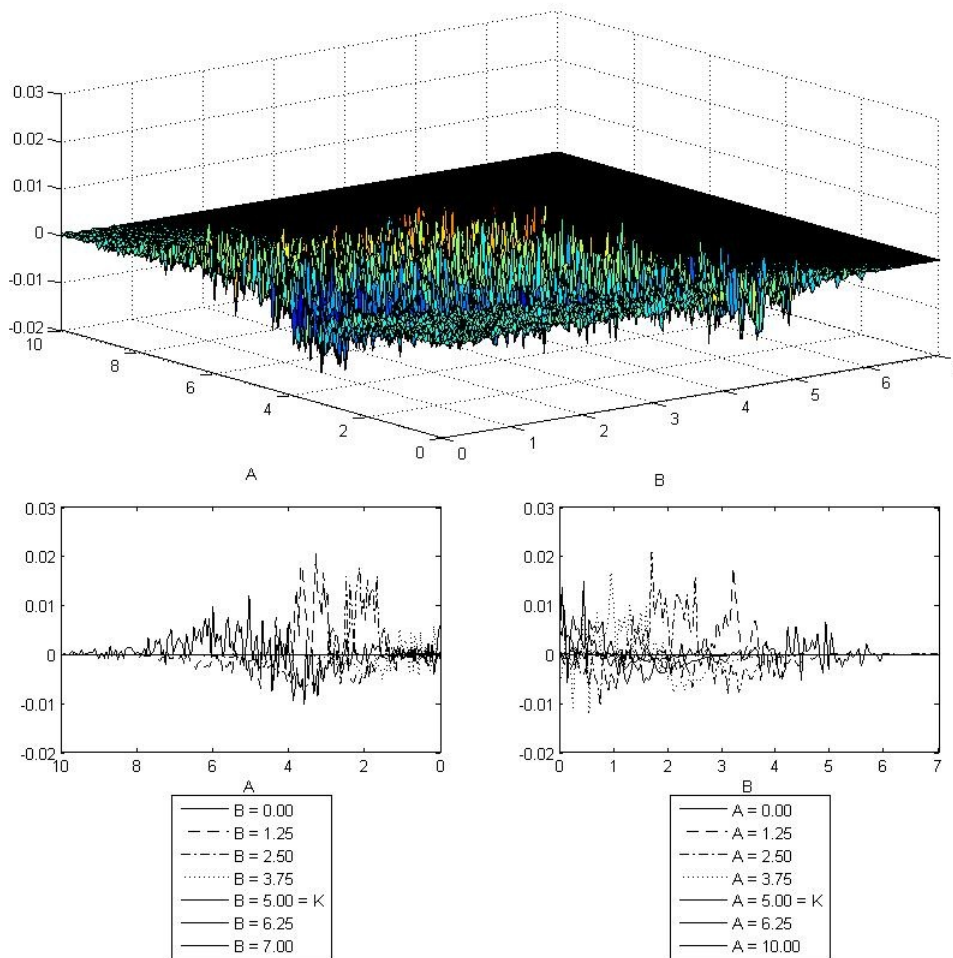
Figure 5.19: Comparison of "gamma" approximation and the Monte Carlo solution. This is most extreme for values of I that are closer to $\gamma = 0.5$ than $\gamma = 0$ or $\gamma = 1$
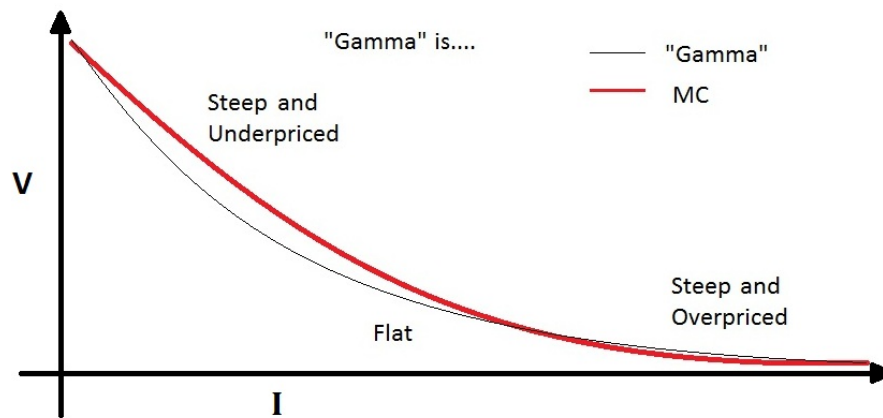
Chapter 6

Comparing Hedging Profits

Now that we have looked at the pricing and delta surfaces for our three two-dimensional methods we will compare the distribution of hedging profits for all of our models first through moments and then through histograms and QQ plots. The models we compare are

1. the truly one-dimensional Black-Scholes model that ignores that there is a dominant asset,

2. the "gamma" model hedging with only the index hedging ratio Δ_I,

3. the "gamma" model hedging with both Δ_A and Δ_B approximated by interpolation,

4. the ADI method solving the two-dimensional PDE, and

5. the Monte Carlo simulation approach to solving the two-dimensional PDE.

We examine the first two moments for both model prices and market prices. This is because, as we saw in the previous section, there are some issues with the option prices, as well as the deltas, in some areas of each model. We would like to complete both the theoretical hedging exercise as well as use market prices with modeled deltas to hedge without compounding any pricing errors. Since this is still a theoretical exercise, but we need market prices for each individual stock path, we will use the Monte Carlo simulated option prices because even though they are not smooth that does not affect hedging and they are the most accurate solution to the PDE that we have[1].

Since we would have the same results for the Monte Carlo simulation hedging for both market and model prices we will look at its hedging results here.

The hedging profit is the amount of money that would be realized upon liquidation of all positions at the maturity of the option. We have rebalanced our hedge daily, or 252 times per year. If we hedge the dominated index option with both individual underlying assets, the dominating stock and the rest of the index, the portfolio that we use is:

$$\Pi(t) = V(t) - \Delta_A(t)A(t) - \Delta_B(t)B(t) \tag{6.1}$$

---

[1]Remember, from equations (4.23), that we simulate our Monte Carlo Delta values separately from the option values; though this still produces non-smooth Delta values they are much more smooth that would be produced by calculating Deltas from the Monte Carlo simulated pricing surface.

Table 6.1: Parameters used for Hedging Comparisons

r	=	5%
K	=	$5
T	=	1 year
σ_A	=	0.25
σ_B	=	0.10
ρ	=	0.3
f	=	20¢
g	=	20¢
h	=	dt = 1 business day = 1/252 years

The cost to set up these positions is $-\Pi(t) = \Delta_A(t)A(t) + \Delta_B(t)B(t) - V(t)$, so this is the cash balance at inception. The cash balance at each step is the previous balance charged, or earned, the continuously compounded interest rate, r, plus the increase in the short position in both assets:

$$cash(t) = e^{rdt}cash(t-1) + (\Delta_A(t) - \Delta_A(t-1))A(t) + (\Delta_B(t) - \Delta_B(t-1))B(t) \qquad (6.2)$$

At the option's maturity, if the option is in the money, we exercise the option by selling the one of each asset that we own for the strike price and pay off our debt. The cash that is left over is our hedging profit. If the option is out of the money then we will not own any of the assets and the amount of cash is our hedging profit. Figures 6.1 and 6.2 track how all the variables evolve over time for a single sample pair of asset paths under the Monte Carlo simulation driven hedging for various cases of option payouts. The parameters used in all of the examples of this chapter are listed in table 6.1.

In figure 6.1 the option finished out of the money, but barely, and in figure 6.2 the option finished in the money, again only barely. We know from our previous analysis that the option prices are accurate and they appear to be here as well. It is hard to see that both Δ_A and Δ_B are plotted in the third plot on the left because they stay so close together. This follows from the reasoning that developed the "gamma" model, and close together but imperfect is what we expect to see. The cash amount, or bank account balance, is always negative when hedging a put as money is borrowed for both the option purchase and to buy stock. The cash value is a deterministic function of interest and the values plotted on the left side. The final plot is the portfolio's liquidation value. At times before maturity this value would be achieved by liquidating the whole portfolio and paying off all debts. At maturity either the option is in the money and we exercise and liquidate other positions, if the option is out of the money its price should be near zero, the result is the hedging strategy's profits.

The hedging profits from a vanilla call or put on a single underlying asset for a single time step driven by one GBM will have hedging profits that are distributed χ^2 with one degree of freedom translated to have a mean of zero at each time step[10]; the sum of all time steps, by the central limit theorem, will be normally distrubuted[2]. For one year of hedging with daily rebalancing the normal distribution is not quite acheived, but the hedging profits in that case are

Figure 6.1: Hedging Profit Evolution Sample Path, Out of the Money, Parameters as in Table 6.1
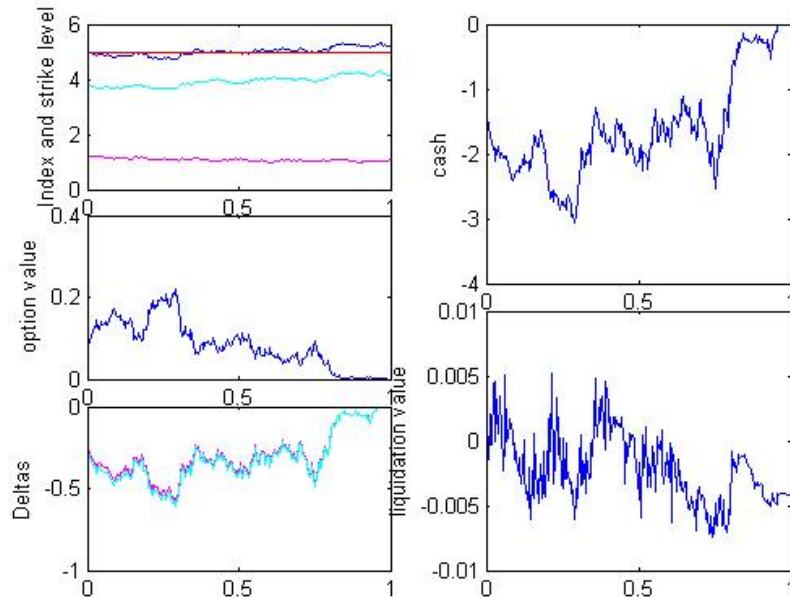
Figure 6.2: Hedging Profit Evolution Sample Path, In the Money, Parameters as in Table 6.1
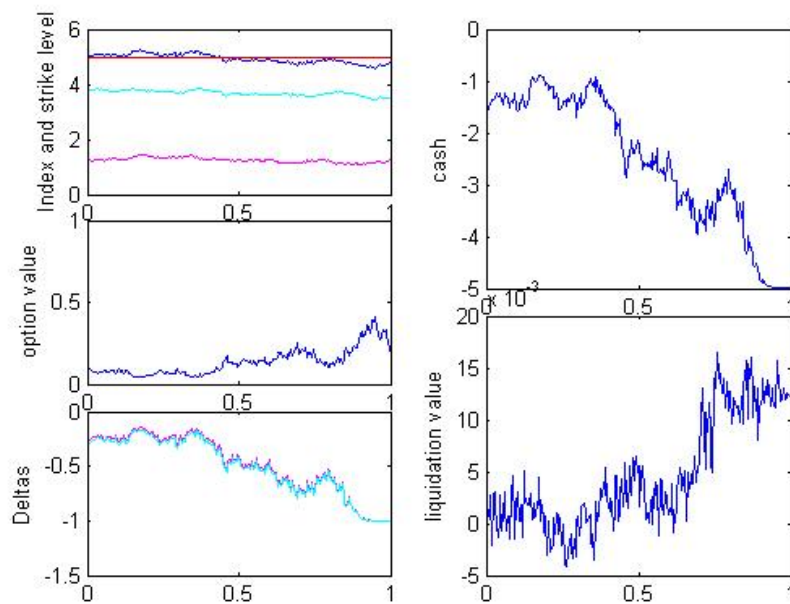
Figure 6.3: Histogram of Hedging Profits from Monte Carlo Solution, Parameters as in Table 6.1
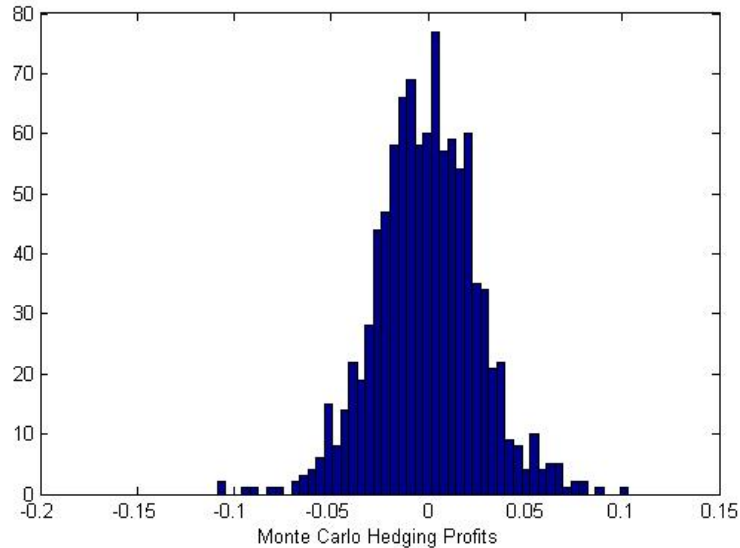
distributed $\chi^2(252)$ which appears to be normal at first glance. This is examined in Appendix E. Next we compare the hedging profits of $10,000$ sample paths to χ^2 with 252 degrees of freedom. From figures 6.3 and 6.4 we can see that the hedging profits appear to be distributed similar to $\chi^2(252)$; the QQ plot, figure 6.5 confirms this.

Figure 6.3 shows the histogram for the hedging profits of the Monte Carlo simulation. We can see that it is centred around zero and has very little skewness. Now let's look at the χ^2 distribution for 252 degrees of freedom in figure 6.4. This looks like it has the same shape as the previous figure but without the translation it does not have a mean of zero. Let's compare using the QQ plot in figure 6.5 instead.

From figure 6.5 we can see that χ^2 with 252 degrees of freedom is a fairly good fit for the hedging profits of an index option with an overweighted constituent stock using the Monte Carlo solution. The tails of the hedging profits are a little fatter than of the $\chi^2(252)$ but this is to be expected because the tails of the sum of two assets are fatter than the tails of a single asset.

6.1 Model Prices

We start by hedging with the model prices paired with that models' deltas. The implications of this is that if the option was under or over priced at inception then the portfolio was set up at the wrong price. This means that the strategy could have fairly correct delta values and still gain or lose money based on a price the option could not actually be bought or sold for. We look at this, understanding its shortcomings, to test the stand alone accuracy of each model; later we will use market prices to adjust for this.

Figure 6.4: Histogram of $\chi^2(252)$
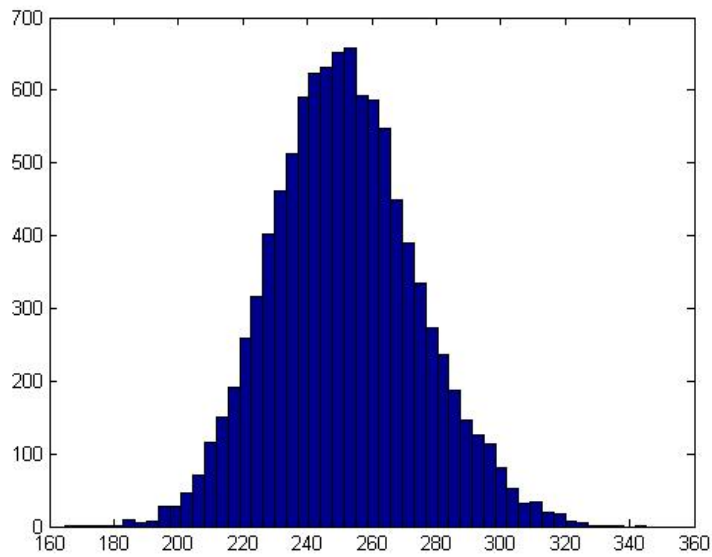
Figure 6.5: QQ Plot of $\chi^2(252)$ and Monte Carlo Hedging Profits, Monte Carlo Sample Size of 1000 with 1000 points plotted, Parameters as in Table 6.1
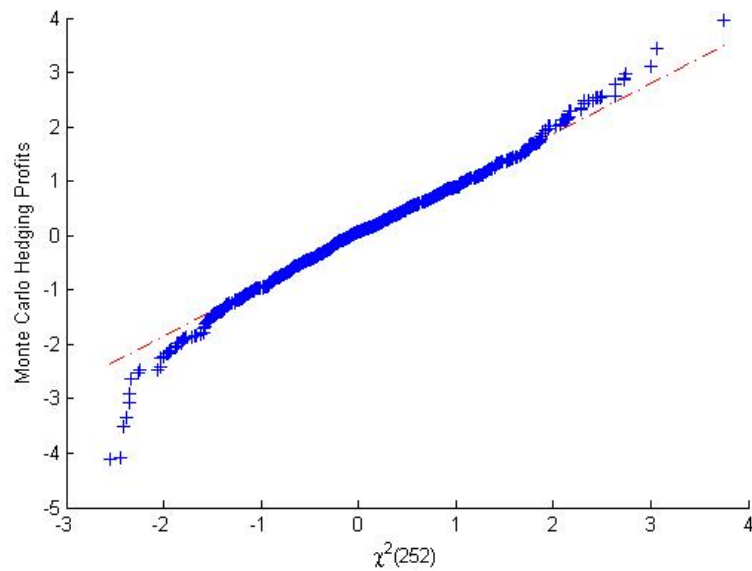
Table 6.2: Hedging Profits as r Varies

Model Price		1D BS	γ: Δ_I	γ: Δ_A, Δ_B	ADI	MC
		V_0	V_0	V_0	V_0	V_0
r varies		mean	mean	mean	mean	mean
		st.dev.	st.dev.	st.dev.	st.dev.	st.dev.
$r = 0.12$	$\rho = 0.30$	0.0273	0.0373	0.0373	0.1067	0.0352
$\sigma_A = 0.25$	$\sigma_B = 0.10$	−0.0003	−0.0025	−0.5588	−0.0680	0.0001
$\gamma = 0.25$	$I_0/K = 1.00$	0.0045	0.0653	1.4794	0.1076	0.0081
$r = 0.10$	$\rho = 0.30$	0.0671	0.0530	0.0530	0.1088	0.0545
$\sigma_A = 0.25$	$\sigma_B = 0.10$	−0.0004	−0.0002	−0.5922	−0.0383	0.0001
$\gamma = 0.25$	$I_0/K = 1.00$	0.0051	0.0094	1.4981	0.0901	0.0091
$r = 0.08$	$\rho = 0.30$	0.0848	0.0736	0.0736	0.1110	0.0736
$\sigma_A = 0.25$	$\sigma_B = 0.10$	−0.0005	−0.0000	−0.5133	−0.0124	0.0001
$\gamma = 0.25$	$I_0/K = 1.00$	0.0050	0.0097	1.4118	0.0701	0.0091
$r = 0.06$	$\rho = 0.30$	0.1015	0.1000	0.1000	0.1133	0.0989
$\sigma_A = 0.25$	$\sigma_B = 0.10$	−0.0003	0.0006	−0.6288	0.0090	0.0008
$\gamma = 0.25$	$I_0/K = 1.00$	0.0056	0.0109	1.5591	0.0609	0.0103
$r = 0.04$	$\rho = 0.30$	0.1532	0.1330	0.1330	0.1155	0.1338
$\sigma_A = 0.25$	$\sigma_B = 0.10$	−0.0003	0.0007	−0.5608	0.0313	0.0003
$\gamma = 0.25$	$I_0/K = 1.00$	0.0055	0.0110	1.4823	0.0520	0.0101
$r = 0.02$	$\rho = 0.30$	0.1777	0.1733	0.1733	0.1179	0.1742
$\sigma_A = 0.25$	$\sigma_B = 0.10$	0.0001	0.0010	−0.5009	0.0568	0.0002
$\gamma = 0.25$	$I_0/K = 1.00$	0.0051	0.0108	1.4104	0.0473	0.0100
$r = 0.00$	$\rho = 0.30$	0.2283	0.2215	0.2215	0.1203	0.2143
$\sigma_A = 0.25$	$\sigma_B = 0.10$	−0.0000	0.0009	−0.6070	0.0816	0.0004
$\gamma = 0.25$	$I_0/K = 1.00$	0.0054	0.0118	1.5472	0.0552	0.0108

In the following sections we vary each of six important parameters and compare each model's option price at inception which is the price paid for the option to set up the portfolio, V_0, and the mean and standard deviation of the hedging profits over the life of the option to see how effective the hedging strategy was.

The results for the Monte Carlo simulation, our benchmark, is in the right hand column and the variable values used for each case are on the left.

Risk Free Rate, r

In table 6.2 we can see that, given the other parameters used, for interest rates of 6% or higher the numerical ADI model is over priced, at $r = 12\%$ by 7¢ more than a 200% error on a 3¢ option. With the same other parameters, for an interest rate of 4% or lower the numerical ADI model is under priced, by as much as 11¢ at a nominal interest rate of 0%, this is now on a 21¢ option so we have an error of just over 50%. We can see that all of the Black-Scholes one-dimensional approximations price the option fairly accurately, but the "gamma" solutions

are slightly more accurate.

Even with such accurate pricing the Black-Scholes approximation and the "gamma" solution with one delta value still seem to lose money slightly more often than not, especially at high interest rates. What is especially noteworthy is that the "gamma" approximation that uses both deltas loses around 50¢ for every case, increasing as the option price increases or the interest rate decreases, losing from 2.8% up to over 15% of the option price. Whereas the numerical ADI solution loses less than it is mispriced by for interest rates of 4% or more. Lower than that it makes a little less money than it should given the mispricing.

The other interesting values on this chart are the standard deviations, especially of the hedging profits for the "gamma" approximation using both deltas. The standard deviation for the Monte Carlo simulations are of the order of magnitude of 10^{-3} to 10^{-2}. The one-dimensional Black-Scholes approximation and the "gamma" approximation that use only one delta value to hedge are on the order of 10^{-3} to 10^{-2} as well. But the numerical ADI solution is higher by a factor of 10 for $r = 10\%$, more for higher interest rates, less for lower ones, for $r = 0\%$ it is higher by a factor of only 5. And for the "gamma" approximation with both deltas it is of order 10^0, 7 to 50 times the correct option price, and this is consistent across all interest rates. This tells us, that even though we have an analytic approximation to the option price that is close to the true option value, that the hedging strategy derived from its pricing surface is very risky.

Correlation, ρ

Now we will vary the correlation between the dominant asset and the basket that is the rest of the stocks in the index.

In table 6.3 we see some of the same issues as before: huge losses with an even larger standard deviation for the "gamma" approximation that uses both deltas, mispricing in the ADI numerical solution, and profits that may not line up with over and under payment to set up the initial positions.

The loss and the deviation of that loss in the "gamma" approximation that uses both deltas increase as ρ increases. The numerical ADI solution is over priced for $\rho < 0.3$ and under priced for $\rho > 0.3$, there is a decrease in profits that follows from this over and under pricing but even considering that this strategy still seems to make a little money more often than not - though we will examine this later. The option prices in these cases are higher here than some in the previous section so the relative profits and standard deviations are lower. Though the ADI numerical solution's standard deviations are still on the same order of magnitude as the option prices.

Dominant Stock Volatility σ_A

Next we look at the volatility of the stock that dominates the index in table 6.4.

Again the "gamma" approximation that uses both deltas loses money on the scale of 3 to 5 times the option price, the loss increases as the volatility increases as does the standard deviation of the loss. The ADI method is a little over priced for high volatility but fairly well priced for low volatility but the standard deviation decreases relative to the option price as the volatility increases.

Table 6.3: Hedging Profits as ρ Varies

Model Price		1D BS	γ: Δ_I	γ: Δ_A, Δ_B	ADI	MC
		V_0	V_0	V_0	V_0	V_0
ρ varies		mean	mean	mean	mean	mean
		st.dev.	st.dev.	st.dev.	st.dev.	st.dev.
$r = 0.05$	$\rho = 0.90$	0.1456	0.1566	0.1566	0.1306	0.1522
$\sigma_A = 0.25$	$\sigma_B = 0.10$	−0.0007	−0.0000	−0.7001	0.0486	0.0005
$\gamma = 0.25$	$I_0/K = 1.00$	0.0062	0.0127	1.6378	0.0623	0.0127
$r = 0.05$	$\rho = 0.70$	0.1614	0.1437	0.1437	0.1254	0.1435
$\sigma_A = 0.25$	$\sigma_B = 0.10$	−0.0005	0.0001	−0.7118	0.0372	0.0005
$\gamma = 0.25$	$I_0/K = 1.00$	0.0059	0.0122	1.6396	0.0619	0.0118
$r = 0.05$	$\rho = 0.50$	0.1274	0.1301	0.1301	0.1200	0.1267
$\sigma_A = 0.25$	$\sigma_B = 0.10$	−0.0005	0.0006	−0.6951	0.0321	0.0006
$\gamma = 0.25$	$I_0/K = 1.00$	0.0058	0.0121	1.6292	0.0597	0.0115
$r = 0.05$	$\rho = 0.30$	0.1217	0.1157	0.1157	0.1144	0.1125
$\sigma_A = 0.25$	$\sigma_B = 0.10$	−0.0003	0.0003	−0.4883	0.0229	0.0000
$\gamma = 0.25$	$I_0/K = 1.00$	0.0058	0.0111	1.3767	0.0518	0.0102
$r = 0.05$	$\rho = 0.10$	0.1007	0.1004	0.1004	0.1086	0.0981
$\sigma_A = 0.25$	$\sigma_B = 0.10$	−0.0001	0.0009	−0.4083	0.0139	0.0005
$\gamma = 0.25$	$I_0/K = 1.00$	0.0043	0.0091	1.2747	0.0496	0.0083

Remaining Index Volatility σ_B

Next we look at the volatility of the remaining basket of stocks in the index after the dominating asset is separated in table 6.5.

The losses in the "gamma" approximation that uses both deltas are still 2 to 5 times the option price and increase as the volatility increases and the standard deviations also increase with the volatilities.

For the ADI numerical solution, unlike when the volatility of the dominant asset changed, at high volatilities of the remaining basket of assets the ADI price is significantly under priced and at low volatilities it is under priced. The standard deviations increase with the volatility and the means are higher than the over and under pricing implies.

Level of Dominance, γ

This brings us to consider the proportion of this index that is taken up by the dominating stock.

We see in table 6.6 that the "gamma" approximation that uses both deltas behaves similarly to the previous cases, the losses increase in magnitude and variation as γ increases. The ADI numerical solution is significantly under priced where γ is high, when the dominant stock that is more volatile than the rest makes up a higher proportion of the index. It is over priced where γ is low. For the Monte Carlo simulation and all three analytic approximations the option price at inception is surprisingly not monotone with respect to γ, though it is for the ADI solution.

Table 6.4: Hedging Profits as σ_A Varies

Model Price		1D BS	$\gamma: \Delta_I$	$\gamma: \Delta_A, \Delta_B$	ADI	MC
		V_0	V_0	V_0	V_0	V_0
σ_A varies		*mean*	*mean*	*mean*	*mean*	*mean*
		st.dev.	*st.dev.*	*st.dev.*	*st.dev.*	*st.dev.*
$r = 0.05$	$\rho = 0.03$	0.1556	0.1621	0.1621	0.1777	0.1523
$\sigma_A = 0.45$	$\sigma_B = 0.10$	-0.0039	-0.0096	-0.7856	0.0005	0.0002
$\gamma = 0.25$	$I_0/K = 1.00$	0.0095	0.1373	1.7031	0.0614	0.0130
$r = 0.05$	$\rho = 0.03$	0.1405	0.1260	0.1260	0.1417	0.1243
$\sigma_A = 0.35$	$\sigma_B = 0.10$	-0.0019	-0.0006	-0.6541	0.0054	0.0007
$\gamma = 0.25$	$I_0/K = 1.00$	0.0063	0.0160	1.5978	0.0533	0.0102
$r = 0.05$	$\rho = 0.03$	0.1006	0.0948	0.0948	0.1066	0.0954
$\sigma_A = 0.25$	$\sigma_B = 0.10$	-0.0003	0.0004	-0.3909	0.0120	0.0000
$\gamma = 0.25$	$I_0/K = 1.00$	0.0043	0.0086	1.2604	0.0458	0.0078
$r = 0.05$	$\rho = 0.03$	0.0734	0.0710	0.0710	0.0745	0.0714
$\sigma_A = 0.15$	$\sigma_B = 0.10$	0.0002	0.0003	-0.2597	0.0174	0.0002
$\gamma = 0.25$	$I_0/K = 1.00$	0.0032	0.0060	0.9988	0.0488	0.0062
$r = 0.05$	$\rho = 0.03$	0.0528	0.0574	0.0574	0.0527	0.0559
$\sigma_A = 0.05$	$\sigma_B = 0.10$	-0.0000	-0.0012	-0.1573	0.0203	0.0004
$\gamma = 0.25$	$I_0/K = 1.00$	0.0027	0.0052	0.8007	0.0435	0.0050

Table 6.5: Hedging Profits as σ_B Varies

Model Price		1D BS	$\gamma: \Delta_I$	$\gamma: \Delta_A, \Delta_B$	ADI	MC
		V_0	V_0	V_0	V_0	V_0
σ_B varies		*mean*	*mean*	*mean*	*mean*	*mean*
		st.dev.	*st.dev.*	*st.dev.*	*st.dev.*	*st.dev.*
$r = 0.05$	$\rho = 0.03$	0.2711	0.2379	0.2379	0.1309	0.2375
$\sigma_A = 0.25$	$\sigma_B = 0.22$	-0.0004	-0.0013	-1.1526	0.1171	0.0005
$\gamma = 0.25$	$I_0/K = 1.00$	0.0092	0.0188	2.0382	0.1143	0.0184
$r = 0.05$	$\rho = 0.03$	0.1624	0.1619	0.1619	0.1171	0.1663
$\sigma_A = 0.25$	$\sigma_B = 0.16$	0.0003	-0.0041	-0.8076	0.0584	0.0004
$\gamma = 0.25$	$I_0/K = 1.00$	0.0065	0.1583	1.7723	0.0792	0.0131
$r = 0.05$	$\rho = 0.03$	0.0888	0.0948	0.0948	0.1066	0.0945
$\sigma_A = 0.25$	$\sigma_B = 0.10$	-0.0003	0.0005	-0.4298	0.0124	-0.0001
$\gamma = 0.25$	$I_0/K = 1.00$	0.0046	0.0090	1.3018	0.0476	0.0084
$r = 0.05$	$\rho = 0.03$	0.0516	0.0478	0.0478	0.1002	0.0426
$\sigma_A = 0.25$	$\sigma_B = 0.04$	-0.0011	-0.0021	-0.1017	-0.0216	0.0004
$\gamma = 0.25$	$I_0/K = 1.00$	0.0028	0.0082	0.6215	0.0342	0.0037

Table 6.6: Hedging Profits as γ Varies

Model Price		1D BS	γ: Δ_I	γ: Δ_A, Δ_B	ADI	MC
		V_0	V_0	V_0	V_0	V_0
γ varies		*mean*	*mean*	*mean*	*mean*	*mean*
		st.dev.	*st.dev.*	*st.dev.*	*st.dev.*	*st.dev.*
$r = 0.05$	$\rho = 0.03$	0.1828	0.2176	0.2176	0.1066	0.2143
$\sigma_A = 0.25$	$\sigma_B = 0.10$	−0.0012	−0.0045	−0.8192	0.1291	0.0020
$\gamma = 0.65$	$I_0/K = 1.00$	0.0083	0.1603	1.7724	0.1176	0.0165
$r = 0.05$	$\rho = 0.03$	0.1952	0.1783	0.1783	0.1066	0.1785
$\sigma_A = 0.25$	$\sigma_B = 0.10$	−0.0018	−0.0007	−0.7003	0.0860	0.0012
$\gamma = 0.55$	$I_0/K = 1.00$	0.0072	0.0167	1.6486	0.0883	0.0127
$r = 0.05$	$\rho = 0.03$	0.1635	0.1434	0.1434	0.1066	0.1391
$\sigma_A = 0.25$	$\sigma_B = 0.10$	−0.0015	−0.0019	−0.5762	0.0499	0.0000
$\gamma = 0.45$	$I_0/K = 1.00$	0.0068	0.0158	1.4977	0.0709	0.0112
$r = 0.05$	$\rho = 0.03$	0.1208	0.1147	0.1147	0.1066	0.1159
$\sigma_A = 0.25$	$\sigma_B = 0.10$	−0.0010	0.0002	−0.6107	0.0266	0.0006
$\gamma = 0.35$	$I_0/K = 1.00$	0.0055	0.0117	1.5291	0.0608	0.0091
$r = 0.05$	$\rho = 0.03$	0.0872	0.0948	0.0948	0.1066	0.0929
$\sigma_A = 0.25$	$\sigma_B = 0.10$	−0.0003	0.0007	−0.4593	0.0092	0.0001
$\gamma = 0.25$	$I_0/K = 1.00$	0.0047	0.0091	1.3057	0.0509	0.0086
$r = 0.05$	$\rho = 0.03$	0.0736	0.0861	0.0861	0.1060	0.0864
$\sigma_A = 0.25$	$\sigma_B = 0.10$	0.0001	0.0003	−0.4699	0.0018	−0.0000
$\gamma = 0.15$	$I_0/K = 1.00$	0.0039	0.0073	1.3681	0.0512	0.0076
$r = 0.05$	$\rho = 0.03$	0.0953	0.0899	0.0899	0.0912	0.0892
$\sigma_A = 0.25$	$\sigma_B = 0.10$	0.0001	−0.0001	−0.5034	0.0108	0.0001
$\gamma = 0.05$	$I_0/K = 1.00$	0.0042	0.0084	1.3988	0.0429	0.0084

The ADI numerical solution, as before, makes money more often than it loses money in all cases.

In the Moneyness, I_0/K

Finally we look at how being in or out of the money at inception effects the outcomes in table 6.7. Now we compare the hedging profits of a strategy when the option does not necessarily start at the money. As should be expected if the option starts too far out of the money the option is worthless, and this holds across all the pricing models. As the option moves closer to being at the money the option starts to have a non-zero price, the ADI numerical solution undervalues these options but the rest of the methods price these fairly well. At the money the ADI numerical solution begins to over price the option and as it moves further into the money the over pricing becomes more drastic. The option prices are fairly accurate for the other models when the option is at or in the money.

   The ADI numerical solution as before seems to lose less than the over and under pricing should imply with deviations that are highest at the money and decrease as the option starts further in or out of the money. This is the same for the standard deviations of the "gamma" approximation that uses both deltas but the losses increase as the option moves from out of the money toward being in the money. The standard deviations of the "gamma" approximation that uses only one delta increase by an order of magnitude when the option starts in the money, but stays constant for all levels of being in the money.

6.2   Market Price

Now we use the option prices from the Monte Carlo simulation as the price of our options to set up our portfolio, this means that the fair price is actually used in the hedging strategy and that any profits that came from mispricing, and then the interest on that mispricing, are eliminated. This also means that if the market prices were truly the price that the option traded at then the portfolio could really be set up, whereas model prices are not the price you can pay for an option.

   Since the ADI numerical solution is the only model that has drastic mispricing issues we will focus mostly on that model here since the other trends have already been discussed.

Risk Free Rate, r

In table 6.8 we can see that without mispricing of the option to set up the portfolio that the profits and losses from the ADI numerical solution are much more modest than before. In all these cases the mean profit is less that 100% of the option price. For high interest rates this strategy makes a little money on average and for low interest rates it loses a little money on average. This must come from inaccuracies in the deltas because the strategy requires borrowing money so the effect that low interest rates have on that should increase profits relative to high interest rates.

   That the ADI method still has higher profits than the other methods on average is because of the extreme changes in the Delta values near the strike price. This causes the strategy to

Table 6.7: Hedging Profits as I_0/K Varies

Model Price		1D BS	$\gamma: \Delta_I$	$\gamma: \Delta_A, \Delta_B$	ADI	MC
		V_0	V_0	V_0	V_0	V_0
I_0/K varies		*mean*	*mean*	*mean*	*mean*	*mean*
		st.dev.	*st.dev.*	*st.dev.*	*st.dev.*	*st.dev.*
$r = 0.05$	$\rho = 0.03$	0.0000	0.0000	0.0000	0.0000	0.0000
$\sigma_A = 0.25$	$\sigma_B = 0.10$	−0.0000	−0.0000	0.0000	0.0000	0.0000
$\gamma = 0.25$	$I_0/K = 1.50$	0.0000	0.0000	0.0000	0.0000	0.0000
$r = 0.05$	$\rho = 0.03$	0.0000	0.0000	0.0000	0.0000	0.0000
$\sigma_A = 0.25$	$\sigma_B = 0.10$	−0.0000	−0.0000	0.0000	0.0000	−0.0000
$\gamma = 0.25$	$I_0/K = 1.40$	0.0000	0.0000	0.0000	0.0000	0.0000
$r = 0.05$	$\rho = 0.03$	0.0001	0.0001	0.0001	0.0000	0.0002
$\sigma_A = 0.25$	$\sigma_B = 0.10$	−0.0000	−0.0000	0.0000	0.0000	−0.0000
$\gamma = 0.25$	$I_0/K = 1.30$	0.0001	0.0001	0.0001	0.0001	0.0001
$r = 0.05$	$\rho = 0.03$	0.0019	0.0017	0.0017	0.0001	0.0015
$\sigma_A = 0.25$	$\sigma_B = 0.10$	−0.0001	−0.0001	−0.0048	0.0006	−0.0000
$\gamma = 0.25$	$I_0/K = 1.20$	0.0007	0.0011	0.1581	0.0035	0.0010
$r = 0.05$	$\rho = 0.03$	0.0226	0.0160	0.0160	0.0055	0.0153
$\sigma_A = 0.25$	$\sigma_B = 0.10$	−0.0000	0.0001	−0.0312	0.0066	0.0002
$\gamma = 0.25$	$I_0/K = 1.10$	0.0020	0.0032	0.3432	0.0137	0.0032
$r = 0.05$	$\rho = 0.03$	0.1037	0.0948	0.0948	0.1066	0.0933
$\sigma_A = 0.25$	$\sigma_B = 0.10$	−0.0002	0.0007	−0.4519	0.0137	0.0002
$\gamma = 0.25$	$I_0/K = 1.00$	0.0046	0.0092	1.3194	0.0465	0.0087
$r = 0.05$	$\rho = 0.03$	0.3387	0.3387	0.3387	0.4805	0.3415
$\sigma_A = 0.25$	$\sigma_B = 0.10$	0.0002	−0.0027	−1.9900	−0.0964	0.0003
$\gamma = 0.25$	$I_0/K = 0.90$	0.0061	0.1581	2.2830	0.0858	0.0112
$r = 0.05$	$\rho = 0.03$	0.7650	0.7632	0.7632	0.9513	0.7678
$\sigma_A = 0.25$	$\sigma_B = 0.10$	0.0006	−0.0002	−3.8670	−0.1798	0.0002
$\gamma = 0.25$	$I_0/K = 0.80$	0.0052	0.0608	1.9804	0.0502	0.0079
$r = 0.05$	$\rho = 0.03$	1.2563	1.2563	1.2563	1.4267	1.2559
$\sigma_A = 0.25$	$\sigma_B = 0.10$	0.0002	−0.0048	−4.9037	−0.1866	−0.0000
$\gamma = 0.25$	$I_0/K = 0.70$	0.0018	0.1581	0.6383	0.0134	0.0042
$r = 0.05$	$\rho = 0.03$	1.7561	1.7561	1.7561	1.9021	1.7551
$\sigma_A = 0.25$	$\sigma_B = 0.10$	0.0000	−0.0050	−4.9914	−0.1611	−0.0001
$\gamma = 0.25$	$I_0/K = 0.60$	0.0003	0.1581	0.1940	0.0096	0.0032
$r = 0.05$	$\rho = 0.03$	2.2561	2.2561	2.2561	2.3774	2.2561
$\sigma_A = 0.25$	$\sigma_B = 0.10$	0.0000	−0.0050	−4.9950	−0.1341	−0.0000
$\gamma = 0.25$	$I_0/K = 0.50$	0.0000	0.1581	0.1581	0.0084	0.0026

Table 6.8: Hedging Profits as r Varies

Market Price		1D BS	$\gamma: \Delta_I$	$\gamma: \Delta_A, \Delta_B$	ADI	MC
		V_0	V_0	V_0	V_0	V_0
r varies		mean	mean	mean	mean	mean
		st.dev.	st.dev.	st.dev.	st.dev.	st.dev.
$r = 0.12$	$\rho = 0.30$	0.0362	0.0362	0.0362	0.0362	0.0362
$\sigma_A = 0.25$	$\sigma_B = 0.10$	0.0011	0.0000	−0.6143	0.0178	−0.0006
$\gamma = 0.25$	$I_0/K = 1.00$	0.0088	0.0094	1.5167	0.1031	0.0087
$r = 0.10$	$\rho = 0.30$	0.0491	0.0491	0.0491	0.0491	0.0491
$\sigma_A = 0.25$	$\sigma_B = 0.10$	0.0016	0.0004	−0.5093	0.0238	−0.0005
$\gamma = 0.25$	$I_0/K = 1.00$	0.0094	0.0093	1.3870	0.0899	0.0087
$r = 0.08$	$\rho = 0.30$	0.0726	0.0726	0.0726	0.0726	0.0726
$\sigma_A = 0.25$	$\sigma_B = 0.10$	0.0025	0.0014	−0.6183	0.0229	0.0004
$\gamma = 0.25$	$I_0/K = 1.00$	0.0104	0.0100	1.5498	0.0744	0.0091
$r = 0.06$	$\rho = 0.30$	0.1010	0.1010	0.1010	0.1010	0.1010
$\sigma_A = 0.25$	$\sigma_B = 0.10$	0.0021	0.0012	−0.6024	0.0195	0.0003
$\gamma = 0.25$	$I_0/K = 1.00$	0.0118	0.0106	1.5312	0.0651	0.0098
$r = 0.04$	$\rho = 0.30$	0.1337	0.1337	0.1337	0.1337	0.1337
$\sigma_A = 0.25$	$\sigma_B = 0.10$	0.0019	0.0011	−0.5393	0.0127	0.0003
$\gamma = 0.25$	$I_0/K = 1.00$	0.0121	0.0109	1.4699	0.0475	0.0097
$r = 0.02$	$\rho = 0.30$	0.1745	0.1745	0.1745	0.1745	0.1745
$\sigma_A = 0.25$	$\sigma_B = 0.10$	0.0012	0.0007	−0.5230	−0.0007	0.0001
$\gamma = 0.25$	$I_0/K = 1.00$	0.0125	0.0113	1.4392	0.0488	0.0102
$r = 0.00$	$\rho = 0.30$	0.2206	0.2206	0.2206	0.2206	0.2206
$\sigma_A = 0.25$	$\sigma_B = 0.10$	0.0010	0.0009	−0.6195	−0.0196	0.0003
$\gamma = 0.25$	$I_0/K = 1.00$	0.0127	0.0118	1.5469	0.0570	0.0108

Table 6.9: Hedging Profits as ρ Varies

Market Price		1D BS	γ: Δ_I	γ: Δ_A, Δ_B	ADI	MC
		V_0	V_0	V_0	V_0	V_0
ρ varies		mean	mean	mean	mean	mean
		st.dev.	st.dev.	st.dev.	st.dev.	st.dev.
$r = 0.05$	$\rho = 0.90$	0.1509	0.1509	0.1509	0.1509	0.1509
$\sigma_A = 0.25$	$\sigma_B = 0.10$	0.0025	0.0015	−0.8123	0.0178	0.0001
$\gamma = 0.25$	$I_0/K = 1.00$	0.0145	0.0136	1.7585	0.0682	0.0130
$r = 0.05$	$\rho = 0.70$	0.1402	0.1402	0.1402	0.1402	0.1402
$\sigma_A = 0.25$	$\sigma_B = 0.10$	0.0030	0.0019	−0.6550	0.0176	0.0007
$\gamma = 0.25$	$I_0/K = 1.00$	0.0131	0.0123	1.6086	0.0605	0.0114
$r = 0.05$	$\rho = 0.50$	0.1284	0.1284	0.1284	0.1284	0.1284
$\sigma_A = 0.25$	$\sigma_B = 0.10$	0.0018	0.0011	−0.6886	0.0136	−0.0002
$\gamma = 0.25$	$I_0/K = 1.00$	0.0129	0.0119	1.6102	0.0601	0.0108
$r = 0.05$	$\rho = 0.30$	0.1155	0.1155	0.1155	0.1155	0.1155
$\sigma_A = 0.25$	$\sigma_B = 0.10$	0.0019	0.0014	−0.5549	0.0185	0.0004
$\gamma = 0.25$	$I_0/K = 1.00$	0.0120	0.0104	1.4496	0.0539	0.0094
$r = 0.05$	$\rho = 0.10$	0.0993	0.0993	0.0993	0.0993	0.0993
$\sigma_A = 0.25$	$\sigma_B = 0.10$	0.0015	0.0005	−0.4484	0.0198	−0.0001
$\gamma = 0.25$	$I_0/K = 1.00$	0.0111	0.0097	1.3120	0.0525	0.0092

over hedge; since hedging follows the "buy low, sell high" strategy it usually makes money. Unfortunately this is also an imperfect hedge, so there is a chance of large tail losses too; we will look at this more later.

The standard deviations did not change from the model price, the standard deviation is high for high interest rates and lower for low interest rates. For all of the cases here a 68% confidence interval of the hedging profits contains zero. This means that we can not reject a null hypothesis of hedging profits equal zero. This is also the case for all of the other models, though the high standard deviation of the "gamma" model that uses both deltas is needed to balance out the consistently large losses.

Correlation, ρ

Next we look at table 6.9. For various levels of ρ the hedging on the ADI numerical solution consistently makes about 2¢ with a standard deviation of 5¢ to 7¢. The profit increases slightly and the standard deviation decreases as ρ decreases. The other three models behave as they did for the model prices.

Dominant Stock Volatility σ_A

In table 6.10 we see that hedging on the ADI numerical solution again yields a fairly consistent profit of about 2¢ on options that range in initial price from 5¢ to 15¢, which increases slightly

Table 6.10: Hedging Profits as σ_A Varies

Market Price		1D BS	γ: Δ_I	γ: Δ_A, Δ_B	ADI	MC
		V_0	V_0	V_0	V_0	V_0
σ_A varies		*mean*	*mean*	*mean*	*mean*	*mean*
		st.dev.	*st.dev.*	*st.dev.*	*st.dev.*	*st.dev.*
$r = 0.05$	$\rho = 0.03$	0.1590	0.1590	0.1590	0.1590	0.1590
$\sigma_A = 0.45$	$\sigma_B = 0.10$	0.0081	0.0020	-0.7645	0.0211	0.0000
$\gamma = 0.25$	$I_0/K = 1.00$	0.0329	0.0285	1.6868	0.0634	0.0128
$r = 0.05$	$\rho = 0.03$	0.1235	0.1235	0.1235	0.1235	0.1235
$\sigma_A = 0.35$	$\sigma_B = 0.10$	0.0041	0.0016	-0.6065	0.0212	0.0001
$\gamma = 0.25$	$I_0/K = 1.00$	0.0181	0.0165	1.5308	0.0555	0.0107
$r = 0.05$	$\rho = 0.03$	0.0952	0.0952	0.0952	0.0952	0.0952
$\sigma_A = 0.25$	$\sigma_B = 0.10$	0.0022	0.0015	-0.3888	0.0226	0.0008
$\gamma = 0.25$	$I_0/K = 1.00$	0.0107	0.0092	1.2378	0.0461	0.0084
$r = 0.05$	$\rho = 0.03$	0.0697	0.0697	0.0697	0.0697	0.0697
$\sigma_A = 0.15$	$\sigma_B = 0.10$	0.0002	0.0001	-0.2787	0.0190	0.0002
$\gamma = 0.25$	$I_0/K = 1.00$	0.0076	0.0065	1.0614	0.0448	0.0066
$r = 0.05$	$\rho = 0.03$	0.0562	0.0562	0.0562	0.0562	0.0562
$\sigma_A = 0.05$	$\sigma_B = 0.10$	0.0001	-0.0003	-0.1726	0.0168	0.0005
$\gamma = 0.25$	$I_0/K = 1.00$	0.0066	0.0053	0.8209	0.0458	0.0050

for high volatilities of the dominant stock or high option prices, as does their standard deviations. A 68% confidence interval still contains zero so we can not say with much confidence that the profits do not equal zero.

Remaining Index Volatility σ_B

In table 6.11 we see that for high volatility of the remaining stocks in the index after the dominating stock is removed the mean profit is very close to zero but has a very high standard deviation. As the volatility decreases the mean profit increases and its standard deviation decreases. These same trends are followed by the large losses and larger standard deviations of the "gamma" model that uses both deltas.

At the lowest volatility, $\sigma_B = 0.04$, a 68% confidence interval no longer contains zero so at that confidence level we can now reject a null hypothesis that the hedging profits equal zero.

Level of Dominance, γ

As γ increases in table 6.12 the dominant asset takes over more of the index at the option's inception. At high values of γ the standard deviation of hedging profits from the ADI numerical solution are high, though only half the option's price. As γ decreases so does the variation, but the option price also decreases at the same rate. As the option price and γ decrease the mean hedging profits increase, except at very low levels of γ, where the profit decreases sharply. The

Table 6.11: Hedging Profits as σ_B Varies

Market Price		1D BS	$\gamma: \Delta_I$	$\gamma: \Delta_A, \Delta_B$	ADI	MC
		V_0	V_0	V_0	V_0	V_0
σ_B varies		mean	mean	mean	mean	mean
		st.dev.	st.dev.	st.dev.	st.dev.	st.dev.
$r = 0.05$	$\rho = 0.03$	0.2277	0.2277	0.2277	0.2277	0.2277
$\sigma_A = 0.25$	$\sigma_B = 0.22$	−0.0019	−0.0018	−1.2049	0.0068	−0.0002
$\gamma = 0.25$	$I_0/K = 1.00$	0.0197	0.0200	2.0592	0.1186	0.0189
$r = 0.05$	$\rho = 0.03$	0.1619	0.1619	0.1619	0.1619	0.1619
$\sigma_A = 0.25$	$\sigma_B = 0.16$	−0.0006	−0.0003	−0.8273	0.0178	−0.0000
$\gamma = 0.25$	$I_0/K = 1.00$	0.0139	0.0135	1.7624	0.0775	0.0136
$r = 0.05$	$\rho = 0.03$	0.0948	0.0948	0.0948	0.0948	0.0948
$\sigma_A = 0.25$	$\sigma_B = 0.10$	0.0020	0.0013	−0.3916	0.0214	0.0004
$\gamma = 0.25$	$I_0/K = 1.00$	0.0108	0.0092	1.2058	0.0512	0.0082
$r = 0.05$	$\rho = 0.03$	0.0444	0.0444	0.0444	0.0444	0.0444
$\sigma_A = 0.25$	$\sigma_B = 0.04$	0.0055	0.0015	−0.1074	0.0360	0.0000
$\gamma = 0.25$	$I_0/K = 1.00$	0.0106	0.0080	0.6180	0.0343	0.0036

same trends are followed by the hedging profits of the "gamma" model that uses both deltas.

In the Moneyness, I_0/K

Lats we look at table 6.13. As the option moves less out of the money at inception the option price rises from zero but the mean hedging profit for the ADI numerical solution stays around zero while the standard deviation grows. The standard deviation grows until it is slightly in the money, and then it decreases as it goes further into the money. At the money and a little bit in the money the mean hedging profit is positive, but only a little bit relative to the option price. For options that start deep in or out of the money there are very slight hedging losses.

The hedging profits from the "gamma" model that uses both deltas have increasingly large losses as the option price increases but the deviations increase toward par value.

6.3 Hedging Profit Distributions

Finally we will compare the hedging profit distributions, using the market price convention[2], rather than just expected profits. This was done as an example for the Monte Carlo solution at the beginning of this section, see figures 6.3 and 6.5. From the Monte Carlo Simulations we saw that the fat tailedness of the sum of two lognormal random variables, or the dominating stock and the rest, comes through in the hedging profits, the tails were fatter than $\chi^2(252)$ which

---

[2]The market price convention is what was used in section 6.2 rather than 6.1; using option prices to setup the portfolio that are from Monte Carlo rather than each model. This gives a price that the option could really be purchased for rather than just the price that a model produces.

Table 6.12: Hedging Profits as γ Varies

Market Price		1D BS	$\gamma: \Delta_I$	$\gamma: \Delta_A, \Delta_B$	ADI	MC
		V_0	V_0	V_0	V_0	V_0
γ varies		mean	mean	mean	mean	mean
		st.dev.	st.dev.	st.dev.	st.dev.	st.dev.
$r = 0.05$	$\rho = 0.03$	0.2143	0.2143	0.2143	0.2143	0.2143
$\sigma_A = 0.25$	$\sigma_B = 0.10$	0.0026	0.0005	-0.8595	0.0099	-0.0002
$\gamma = 0.65$	$I_0/K = 1.00$	0.0223	0.0211	1.7952	0.1097	0.0166
$r = 0.05$	$\rho = 0.03$	0.1760	0.1760	0.1760	0.1760	0.1760
$\sigma_A = 0.25$	$\sigma_B = 0.10$	0.0047	0.0017	-0.7620	0.0115	0.0003
$\gamma = 0.55$	$I_0/K = 1.00$	0.0206	0.0186	1.6942	0.0877	0.0131
$r = 0.05$	$\rho = 0.03$	0.1436	0.1436	0.1436	0.1436	0.1436
$\sigma_A = 0.25$	$\sigma_B = 0.10$	0.0046	0.0018	-0.6701	0.0130	0.0004
$\gamma = 0.45$	$I_0/K = 1.00$	0.0171	0.0154	1.5958	0.0722	0.0113
$r = 0.05$	$\rho = 0.03$	0.1150	0.1150	0.1150	0.1150	0.1150
$\sigma_A = 0.25$	$\sigma_B = 0.10$	0.0037	0.0022	-0.5148	0.0201	0.0006
$\gamma = 0.35$	$I_0/K = 1.00$	0.0132	0.0116	1.4174	0.0563	0.0092
$r = 0.05$	$\rho = 0.03$	0.0944	0.0944	0.0944	0.0944	0.0944
$\sigma_A = 0.25$	$\sigma_B = 0.10$	0.0007	0.0000	-0.4916	0.0217	-0.0005
$\gamma = 0.25$	$I_0/K = 1.00$	0.0102	0.0091	1.3670	0.0479	0.0082
$r = 0.05$	$\rho = 0.03$	0.0841	0.0841	0.0841	0.0841	0.0841
$\sigma_A = 0.25$	$\sigma_B = 0.10$	0.0001	0.0001	-0.3747	0.0255	0.0001
$\gamma = 0.15$	$I_0/K = 1.00$	0.0083	0.0076	1.2214	0.0464	0.0076
$r = 0.05$	$\rho = 0.03$	0.0899	0.0899	0.0899	0.0899	0.0899
$\sigma_A = 0.25$	$\sigma_B = 0.10$	0.0002	0.0001	-0.4808	0.0136	0.0003
$\gamma = 0.05$	$I_0/K = 1.00$	0.0093	0.0085	1.3821	0.0462	0.0086

Table 6.13: Hedging Profits as I_0/K Varies

Market Price		1D BS	$\gamma: \Delta_I$	$\gamma: \Delta_A, \Delta_B$	ADI	MC
		V_0	V_0	V_0	V_0	V_0
I_0/K varies		mean	mean	mean	mean	mean
		st.dev.	st.dev.	st.dev.	st.dev.	st.dev.
$r = 0.05$	$\rho = 0.03$	0.0000	0.0000	0.0000	0.0000	0.0000
$\sigma_A = 0.25$	$\sigma_B = 0.10$	0.0000	0.0000	0.0000	−0.0000	0.0000
$\gamma = 0.25$	$I_0/K = 1.50$	0.0000	0.0000	0.0000	0.0000	0.0000
$r = 0.05$	$\rho = 0.03$	0.0000	0.0000	0.0000	0.0000	0.0000
$\sigma_A = 0.25$	$\sigma_B = 0.10$	0.0000	0.0000	0.0000	−0.0000	0.0000
$\gamma = 0.25$	$I_0/K = 1.40$	0.0000	0.0000	0.0000	0.0000	0.0000
$r = 0.05$	$\rho = 0.03$	0.0001	0.0001	0.0001	0.0001	0.0001
$\sigma_A = 0.25$	$\sigma_B = 0.10$	0.0001	0.0000	0.0000	−0.0001	0.0000
$\gamma = 0.25$	$I_0/K = 1.30$	0.0001	0.0001	0.0007	0.0001	0.0001
$r = 0.05$	$\rho = 0.03$	0.0018	0.0018	0.0018	0.0018	0.0018
$\sigma_A = 0.25$	$\sigma_B = 0.10$	0.0003	0.0001	−0.0046	−0.0011	−0.0000
$\gamma = 0.25$	$I_0/K = 1.20$	0.0010	0.0007	0.1573	0.0028	0.0006
$r = 0.05$	$\rho = 0.03$	0.0158	0.0158	0.0158	0.0158	0.0158
$\sigma_A = 0.25$	$\sigma_B = 0.10$	0.0009	0.0003	−0.0447	−0.0041	−0.0000
$\gamma = 0.25$	$I_0/K = 1.10$	0.0046	0.0038	0.3738	0.0164	0.0036
$r = 0.05$	$\rho = 0.03$	0.0947	0.0947	0.0947	0.0947	0.0947
$\sigma_A = 0.25$	$\sigma_B = 0.10$	0.0013	0.0005	−0.4727	0.0200	−0.0000
$\gamma = 0.25$	$I_0/K = 1.00$	0.0101	0.0087	1.3562	0.0488	0.0080
$r = 0.05$	$\rho = 0.03$	0.3453	0.3453	0.3453	0.3453	0.3453
$\sigma_A = 0.25$	$\sigma_B = 0.10$	−0.0000	0.0005	−1.7819	0.0538	−0.0004
$\gamma = 0.25$	$I_0/K = 0.90$	0.0115	0.0130	2.2298	0.0863	0.0114
$r = 0.05$	$\rho = 0.03$	0.7660	0.7660	0.7660	0.7660	0.7660
$\sigma_A = 0.25$	$\sigma_B = 0.10$	0.0004	−0.0042	−4.0089	0.0134	−0.0000
$\gamma = 0.25$	$I_0/K = 0.80$	0.0073	0.1585	1.8736	0.0500	0.0085
$r = 0.05$	$\rho = 0.03$	1.2629	1.2629	1.2629	1.2629	1.2629
$\sigma_A = 0.25$	$\sigma_B = 0.10$	0.0003	−0.0046	−4.9086	−0.0061	0.0002
$\gamma = 0.25$	$I_0/K = 0.70$	0.0044	0.1586	0.6010	0.0196	0.0047
$r = 0.05$	$\rho = 0.03$	1.7577	1.7577	1.7577	1.7577	1.7577
$\sigma_A = 0.25$	$\sigma_B = 0.10$	0.0001	−0.0049	−4.9949	−0.0078	0.0001
$\gamma = 0.25$	$I_0/K = 0.60$	0.0031	0.1582	0.1581	0.0103	0.0031
$r = 0.05$	$\rho = 0.03$	2.2534	2.2534	2.2534	2.2534	2.2534
$\sigma_A = 0.25$	$\sigma_B = 0.10$	−0.0001	−0.0051	−4.9951	−0.0067	−0.0001
$\gamma = 0.25$	$I_0/K = 0.50$	0.0025	0.1580	0.1581	0.0084	0.0026

would be the distribution of the hedging profits for a put on a single underlying asset. Now we will look at the other methods.

First we look at the simplest method, the one-dimensional Black-Scholes method which ignores that an underlying asset is not evenly weighted and prices the option as though there was a single underlying asset. From figures 6.6 and 6.7 we see that the hedging profits look fairly normally distributed, and the QQ plot shows us that the hedging profits are very similarly distributed to $\chi^2(252)$.

Next we look at figures 6.8 and 6.9, the "gamma" approximation with only one delta value, Δ_I, which can be found analytically and it varies as the make up of the index varies through varying σ^*. This is also close to the $\chi^2(252)$ distribution with slightly fatter tails.

In contrast, in figures 6.10 and 6.11, using the "gamma" approximation with both underlying deltas, Δ_A and Δ_B which need to be numerically approximated, there is a much larger chance of a big loss. This is from numerical error in approximating the delta values and using a pricing surface that is designed by one delta value and forcing two deltas out of it. The tails here are not balanced, there is no balancing possibility of a large gain.

The ADI solution, in figures 6.12 and 6.13, also has numerical error from interpolating the delta values, but the error is much smoother because the surface was designed by both delta values. Compared to the previous solution this one has similarly fat tails of both the left and the right, meaning that the possibility of a big loss is somewhat balanced out by the possibility of a big gain.

Now that we have examined our various pricing methods and hedging strategies for various theoretical parameters and have seen that our theoretical hedging profits are similar to previous well known results, we are now ready to see how our methods perform for some empirical data.

Figure 6.6: Histogram of Hedging Profits from One-Dimensional Black-Scholes, Parameters as in Table 6.1
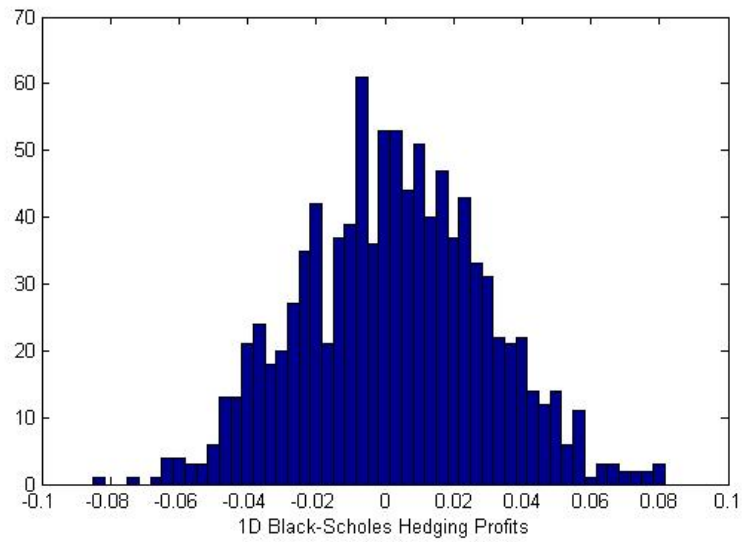

Figure 6.7: QQ Plot of Hedging Profits from One-Dimensional Black-Scholes Against $\chi^2(252)$, Parameters as in Table 6.1

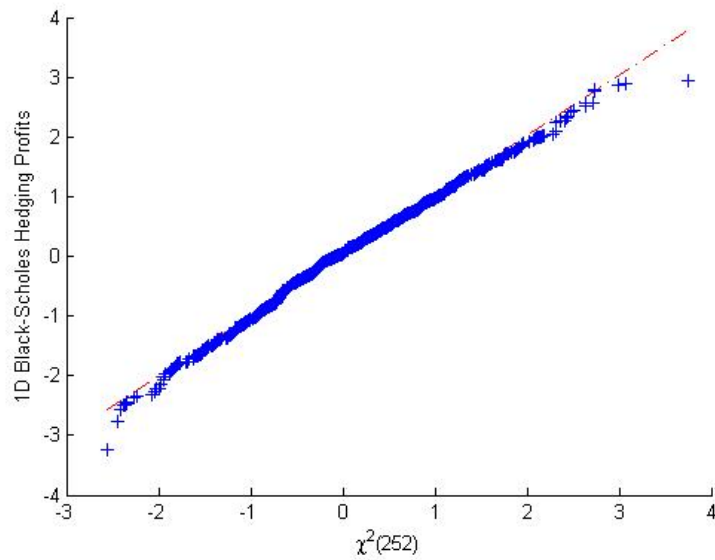Figure 6.8: Histogram of Hedging Profits from One-Dimensional "gamma" Approximation, Parameters as in Table 6.1

Figure 6.9: QQ Plot of Hedging Profits from One-Dimensional "gamma" Approximation Against $\chi^2(252)$, Parameters as in Table 6.1

Figure 6.10: Histogram of Hedging Profits from Two-Dimensional "gamma" Approximation, Parameters as in Table 6.1
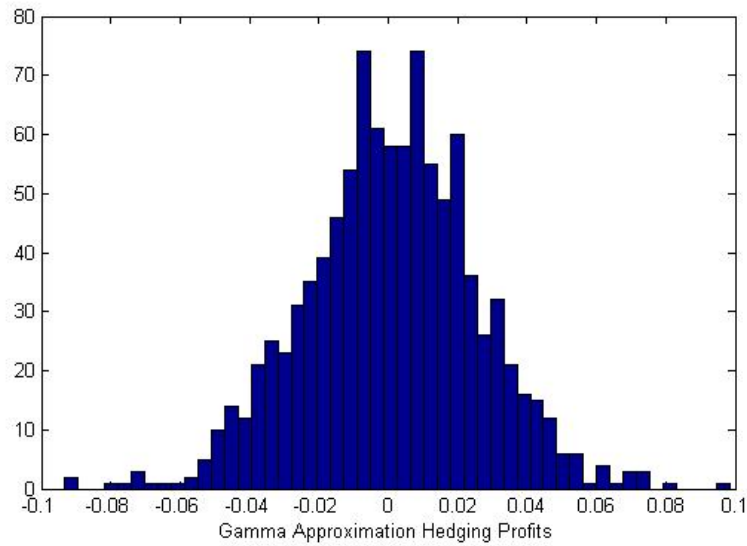

Figure 6.11: QQ Plot of Hedging Profits from Two-Dimensional "gamma" Approximation Against $\chi^2(252)$, Parameters as in Table 6.1
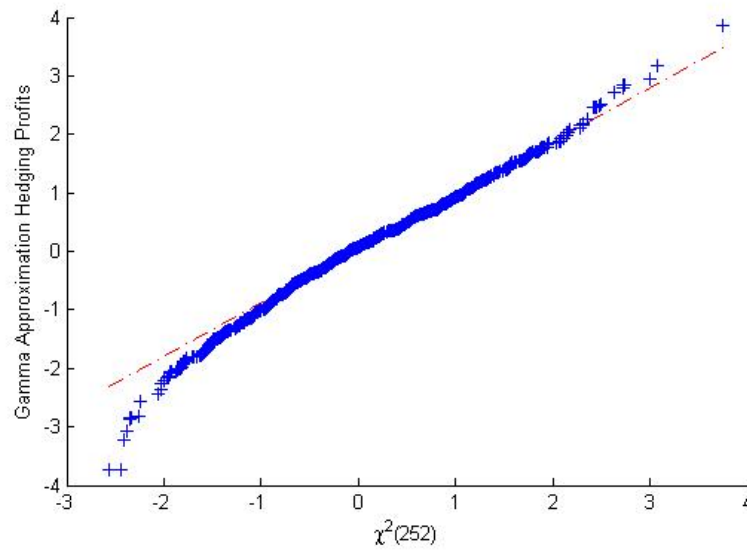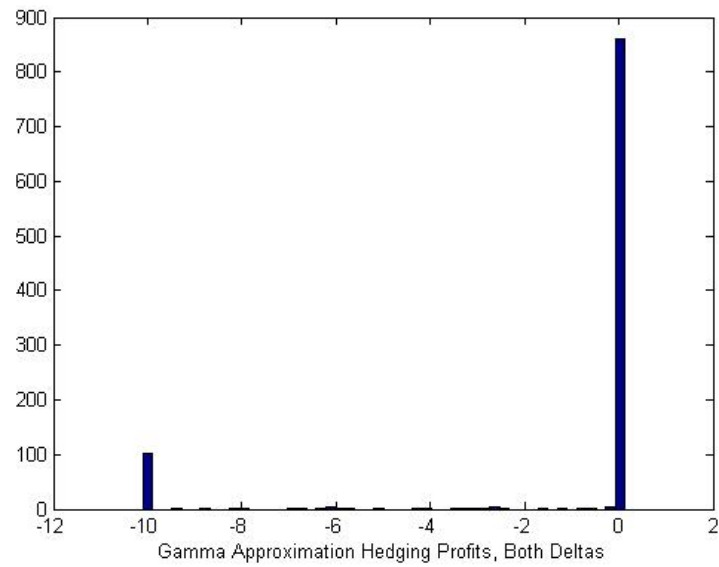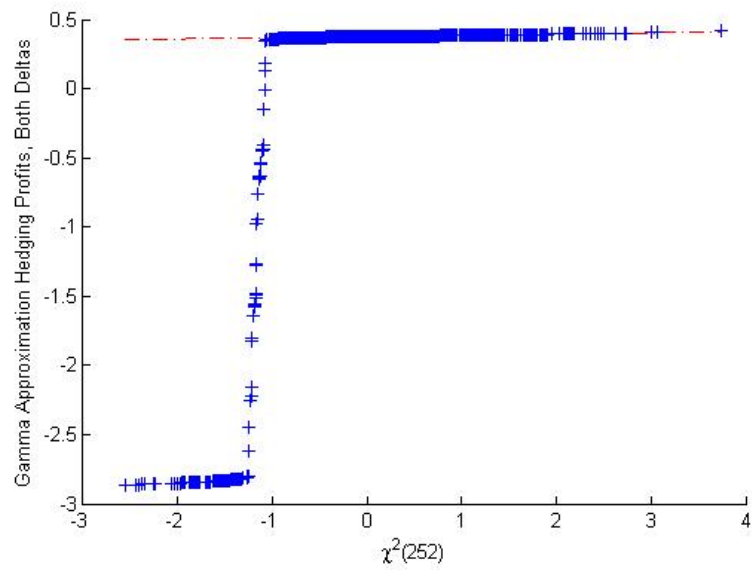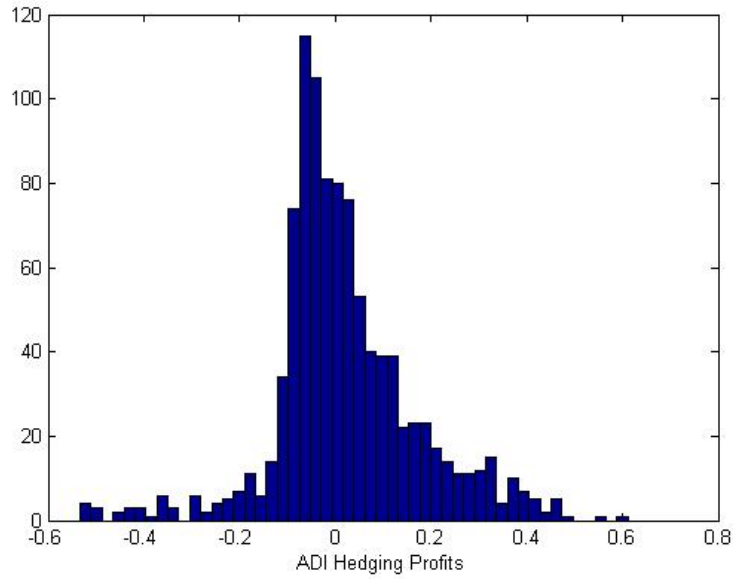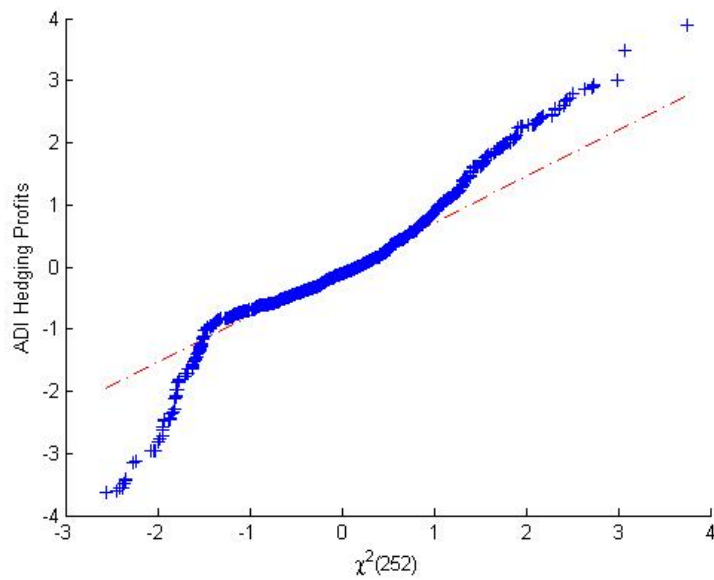
Figure 6.12: Histogram of Hedging Profits from ADI Solution, Parameters as in Table 6.1

Figure 6.13: QQ Plot of Hedging Profits from ADI Solution Against $\chi^2(252)$, Parameters as in Table 6.1

Chapter 7

Empirical Results

Now that we have examined the different methods' pricing surfaces, their deltas and the distributions of the resulting hedging profits, we will look at an empirical example. It is difficult, or costly, to get detailed index weighting data for North American and European indices, but Asian data is more available so we will look at the South Korean Index, KOSPI, almost 20% of which is currently made up by Samsung[11]. We get the sigma values from historical data from 2000 to 2012 and price a one year put from June 2012 to June 2013. We look at two strike prices, one that matures in the money and one that matures out of the money. The Bank of Korea, South Korea's central bank, has their current one week repo rate set at 2.5%, given their high credit rating, in the mid 'A' ratings, this is a reasonable risk free rate for our purpose.

In doing this we must remember that stocks do not really follow GBM, which our models all assume, as we examined in chapter 2. In practice we also have to model based on historical volatilities and correlations, which are not necessarily the same parameters that the stocks will follow for the life of the option, as regimes change over time. Another thing that comes up in practice is that the parameters can be reevaluated at every rebalancing time so that their values do change over time. They could use implied volatilites from other options traded on the same underlying assets or use a rolling window of historical prices that evolves for the life of the option. Either one would be an improvement to the closed form example that we present here, but the amount of data required to do the former is inaccessible and the though the latter is possible it is beyond the scope of the theoretical model that we are exploring in this section.

It is very important to note that this section only includes one sample path. It is not indicative of regular behaviour of the methods as far as returns are concerned. Though the ADI methof often makes money by over hedging this has the same effect on the riskiness of the porfolio as underhedging: it is not a perfect hedge and there is more risk than the other methods. Higher returns have to be paid for by taking more risk. Compounding these issues is that an empirical stock has non-constant volatility and does not follow GBM.

7.1 In the Money

In figures 7.1 to 7.5 we see that the ADI numerical solution made money on the hedging strategy at maturity, whereas none of the other strategies did. The ADI strategy had very strange looking Delta values; but because we know that the deltas in the ADI solution change

Figure 7.1: KOSPI Hedging Profit Evolution from ADI Solution, $K = 1900$, $\sigma_A = 0.27$, $\sigma_B = 0.18$, $\rho = 0.1$
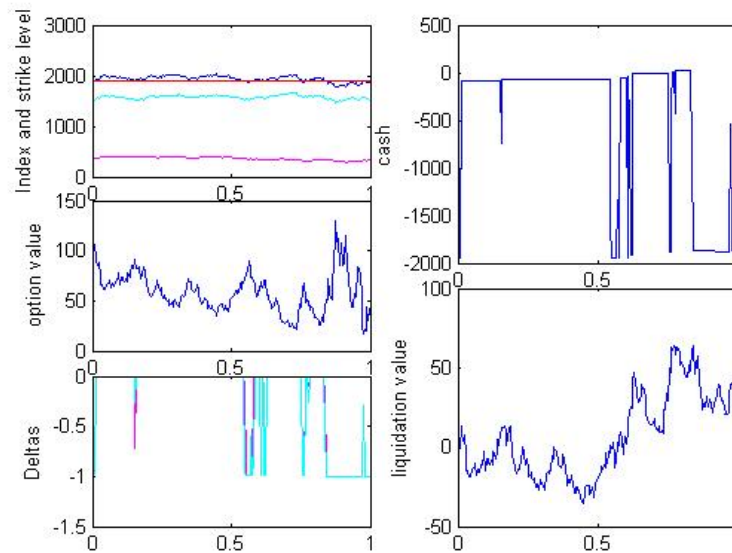
very rapidly for high sigma values near the strike price, we are not too surprised by the extreme jumps in the Delta values seen in figure 7.1. It is the extreme jumps that cause the profit, by buying low and selling high, but there is also a possibility of a large loss because of this extreme action; it was successful in this case though.

7.2   Out of the Money

In figures 7.6 to 7.10 we see that the ADI numerical solution again made money for the same reasons as before.

Figure 7.2: KOSPI Hedging Profit Evolution from Black-Scholes Solution, $K = 1900$

Figure 7.3: KOSPI Hedging Profit Evolution from "gamma" approximation, Δ_I only, $K = 1900$

Figure 7.4: KOSPI Hedging Profit Evolution from "gamma" approximation, both Deltas, $K =$ 1900

Figure 7.5: KOSPI Hedging Profit Evolution from Monte Carlo Simulation, $K = 1900$

Figure 7.6: KOSPI Hedging Profit Evolution from ADI Solution, $K = 1800$
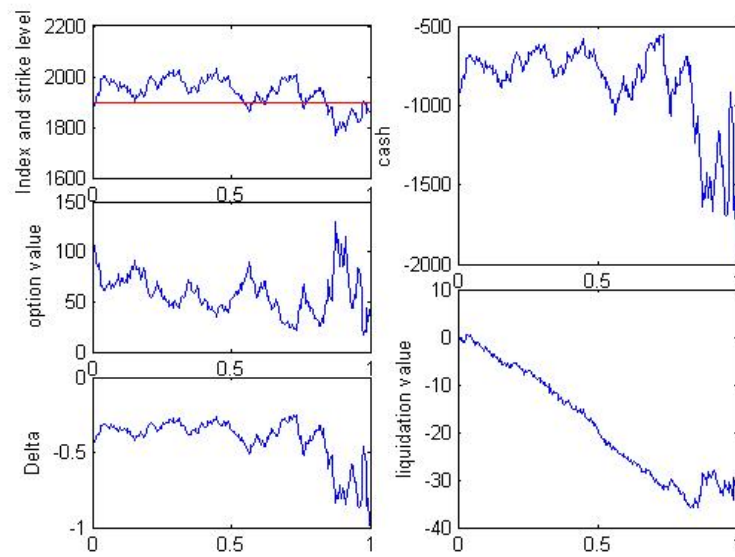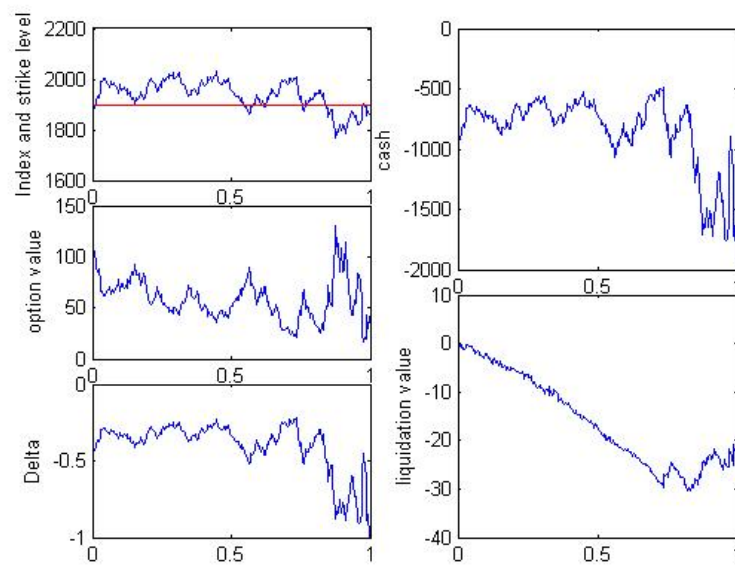

Figure 7.7: KOSPI Hedging Profit Evolution from Black-Scholes Solution, $K = 1800$
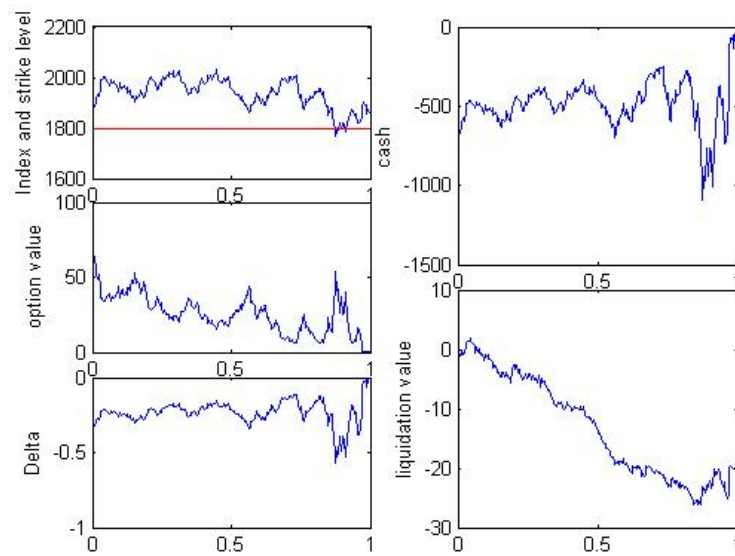
Figure 7.8: KOSPI Hedging Profit Evolution from "gamma" approximation, Δ_I only, $K = 1800$

Figure 7.9: KOSPI Hedging Profit Evolution from "gamma" approximation, both Deltas, $K = 1800$
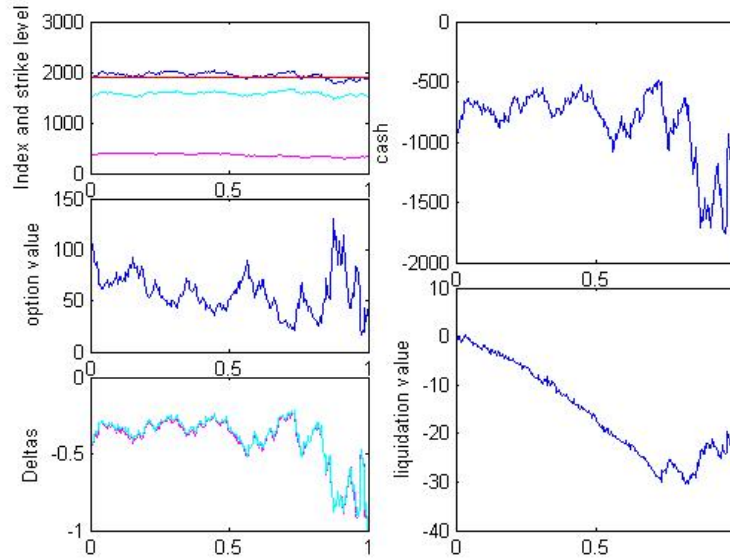
Figure 7.10: KOSPI Hedging Profit Evolution from Monte Carlo Simulation, $K = 1800$
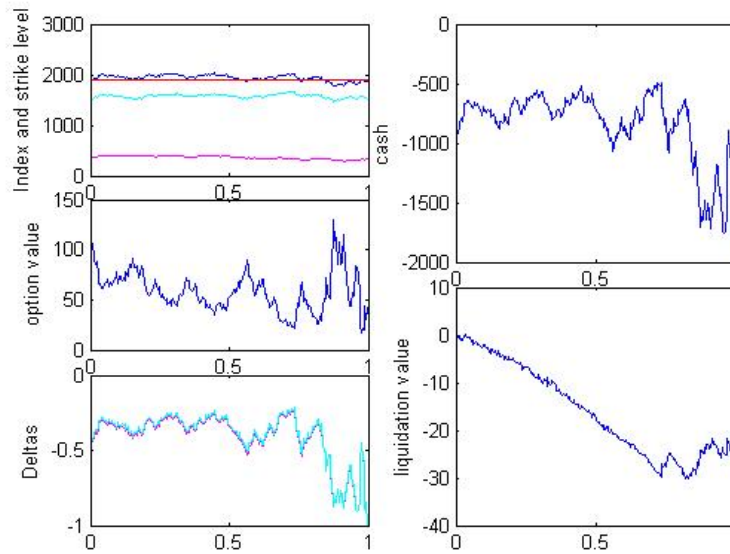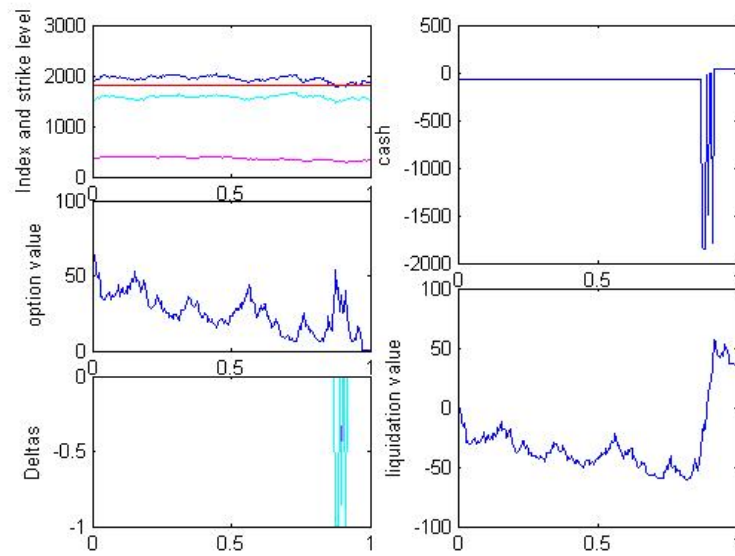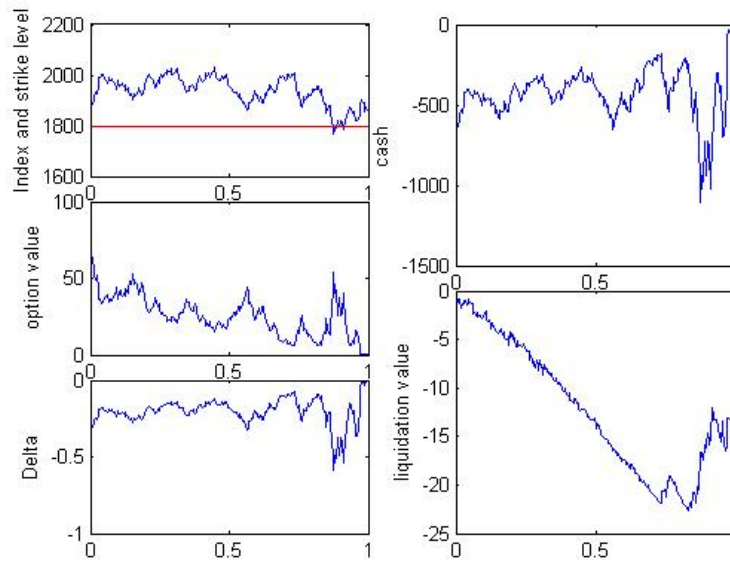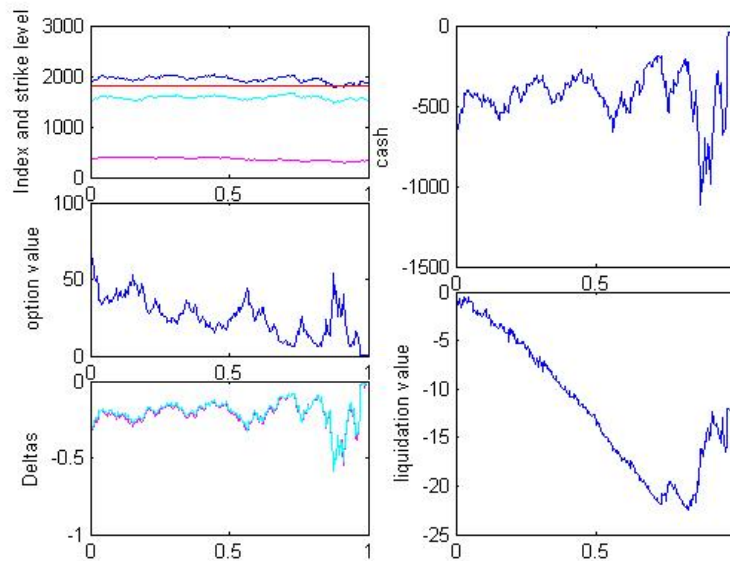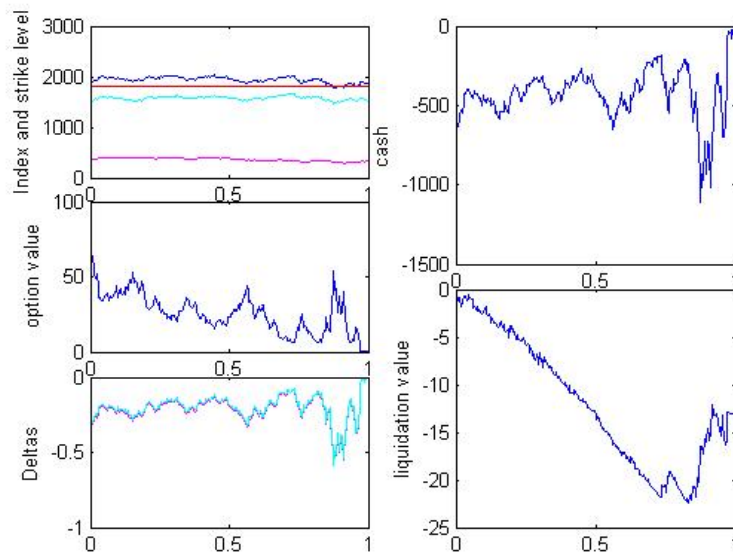
Chapter 8

Conclusion

8.1 Summary

In this thesis we began by discussing the possibility and recent occurences of an index containing a stock that grows relative to the index to the point of dominating the index. We then moved on to examine how individual stocks combine into an index; we saw that many positively correlated, evenly weighted stocks very conveniently combine to create an index that follows the same distribution as the individual stocks, GBM. This was also the case for an empirical example where the assets were fat tailed.

After being introduced to dominated indices we next moved on to develop PDEs that define the value of an option written on such an index. We considered the one and two-dimensional Black-Scholes PDEs as well as a one-dimensional approximation to the two-dimensional PDE which has a parameter that varies two-dimensionally. The one-dimensional PDEs have analytic solutions but we moved on to define two numerical solutions to the two-dimensional PDE. The methods used to numerically solve the two-dimensional PDE are an Alternating Direction Implicit method, a finite difference scheme, and a Monte Carlo method. We also split the one-dimensional approximation to the two-dimensional PDE into two cases; hedging with the index as a whole, as the model was intended, and forcing two separate Delta values from the pricing surface to hedge with both the dominating asset and the rest of the index.

Once these models were all defined we compared the three two-dimensional models' pricing and Delta surfaces and we compared the distribution of the hedging profits resulting from all five models. We finished with a simple empirical example to see how our methods behave with real data.

8.2 Conclusions

From our comparisons of pricing and Delta surfaces, hedging profit distributions, and hedging profit evolution sample paths we can draw come conclusions about each model.

ADI

- Unfortunately the ADI method, being a finite difference method, deals poorly with the non-smooth initial condition that comes with option pricing. Although near the center

90

of the pricing surface, near $\gamma = 0.5$, the discrepancies are prominant; closer to less dominance, lower γ values, the discrepancies are less pronounced. Conveniently this is the more likely case empirically.

- The ADI method does include the curvature of the solution, which the one-dimensional solutions lack.

- Even though this method requires multiple matrix operations at every time step until maturity since the matricies are particularly sparse, tri-diagonal, these operations can be implemented very efficiently; in linear time for every point on the grid. Given that it produces a smooth surface, allowing interpolation, a balance between spatial step size and interpolation coupled with efficient implementation can produce a surprisingly fast solution.

Monte Carlo

- The Monte Carlo solution does not over or under price persistently anywhere and neither are its Deltas systematically incorrect, as long as they are calculated independently of the pricing surface. This is why this method is used as a benchmark.

- However this method is not perfect; the pricing surface is unsmooth at all grid points. This is why the Delta values must be evaluated independently. The more sample paths used in the simulation the smoother the surface will be, but the slower the result is produced.

"gamma" Approximation

- This method has an analytic solution, so speed is not an issue.

- However it loses curvature compared to the Monte Carlo solution, but on a very small scale.

Black-Scholes

- Of course this method also has an analytic solution but it ignores any dominance of a constituent asset and therefore has no curvature across different levels of dominance.

- In the dominant asset case, the "gamma" apprixomation is strictly preferred to this solution.

In conclusion, if a constituent stock in an index dominates, or is unevenly weighted compared to the other constituent stocks, then attention should be paid to how to best price and hedge options on the index. It is not entirely clear which method is best to use, but it is clear that naïve approaches are not sufficient. Monte Carlo is consistently a strong candidate method but it can be time consuming compared to the analytic solutions from Black-Scholes or our "gamma" approximation. The ADI numerical solution does not react well to the non-smooth payoff function but it does often out perform the other methods. However it also has a much higher risk of a much larger loss than the other models.

8.3   Future Work

Here are some extensions of this work that could be pursued in the future:

- The theoretical results could be redone for a fatter tailed distribution than the normal distribution. This would make the results more realistic as most stocks and indices have fat tails although all the moments would be less analytically tractable, if they exist. Also, even the one-dimensional PDEs may not have analytic solutions.

- The empirical testing could be extended to choose parameters dynamically. Optimizing this choice depending on market volatility indicators could be included.

- The complete distribution of hedging profits for a single period on a single underlying asset driven by a normal random variable are distributed $\lambda(\chi^2(1) - 1)$ where λ is a deterministic function of the underlying asset at the beginning of the time interval and the option's second spatial derivative, Gamma[10]. Boyle and Emanuel extended this for multiple time steps, but it could be further extended for two underlying assets, or for an underlying asset driven by a fat tailed distribution, or both of these; these could be extended to multiple time steps.

- It would be very nice to be able to test the current work, or any further extensions, on more diverse empirical data. It is difficult and/or costly to get index weighting data, especially because the most common scheduled rebalancing period is only quarterly, so there are only small windows for which the number of shares in an index are constant.

- The Monte Carlo method used in this thesis was just a crude method; to compare the computation time in an empirical example variance reduction techniques could be employed.

- Another possible extension would be to incorporate rebalancing into the empirical algorithm.

- Improve a finite difference scheme so that it handles non-smooth initial conditions perfectly for two-dimensional heat equations.

Bibliography

[1] Oksendal, Bernt K. (2002), Stochastic Differential Equations: An Introduction with Applications, Springer, ISBN 3-540-63720-6.

[2] Robert V. Hogg and Elliot A. Tanis, Probability and Statistical Inference, 8th edition, Prentice Hall.

[3] iShares by BlackRock, `http://ca.ishares.com/product_info/fund/overview/XEG.htm`, Date accessed: July 25th, 2012.

[4] Yahoo Finance, `http://ca.finance.yahoo.com`, Date accessed: October 17th, 2012.

[5] Hull, John C., Options, Futures, and Other Derivatives (2011), Eighth Edition, Pearson, ISBN: 0-13-216484-9.

[6] Strikwerda, John C., Finite Difference Schemes and Partial Differential Equations (2004), Second Edition, Siam, ISBN: 0-89871-567-9

[7] Peaceman, D. W.; Rachford Jr., H. H. (1955), "The numerical solution of parabolic and elliptic differential equations", Journal of the Society for Industrial and Applied Mathematics 3 (1): 2841

[8] Crank, J.; Nicolson, P. (1947). "A practical method for numerical evaluation of solutions of partial differential equations of the heat conduction type". Proc. Camb. Phil. Soc. 43 (1): 5067

[9] Thomas, L.H. (1949), Elliptic Problems in Linear Differential Equations over a Network, Watson Sci. Comput. Lab Report, Columbia University, New York.

[10] Boyle, P. P. and D. Emanuel, "Discretely Adjusted Option Hedges," Journal of Financial Economics, 8 (September 1980), pp. 259-282.

[11] Bloomberg L.P., "KOSPI Index", Date accessed: June 17th, 2013.

Appendix A

Distributions

A.1 Fat tails

Consider a uniform random variable $X \sim \mathcal{U}(-a, a)$, a double exponenial or Laplacian random variable $Y \sim \mathcal{L}(\mu, \lambda)$, and a normal random variable $Z \sim \mathcal{N}(0, 1)$; three different distributions with different tails. Find a and λ such that they all have mean of zero and variance of one:

$$\mathbf{E}[X] = 0 = \frac{1}{2}(-a + a) = 0$$
$$Var(X) = 1 = \frac{1}{12}(a - (-a))^2$$
$$Var(X) = 1 = \frac{4}{12}a^2 \tag{A.1}$$
$$\Rightarrow a = \sqrt{3}$$

$$\mathbf{E}[Y] = 0 = \mu = 0$$
$$\Rightarrow \mu = 0$$
$$Var(Y) = 1 = 2\frac{1}{\lambda}^2 \tag{A.2}$$
$$\Rightarrow \lambda = \sqrt{2}$$

So we will look at the higher moments of the following distributions to compare the fatness of their tails:

$$\begin{cases} X \sim \mathcal{U}(-\sqrt{3}, \sqrt{3}) \\ Y \sim \mathcal{L}(0, \sqrt{2}) \\ Z \sim \mathcal{N}(0, 1) \end{cases} \tag{A.3}$$

Since all three are symmetric distributions the third moment, skewness, is zero for all three. So now we look at the fourth moment, Kurtosis:

$$K(V) = \frac{m_4}{m_2^2} = \frac{\mathbf{E}[V^4]}{\mathbf{E}[V^2]^2} \tag{A.4}$$

Figure A.1: Standard Lognormal Probability Density Function
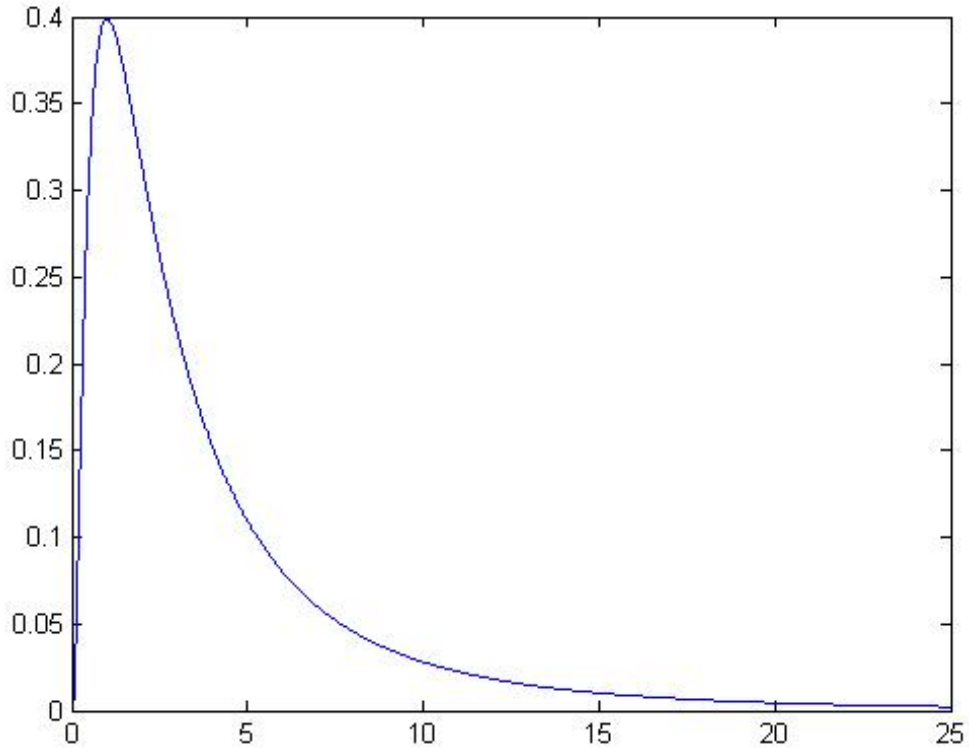
The higher the kurtosis the fatter the tails:

$$\begin{cases} K(X) = 1.8 \\ K(Y) = 6 \\ K(Z) = 3 \end{cases} \tag{A.5}$$

So compared to Z, the normal distribution, X, uniform, has thinner tails and Y, Laplacian, has fatter tails.

A.2 Lognormal Distribution

Here we consider a lognormally distributed random variable, $Y \sim \mathcal{LN}(a, b)$, that is $Y = e^{a+bZ}$, where $Z \sim \mathcal{N}(0, 1)$. See its probability density function in figure A.1.

A.3 Reading QQ Plots

A QQ plot, or a Quantile-Quantile plot, is a scatter plot of all the data points sampled from two distributions, say A and B, put in order and plotted against each other[2]. On the x-axis we have

Figure A.2: The uniform distribution against the normal distribution
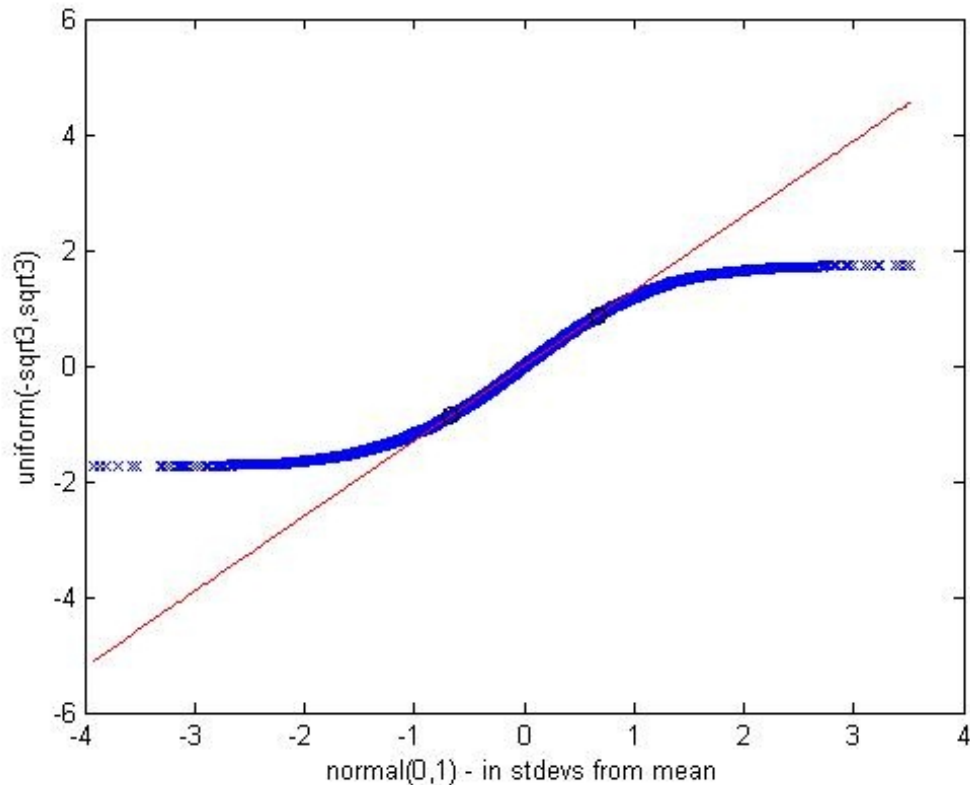
A and on the y-axis we have B. There is also a strait line plotted connecting the data point at the first quantile pair and the third quantile pair, hence the name Quantile-Quantile plot. We compare where the data lies compared to this line to decide which distribution has fatter or thinner tails than the other.

We use QQ plots to compare the tails of two distributions so here we will explain how to read a QQ plot with five examples. Figure A.2 tells us that the normal distribution has fatter tails than the uniform distribution, or that the uniform distribution has thinner tails than the normal distribution.

Figure A.3 tells us that the normal distribution has thinner tails than the Laplacian distribution, or that the Laplacian distribution has fatter tails than the normal distribution. Here one might notice that beyond three standard deviations the pattern changes a little; this is fine. We used a sample size of only 10,000 so events beyond three standard deviations are extremely rare and we really only need to study the shape within three standard deviations.

Figure A.4 tells us that the normal distribution has a thinner right tail and fatter left tail than the lognormal distribution, or that the lognormal distribution has a fatter right tail and thinner left tail than the normal distribution.

Figure A.5 tells us that the normal distribution has a fatter right tail and thinner left tail than the negative lognormal distribution, or that the negative lognormal distribution has a thinner right tail and fatter left tail than the normal distribution.

Figure A.6 tells us that the $\mathcal{N}(0, 1)$ distribution has the same tails as the $\mathcal{N}(3, 6)$ distribution,

Figure A.3: Laplacian distribution against the normal distribution
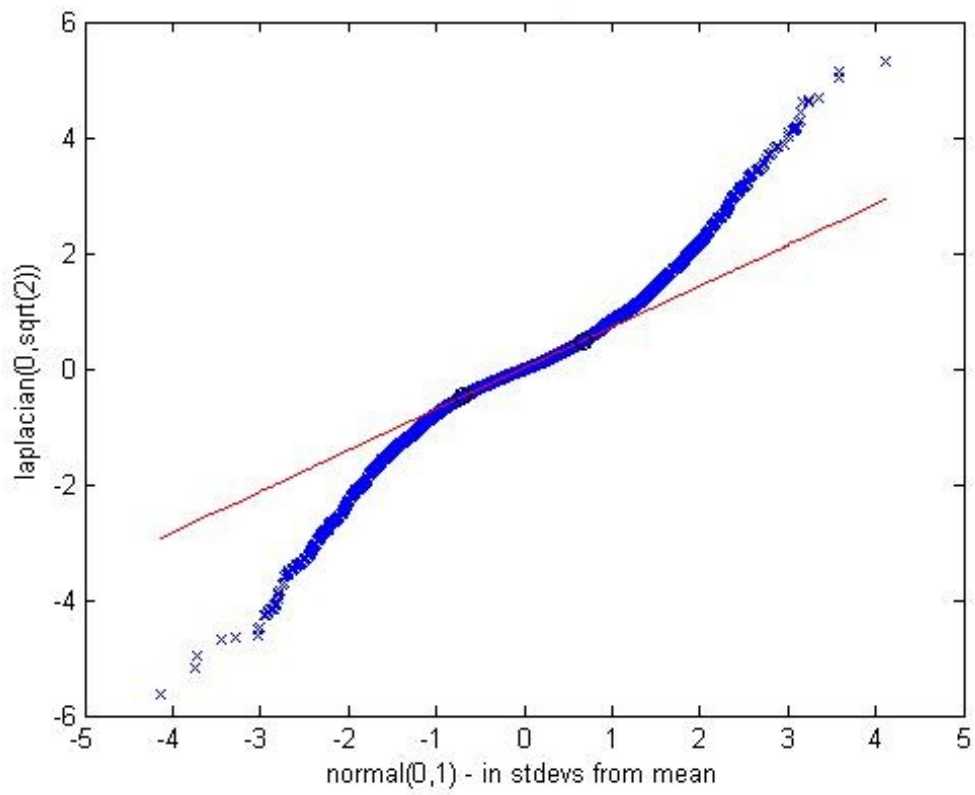
Figure A.4: Lognormal distribution against the normal distribution
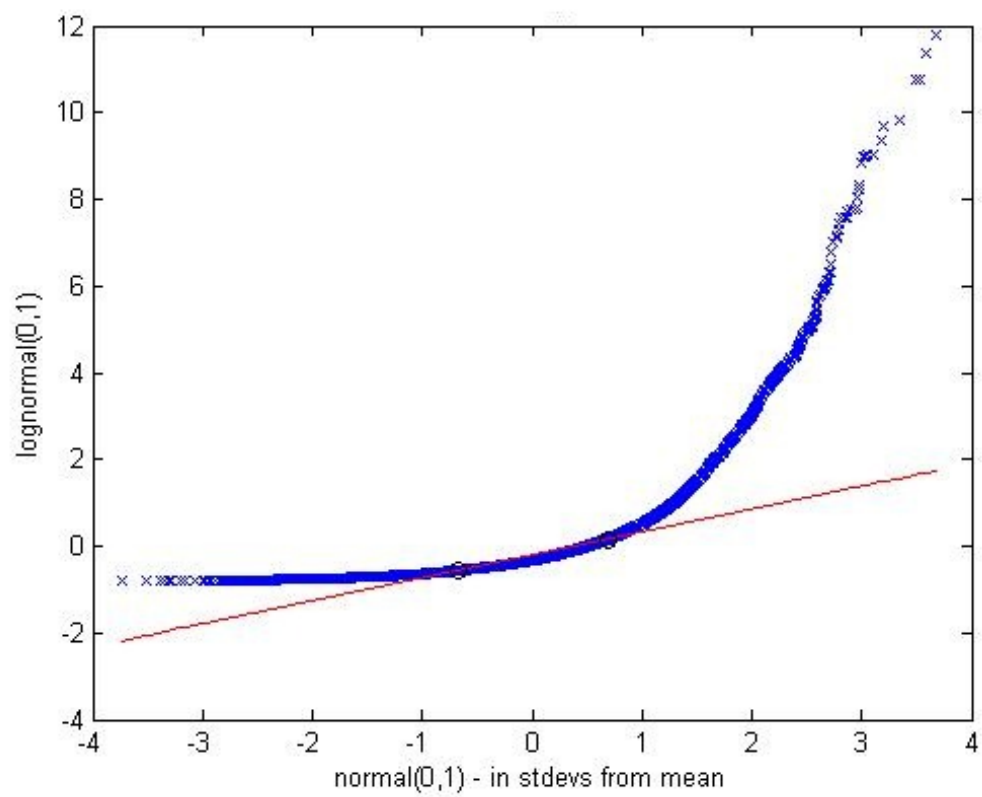
Figure A.5: Negative lognormal distribution against the normal distribution
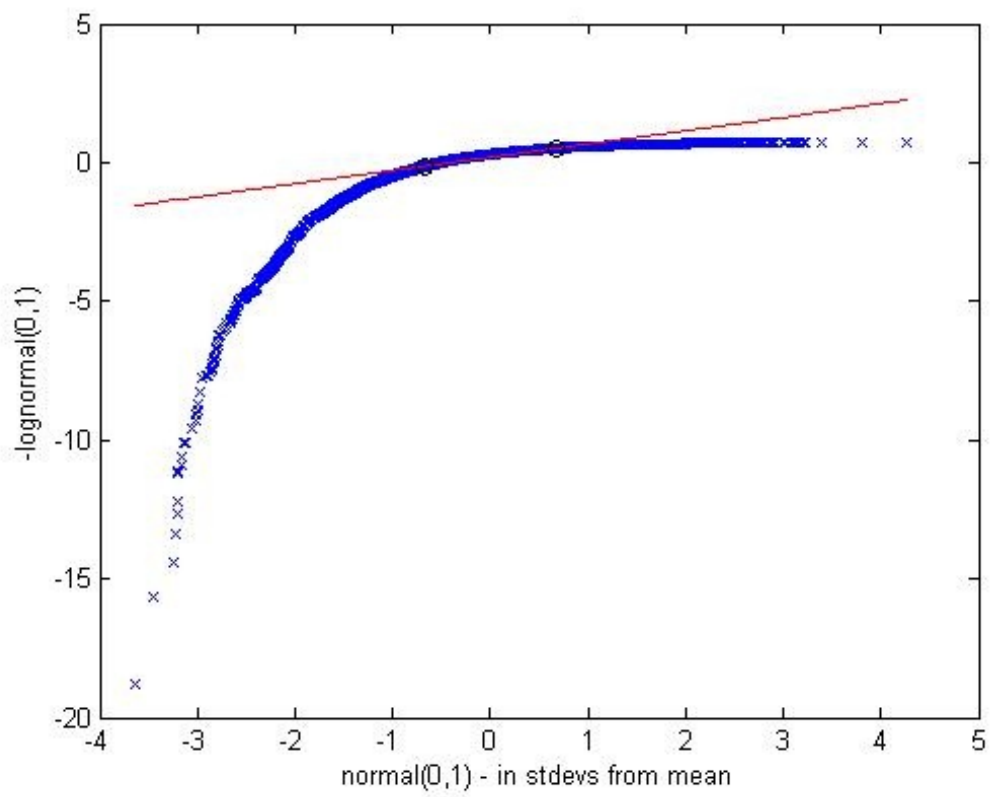
Figure A.6: Non-standard normal distribution against the standard normal distribution
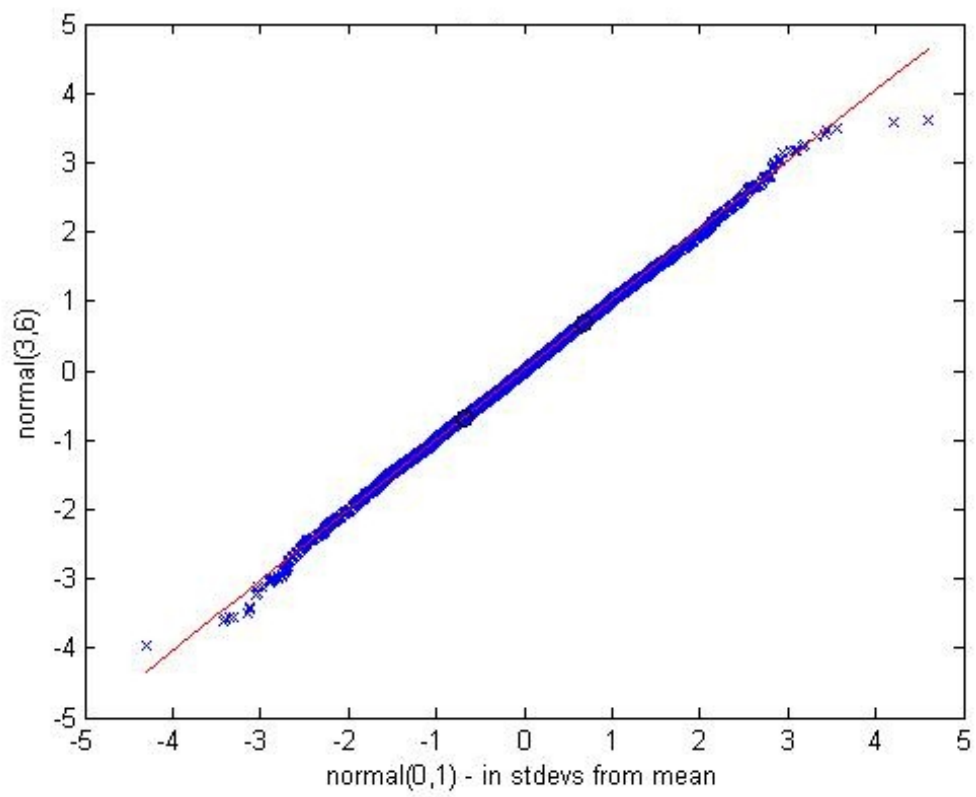
or that the normal distribution has tails that are invariant to linear transformations.

There is also the possibility that on one side two distributions have the same tails and on the other side one is fatter or thinner than the other. Or the case where on one side the difference in the tails is extreme and on the other side it is much more subtle.

A.4 Convolution

If $X, Y \sim f(x), g(y)$ respectively then $X + Y \sim (f * g)(z)$ where

$$(f * g)(z) = \int_{-\infty}^{\infty} f(z - y)g(y)dy \tag{A.6}$$

or

$$(f * g)(z) = \int_{-\infty}^{\infty} f(x)g(z - x)dx \tag{A.7}$$

If $X, Y \sim \mathcal{LN}(0, 1)$ then

$$f(v) = g(v) = \frac{1}{v\sqrt{2\pi}} e^{-\frac{1}{2}(\ln v)^2} \tag{A.8}$$

So $X + Y$ is distributed as follows:

$$(f * f)(z) = \int_{-\infty}^{\infty} \frac{1}{y(z - y)2\pi} e^{-\frac{1}{2}(\ln y)^2 - \frac{1}{2}(\ln(z-y))^2} dy \tag{A.9}$$

This is evaluated numerically, in matlab use quadgk().

Appendix B

Moments

B.1 Moments of a Single Lognormal Random Variable

Here we consider a lognormally distributed random variable, $Y \sim \mathcal{LN}(a, b)$, that is $Y = e^{a+bZ}$, where $Z \sim \mathcal{N}(0, 1)$, and compute its first four moments.

B.1.1 Mean

$$
\begin{aligned}
\mathbf{E}[Y] &= \mathbf{E}[e^{a+bZ}] \\
&= e^a \mathbf{E}[e^{bZ}] \\
&= e^a \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{\frac{-z^2}{2}} e^{bz} dz \\
&= e^a \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{\frac{-(z^2-2b)}{2}} dz \\
&= e^a \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{\frac{-(z^2-2b+b^2)}{2}+\frac{b^2}{2}} dz \\
&= e^{a+\frac{b^2}{2}} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{\frac{-(z-b)^2}{2}} dz \\
\Rightarrow \mathbf{E}[Y] &= e^{a+\frac{b^2}{2}}
\end{aligned}
\tag{B.1}
$$

B.1.2 Variance

$$
\begin{aligned}
var(Y) &= \mathbf{E}[(Y - \mathbf{E}[Y])^2] \\
&= \mathbf{E}[(e^{a+bZ} - e^{a+\frac{b^2}{2}})^2] \\
&= e^{2a} \mathbf{E}[e^{2bZ} - 2e^{\frac{b^2}{2}} e^{bZ} + e^{b^2}] \\
&= e^{2a}(\mathbf{E}[e^{2bZ}] - 2e^{\frac{b^2}{2}} \mathbf{E}[e^{bZ}] + e^{b^2}) \\
&= e^{2a}(e^{2b^2} - 2e^{b^2} + e^{b^2}) \\
\Rightarrow var(Y) &= e^{2a+b^2}(e^{b^2} - 1)
\end{aligned}
\tag{B.2}
$$

B.1.3 Skewness

$$
skew(Y) = \frac{\mathbf{E}[(Y - \mathbf{E}[Y])^3]}{\mathbf{E}[(Y - \mathbf{E}[Y])^2]^{\frac{3}{2}}} = \frac{m_3(Y)}{var(Y)^{\frac{3}{2}}}
\tag{B.3}
$$

where

$$
\begin{aligned}
m_3(Y) &= \mathbf{E}[(Y - \mathbf{E}[Y])^3] \\
&= \mathbf{E}[(e^{a+bZ} - e^{a+\frac{b^2}{2}})^3] \\
&= e^{3a}\mathbf{E}[e^{3bZ} - 3e^{\frac{b^2}{2}}e^{2bZ} + 3e^{b^2}e^{bZ} - e^{\frac{3b^2}{2}}] \\
&= e^{3a}(\mathbf{E}[e^{3bZ}] - 3e^{\frac{b^2}{2}}\mathbf{E}[e^{2bZ}] + 3e^{b^2}\mathbf{E}[e^{bZ}] - e^{\frac{3b^2}{2}}) \\
&= e^{3a}(e^{\frac{9b^2}{2}} - 3e^{\frac{5b^2}{2}} + 3e^{\frac{3b^2}{2}} - e^{\frac{3b^2}{2}}) \\
m_3(Y) &= e^{3a+\frac{3b^2}{2}}(e^{3b^2} - e^{b^2} + 2)
\end{aligned}
$$

(B.4)

So now we can calculate:

$$
\begin{aligned}
skew(Y) &= \frac{m_3(Y)}{var(Y)^{\frac{3}{2}}} \\
&= \frac{e^{3a+\frac{3b^2}{2}}(e^{3b^2}-e^{b^2}+2)}{[e^{2a+b^2}(e^{b^2}-1)]^{\frac{3}{2}}} \\
&= \frac{e^{3a+\frac{3b^2}{2}}(e^{3b^2}-e^{b^2}+2)}{e^{3a+\frac{3b^2}{2}}\sqrt{(e^{b^2}-1)^3}} \\
&= \frac{(e^{b^2}+2)(e^{b^2}-1)^2}{\sqrt{(e^{b^2}-1)^3}} \\
\Rightarrow skew(Y) &= (e^{b^2}+2)\sqrt{e^{b^2}-1}
\end{aligned}
$$

(B.5)

The final factorization of the numerator can easily, though tediously, be found by polynomial division, with the substitution $x = e^{b^2}$ divide by $x - 1$ twice.

B.1.4 Kurtosis

$$
kurt(Y) = \frac{\mathbf{E}[(Y - \mathbf{E}[Y])^4]}{\mathbf{E}[(Y - \mathbf{E}[Y])^2]^2} = \frac{m_4(Y)}{var(Y)^2}
$$

(B.6)

where

$$
\begin{aligned}
m_4(Y) &= \mathbf{E}[(Y - \mathbf{E}[Y])^4] \\
&= \mathbf{E}[(e^{a+bZ} - e^{a+\frac{b^2}{2}})^4] \\
&= e^{4a}\mathbf{E}[e^{4bZ} - 4e^{\frac{b^2}{2}}e^{3bZ} + 6e^{b^2}e^{2bZ} - 4e^{\frac{3b^2}{2}}e^{bZ} + e^{2b^2}] \\
&= e^{4a}(\mathbf{E}[e^{4bZ}] - 4e^{\frac{b^2}{2}}\mathbf{E}[e^{3bZ}] + 6e^{b^2}\mathbf{E}[e^{2bZ}] - 4e^{\frac{3b^2}{2}}\mathbf{E}[e^{bZ}] + e^{2b^2}) \\
&= e^{4a}(e^{8b^2} - 4e^{5b^2} + 6e^{3b^2} - 4e^{2b^2} + e^{2b^2}) \\
m_4(Y) &= e^{4a+2b^2}(e^{6b^2} - 4e^{3b^2} + 6e^{b^2} - 3)
\end{aligned}
$$

(B.7)

So now we can calculate:

$$
\begin{aligned}
kurt(Y) &= \frac{m_4(Y)}{var(Y)^2} \\
&= \frac{e^{4a+2b^2}(e^{6b^2}-4e^{3b^2}+6e^{b^2}-3)}{[e^{2a+b^2}(e^{b^2}-1)]^2} \\
&= \frac{e^{4a+2b^2}(e^{6b^2}-4e^{3b^2}+6e^{b^2}-3)}{e^{4a+2b^2}(e^{b^2}-1)^2} \\
&= \frac{(e^{4b^2}+2e^{3b^2}+3e^{2b^2}-3)(e^{b^2}-1)^2}{(e^{b^2}-1)^2} \\
\Rightarrow kurt(Y) &= e^{4b^2} + 2e^{3b^2} + 3e^{2b^2} - 3
\end{aligned}
$$

(B.8)

The final factorization of the numerator can easily be found by similar polynomial division as for skewness.

B.2   Moments of Sum of Two Identical Evenly Weighted Lognormal Random Variables

Here we compute the first four moments for a sum of two lognormal random variables that are correlated, have the same parameters, and are evenly weighted. So we let $Y_2 = \frac{1}{2}e^{c+dZ_1} + \frac{1}{2}e^{c+dZ_2}$ where $Z_1, Z_2 \sim \mathcal{N}(0, 1)$ are correlated by ρ.

B.2.1   Mean

$$
\begin{aligned}
\mathbf{E}[Y_2] &= \mathbf{E}[\tfrac{1}{2}e^{c+dZ_1} + \tfrac{1}{2}e^{c+dZ_2}] \\
\Rightarrow \mu_2 = \mathbf{E}[Y_2] &= e^{c+\frac{d^2}{2}}
\end{aligned}
\tag{B.9}
$$

B.2.2   Variance

$$
\begin{aligned}
var(Y_2) &= \mathbf{E}[(Y_2 - \mu_2)^2] \\
&= \mathbf{E}[Y_2^2 - 2\mu_2 Y_2 + \mu_2^2] \\
&= \mathbf{E}[Y_2^2] - 2\mu_2 \mathbf{E}[Y_2] + \mu_2^2 \\
\Rightarrow \sigma_2^2 = var(Y_2) &= \mathbf{E}[Y_2^2] - \mu_2^2
\end{aligned}
\tag{B.10}
$$

And using the normal moment generating function again we can find $\mathbf{E}[Y_2^2]$:

$$
\begin{aligned}
\mathbf{E}[Y_2^2] &= \mathbf{E}[(\tfrac{1}{2}e^{c+dZ_1} + \tfrac{1}{2}e^{c+dZ_2})^2] \\
&= \tfrac{1}{4}e^{2c}\mathbf{E}[e^{2dZ_1} + 2e^{dZ_1+dZ_2} + e^{2dZ_2}] \\
&= \tfrac{1}{4}e^{2c}\mathbf{E}[e^{2dZ_1}] + \tfrac{1}{2}e^{2c}\mathbf{E}[e^{dZ_1+dZ_2}] + \tfrac{1}{4}e^{2c}\mathbf{E}[e^{2dZ_2}]
\end{aligned}
\tag{B.11}
$$

And where $i = 1, 2$ we have

$$
\begin{aligned}
\tfrac{1}{4}e^{2c}\mathbf{E}[e^{2dZ_i}] &= \tfrac{1}{4}e^{2c+2d^2} \\
\tfrac{1}{2}e^{2c}\mathbf{E}[e^{dZ_1+dZ_2}] &= \tfrac{1}{2}e^{2c}e^{d^2(1+\rho)}
\end{aligned}
\tag{B.12}
$$

$$
\Rightarrow \sigma_2^2 = \frac{e^{2c+d^2}}{2}(e^{d^2} + e^{d^2\rho} - 2)
\tag{B.13}
$$

B.2.3   Skewness

We know that $skew(Y_2) = \frac{m_3(Y_2)}{\sigma_{Y_2}^3}$ and we will now rearrange m_3 as we did for σ^2:

$$
\begin{aligned}
m_3(X) &= \mathbf{E}[(X - \mu_X)^3] \\
&= \mathbf{E}[X^3 - 3\mu_X X^2 + 3\mu_X^2 X - \mu_X^3] \\
&= \mathbf{E}[X^3] - 3\mu_X[X^2] + 3\mu_X^2 \mathbf{E}[X] - \mu_X^3 \\
&= \mathbf{E}[X^3] - 3\mu_X(\sigma_X^2 + \mu_X^2) + 3\mu_X^2\mu_X - \mu_X^3 \\
\Rightarrow m_3(X) &= \mathbf{E}[X^3] - 3\mu_X\sigma_X^2 - \mu_X^3
\end{aligned}
\tag{B.14}
$$

From here the only term left to solve for is $\mathbf{E}[Y_2^3]$ as mean and variance have already been solved for. From the multivariate normal moment generating function we have :

$$
\begin{aligned}
\mathbf{E}[Y_2^3] &= \mathbf{E}[(\tfrac{1}{2}e^{c+dZ_1} + \tfrac{1}{2}e^{c+dZ_2})^3] \\
&= \tfrac{1}{8}e^{3c}\mathbf{E}[e^{3dZ_1} + 3e^{2dZ_1+dZ_2} + 3e^{dZ_1+2dZ_2} + e^{3dZ_2}] \\
&= \tfrac{1}{8}e^{3c}\mathbf{E}[e^{3dZ_1}] + \tfrac{3}{8}e^{3c}\mathbf{E}[e^{2dZ_1+dZ_2}] + \tfrac{3}{8}e^{3c}\mathbf{E}[e^{dZ_1+2dZ_2}] + \tfrac{1}{8}e^{3c}\mathbf{E}[e^{3dZ_2}]
\end{aligned}
\tag{B.15}
$$

And where $i, j = 1, 2$ we have

$$
\begin{aligned}
\tfrac{1}{8}e^{3c}\mathbf{E}[e^{3dZ_i}] &= \tfrac{1}{8}e^{3c+\frac{9}{2}d^2} \\
\tfrac{3}{8}e^{3c}\mathbf{E}[e^{dZ_i+dZ_j}] &= \tfrac{3}{8}e^{3c}e^{\frac{d^2}{2}(5+4\rho)}
\end{aligned}
\tag{B.16}
$$

B.2.4 Kurtosis

We find kurtosis from $kurt(Y_2) = \frac{m_4(Y_2)}{\sigma_{Y_2}^4}$, and:

$$
\begin{aligned}
m_4(Y_2) &= \mathbf{E}[(Y_2 - \mu_2)^4] \\
&= \mathbf{E}[Y_2^4 - 4\mu_2 Y_2^3 + 6\mu_2^2 Y_2^2 - 4\mu_2^3 Y_2 + \mu_2^4] \\
&= \mathbf{E}[Y_2^4] - 4\mu_2\mathbf{E}[Y_2^3] + 6\mu_2^2\mathbf{E}[Y_2^2] - 4\mu_2^3\mathbf{E}[Y_2] + \mu_2^4 \\
&= \mathbf{E}[Y_2^4] - 4\mu_2\mathbf{E}[Y_2^3] + 6\mu_2^2(\sigma_2^2 + \mu_2^2) - 4\mu_2^3\mu_2 + \mu_2^4 \\
\Rightarrow m_4(Y_2) &= \mathbf{E}[Y_2^4] - 4\mu_2\mathbf{E}[Y_2^3] + 6\mu_2^2\sigma_2^2 + 3\mu_2^4
\end{aligned}
\tag{B.17}
$$

From here the only term left to solve for is $\mathbf{E}[Y_2^4]$:

$$
\begin{aligned}
\mathbf{E}[Y_2^4] &= \mathbf{E}[(\tfrac{1}{2}e^{c+dZ_1} + \tfrac{1}{2}e^{c+dZ_2})^4] \\
&= \tfrac{1}{16}e^{4c}\mathbf{E}[e^{4dZ_1} + 4e^{3dZ_1+dZ_2} + 6e^{2dZ_1+2dZ_2} + 4e^{dZ_1+3dZ_2} + e^{4dZ_2}] \\
&= \tfrac{1}{16}e^{4c}\mathbf{E}[e^{4dZ_1}] + \tfrac{1}{4}e^{4c}\mathbf{E}[e^{3dZ_1+dZ_2}] + \tfrac{3}{8}e^{4c}\mathbf{E}[e^{dZ_1+2dZ_2}] \\
&\quad + \tfrac{1}{4}e^{4c}\mathbf{E}[e^{dZ_1+3dZ_2}] + \tfrac{1}{16}e^{4c}\mathbf{E}[e^{3dZ_2}]
\end{aligned}
\tag{B.18}
$$

And where $i, j = 1, 2$ we have

$$
\begin{aligned}
\tfrac{1}{16}e^{4c}\mathbf{E}[e^{4dZ_i}] &= \tfrac{1}{16}e^{4c+8d^2} \\
\tfrac{1}{4}e^{4c}\mathbf{E}[e^{3dZ_i+dZ_j}] &= \tfrac{1}{4}e^{4c}e^{d^2(5+3\rho)} \\
\tfrac{3}{8}e^{4c}\mathbf{E}[e^{23dZ_i+2dZ_j}] &= \tfrac{3}{8}e^{4c}e^{4d^2(1+\rho)}
\end{aligned}
\tag{B.19}
$$

B.3 Moments of Sum of Two General Lognormal Random Variables

Here we compute the first four moments for a sum of two lognormal random variables that are correlated, have different parameters, and are not evenly weighted. So we let $Y_2 = \alpha e^{c+dZ_1} + (1-\alpha)e^{p+qZ_2}$ where $Z_1, Z_2 \sim \mathcal{N}(0, 1)$ are correlated by ρ.

B.3.1 Mean

$$
\begin{aligned}
\mathbf{E}[Y_2] &= \mathbf{E}[\alpha e^{c+dZ_1} + (1-\alpha)e^{p+qZ_2}] \\
\Rightarrow \mu_2 = \mathbf{E}[Y_2] &= \alpha e^{c+\frac{d^2}{2}} + (1-\alpha)e^{p+q\frac{q^2}{2}}
\end{aligned}
\tag{B.20}
$$

B.3.2 Variance

$$
\begin{aligned}
var(Y_2) &= \mathbf{E}[(Y_2 - \mu_2)^2] \\
\Rightarrow \sigma_2^2 = var(Y_2) &= \mathbf{E}[Y_2^2] - \mu_2^2
\end{aligned} \tag{B.21}
$$

And using the normal moment generating function again we can find $\mathbf{E}[Y_2^2]$. Let $D = \alpha e^{c+dZ_1}$ and $Q = (1 - \alpha)e^{p+qZ_1}$

$$
\begin{aligned}
\mathbf{E}[Y_2^2] &= \mathbf{E}[(D + Q)^2] \\
&= \mathbf{E}[D^2 + 2DQ + Q^2] \\
&= \mathbf{E}[D^2] + 2\mathbf{E}[DQ] + \mathbf{E}[Q^2]
\end{aligned} \tag{B.22}
$$

And we have

$$
\begin{aligned}
\mathbf{E}[D^2] &= \mathbf{E}[(\alpha e^{c+dZ_1})^2] & \mathbf{E}[Q^2] &= \mathbf{E}[((1 - \alpha)e^{p+qZ_1})^2] \\
&= \alpha^2 e^{2c}\mathbf{E}[e^{2dZ_1}] & &= (1 - \alpha)^2 e^{2p}\mathbf{E}[e^{2qZ_1}] \\
&= \alpha^2 e^{2c+2d^2} & &= (1 - \alpha)^2 e^{2p+2q^2}
\end{aligned}
$$

$$
\begin{aligned}
\mathbf{E}[DQ] &= \mathbf{E}[\alpha e^{c+dZ_1}(1 - \alpha)e^{p+qZ_1}] \\
&= \alpha(1 - \alpha)e^{c+p}\mathbf{E}[e^{dZ_1+qZ_2}] \\
&= \alpha(1 - \alpha)e^{c+p}e^{\frac{1}{2}d^2+\rho dq+\frac{1}{2}q^2}
\end{aligned} \tag{B.23}
$$

$$
\Rightarrow \mathbf{E}[Y_2^2] = \alpha^2 e^{2c+2d^2} + 2\alpha(1 - \alpha)e^{c+p}e^{\frac{1}{2}d^2+\rho dq+\frac{1}{2}q^2} + (1 - \alpha)^2 e^{2p+2q^2} \tag{B.24}
$$

B.3.3 Skewness

We know that $skew(Y_2) = \frac{m_3(Y_2)}{\sigma_2^3}$ and that $m_3(Y_2) = \mathbf{E}[Y_2^3] - 3\mu_2\sigma_2^2 - \mu_2^3$ From here the only term left to solve for is $\mathbf{E}[Y_2^3]$

$$
\begin{aligned}
\mathbf{E}[Y_2^3] &= \mathbf{E}[(D + Q)^3] \\
&= \mathbf{E}[D^3 + 3D^2Q + 3DQ^2Q^3] \\
&= \mathbf{E}[D^3] + 3\mathbf{E}[D^2Q] + 3\mathbf{E}[DQ^2] + \mathbf{E}[Q^3]
\end{aligned} \tag{B.25}
$$

And we have

$$
\begin{aligned}
\mathbf{E}[D^3] &= \mathbf{E}[(\alpha e^{c+dZ_1})^3] & \mathbf{E}[Q^3] &= \mathbf{E}[((1 - \alpha)e^{p+qZ_1})^3] \\
&= \alpha^3 e^{3c}\mathbf{E}[e^{3dZ_1}] & &= (1 - \alpha)^3 e^{3p}\mathbf{E}[e^{3qZ_1}] \\
&= \alpha^3 e^{3c+\frac{9}{2}d^2} & &= (1 - \alpha)^3 e^{3p+\frac{9}{2}q^2}
\end{aligned}
$$

$$
\begin{aligned}
\mathbf{E}[D^2Q] &= \mathbf{E}[\alpha^2 e^{2c+2dZ_1}(1 - \alpha)e^{p+qZ_1}] \\
&= \alpha^2(1 - \alpha)e^{2c+p}\mathbf{E}[e^{2dZ_1+qZ_1}] \\
&= \alpha^2(1 - \alpha)e^{2c+p}e^{2d^2+2\rho dq+\frac{1}{2}q^2}
\end{aligned}
$$

$$
\begin{aligned}
\mathbf{E}[DQ^2] &= \mathbf{E}[\alpha e^{c+dZ_1}(1 - \alpha)^2 e^{2p+2qZ_1}] \\
&= \alpha(1 - \alpha)^2 e^{c+2p}\mathbf{E}[e^{dZ_1+2qZ_1}] \\
&= \alpha(1 - \alpha)^2 e^{c+2p}e^{\frac{1}{2}d^2+2\rho dq+2q^2}
\end{aligned} \tag{B.26}
$$

B.3.4 Kurtosis

We know that $kurt(Y_2) = \frac{m_4(Y_2)}{\sigma_2^4}$ and that $m_4(Y_2) = \mathbf{E}[Y_2^4] - 4\mu_2\mathbf{E}[Y_2^3] + 6\mu_2^2\sigma_2^2 + 3\mu_2^4$

From here the only term left to solve for is $\mathbf{E}[Y_2^4]$:

$$
\begin{aligned}
\mathbf{E}[Y_2^4] &= \mathbf{E}[(D+Q)^4] \\
&= \mathbf{E}[D^4 + 4D^3Q + 6D^2Q^2 + 4DQ^3 + Q^4] \\
&= \mathbf{E}[D^4] + 4\mathbf{E}[D^3Q] + 6\mathbf{E}[D^2Q^2] + +4\mathbf{E}[DQ^3] + \mathbf{E}[Q^4]
\end{aligned}
\tag{B.27}
$$

And we have

$$
\begin{aligned}
\mathbf{E}[D^4] &= \mathbf{E}[(\alpha e^{c+dZ_1})^4] & \mathbf{E}[Q^4] &= \mathbf{E}[((1-\alpha)e^{p+qZ_1})^4] \\
&= \alpha^4 e^{4c}\mathbf{E}[e^{4dZ_1}] & &= (1-\alpha)^4 e^{4p}\mathbf{E}[e^{4qZ_1}] \\
&= \alpha^4 e^{4c+8d^2} & &= (1-\alpha)^4 e^{4p+8q^2}
\end{aligned}
$$

$$
\begin{aligned}
\mathbf{E}[D^3Q] &= \mathbf{E}[\alpha^3 e^{3c+3dZ_1}(1-\alpha)e^{p+qZ_1}] \\
&= \alpha^3(1-\alpha)e^{3c+p}\mathbf{E}[e^{3dZ_1+qZ_1}] \\
&= \alpha^3(1-\alpha)e^{3c+p}e^{\frac{9}{2}d^2+3\rho dq+\frac{1}{2}q^2}
\end{aligned}
$$

$$
\begin{aligned}
\mathbf{E}[DQ^3] &= \mathbf{E}[\alpha e^{c+dZ_1}(1-\alpha)^3 e^{3p+3qZ_1}] \\
&= \alpha(1-\alpha)^3 e^{c+3p}\mathbf{E}[e^{dZ_1+3qZ_1}] \\
&= \alpha(1-\alpha)^3 e^{c+3p}e^{\frac{1}{2}d^2+3\rho dq+\frac{9}{2}q^2}
\end{aligned}
$$

$$
\begin{aligned}
\mathbf{E}[D^2Q^2] &= \mathbf{E}[\alpha^2 e^{2c+2dZ_1}(1-\alpha)^2 e^{2p+2qZ_1}] \\
&= \alpha^2(1-\alpha)^2 e^{2c+2p}\mathbf{E}[e^{dZ_1+2qZ_1}] \\
&= \alpha^2(1-\alpha)^2 e^{2c+2p}e^{2d^2+4\rho dq+2q^2}
\end{aligned}
\tag{B.28}
$$

B.4 Moments of Sum of n Identical Lognormal Random Variables

Here we compute the first four moments for a sum of n lognormal random variables that are all identically correlated, have the same parameters, and are evenly weighted. So we let $Y_n = \sum_{i=1}^{n} \frac{1}{n} e^{c+dZ_i}$. The mean does not change, but the higher moments do change with n.

B.4.1 Mean

$$
\begin{aligned}
\mathbf{E}[Y_n] &= \mathbf{E}[\frac{1}{n}\sum_{i=1}^{n} e^{c+dZ_i}] \\
&= \frac{1}{n}n\mathbf{E}[e^{c+dZ_i}] \\
&= e^{c+\frac{d^2}{2}}
\end{aligned}
\tag{B.29}
$$

The $\frac{1}{n}$ term gets cancelled by the n that arises from having a summation which has n terms with identical expectations. For later work we will let $\mu_{Y_n} = e^{c+\frac{d^2}{2}}$

B.4.2   Variance

Here we want to use the same problem solving as with mean: "How many terms identical in expectation of each type do we have?", but because variance is a central moment we first rewrite it as before:

$$var(Y_n) = \mathbf{E}[(Y_n - \mu_{Y_n})^2] = \mathbf{E}[Y_n^2] - \mu_{Y_n}^2 \tag{B.30}$$

Now we need to understand the terms that arise in Y_n^2

$$
\begin{aligned}
Y_n^2 &= \tfrac{1}{n^2}(\sum_{i=1}^n e^{c+dZ_i})^2 \\
&= \tfrac{1}{n^2}(\sum_{i=1}^n (e^{c+dZ_i})^2 + 2\sum_{i=1}^n \sum_{j=i+1}^n e^{c+dZ_i}e^{c+dZ_j})
\end{aligned}
\tag{B.31}
$$

There are n terms in the first summation and $\binom{n}{2}$ terms in the second, there are a few intuitions by which this can be realized. First, as it is written above, the nth triangle number ($n + (n - 1) + (n - 2) + ... + 2 + 1$) is given by $\binom{n}{2}$ as seen in Pascal's triangle. Second is that the second summations can be thought of as $\sum_{i \neq j} e^{c+dZ_i}e^{c+dZ_j}$ where every pair of two different i and j are selected and order does not matter, the number of resulting terms is the definition of 'n-choose-2' or $\binom{n}{2} = \frac{n!}{2!(n-2)!}$. And we know from previous work that the terms of each summation are identical in expectation:

$$
\begin{aligned}
\mathbf{E}[(e^{c+dZ_i})^2] &= \mathbf{E}[e^{2c+2dZ_i}] \\
&= e^{2c+2d^2} \\
\mathbf{E}[e^{c+dZ_i}e^{c+dZ_j}] &= e^{2c+d^2(1+\rho)}
\end{aligned}
\tag{B.32}
$$

So:

$$
\begin{aligned}
\mathbf{E}[Y_n^2] &= \tfrac{1}{n^2}(\sum_{i=1}^n \mathbf{E}[(e^{c+dZ_i})^2] + 2\sum_{i=1}^n \sum_{j=i+1}^n \mathbf{E}[e^{c+dZ_i}e^{c+dZ_j}]) \\
&= \tfrac{1}{n^2}(\sum_{i=1}^n e^{2c+2d^2} + 2\sum_{i=1}^n \sum_{j=i+1}^n e^{2c+d^2(1+\rho)}) \\
&= \tfrac{1}{n^2}(ne^{2c+2d^2} + 2\binom{n}{2}e^{2c+d^2(1+\rho)}) \\
\Rightarrow \mathbf{E}[Y_n^2] &= \tfrac{e^{2c+d^2}}{n^2}(ne^{d^2} + 2\binom{n}{2}e^{d^2\rho})
\end{aligned}
\tag{B.33}
$$

Finally we substitute this back into the variance:

$$
\begin{aligned}
var(Y_n) &= \tfrac{e^{2c+d^2}}{n^2}(ne^{d^2} + 2\binom{n}{2}e^{d^2\rho}) - e^{2c+d^2} \\
\Rightarrow var(Y_n) &= \tfrac{e^{2c+d^2}}{n^2}(ne^{d^2} + 2\binom{n}{2}e^{d^2\rho} - n^2) \\
<< OR >> var(Y_n) &= \tfrac{e^{2c+d^2}}{n}(e^{d^2} + (n - 1)e^{d^2\rho} - n)
\end{aligned}
\tag{B.34}
$$

We let $\sigma_{Y_n}^2 = var(Y_n)$ for simplicity later.

B.4.3   Skewness

For skewness we follow the same method as we did for variance. We start with $skew(Y_n) = \frac{m_3(Y_n)}{\sigma_{Y_n}^3}$ and we rewrite $m_3(Y_n)$ in a more useful form:

$$
\begin{aligned}
m_3(Y_n) &= \mathbf{E}[(Y_n - \mu_{Y_n})^3] \\
&= \mathbf{E}[Y_n^3] - 3\mu_{Y_n}\sigma_{Y_n}^2 - \mu_{Y_n}^3
\end{aligned}
\tag{B.35}
$$

Now we look at the terms that arise in Y_n^3

$$
\begin{aligned}
Y_n^3 &= \tfrac{1}{n^3}(\sum_{i=1}^n e^{c+dZ_i})^3 \\
&= \tfrac{1}{n^3}(\sum_{i=1}^n (e^{c+dZ_i})^3 + 3\sum_{i=1}^n \sum_{j=1, j\neq i}^n (e^{c+dZ_i})^2 e^{c+dZ_j} \\
&\quad + 6\sum_{i=1}^n \sum_{j=i+1}^n \sum_{k=j+1}^n e^{c+dZ_i}e^{c+dZ_j}e^{c+dZ_k})
\end{aligned}
\tag{B.36}
$$

There are n terms in the first summation, 'n-permute-2', $_nP_2$, terms in the second and 'n-choose-3', $\binom{n}{3}$ in the third. The intuition for the second summation is that it can be thought of as every ordered pair of two different i and j, the resulting number of terms is defined by 'n-permute-2' or $_nP_2 = \frac{n!}{(n-2)!}$. The intuition for the third summation is that every triple of three different i, j and k are selected and order does not matter, the number of resulting terms is the definition of 'n-choose-3' or $\binom{n}{3}$. And we know that the terms of each summation are identical in expectation and from the multivariate normal moment generating function they are:

$$
\begin{aligned}
\mathbf{E}[(e^{c+dZ_i})^3] &= e^{3c+\frac{9d^2}{2}} \\
\mathbf{E}[e^{2c+2dZ_i}e^{c+dZ_j}] &= e^{3c+\frac{d^2}{2}(5+4\rho)} \\
\mathbf{E}[e^{c+dZ_i}e^{c+dZ_j}e^{c+dZ_k}] &= e^{3c+\frac{d^2}{2}(3+6\rho)}
\end{aligned}
\tag{B.37}
$$

So:

$$
\begin{aligned}
\mathbf{E}[Y_n^3] &= \tfrac{1}{n^3}(\sum_{i=1}^n \mathbf{E}[(e^{c+dZ_i})^3] + 3\sum_{i=1}^n \sum_{j=1,j\neq i}^n \mathbf{E}[(e^{c+dZ_i})^2 e^{c+dZ_j}] \\
&\quad +6\sum_{i=1}^n \sum_{j=i+1}^n \sum_{k=j+1}^n \mathbf{E}[e^{c+dZ_i}e^{c+dZ_j}e^{c+dZ_k}]) \\
&= \tfrac{1}{n^3}(\sum_{i=1}^n e^{3c+\frac{9d^2}{2}} + 3\sum_{i=1}^n \sum_{j=1,j\neq i}^n e^{3c+\frac{d^2}{2}(5+4\rho)} \\
&\quad +6\sum_{i=1}^n \sum_{j=i+1}^n \sum_{k=j+1}^n e^{3c+\frac{d^2}{2}(3+6\rho)}) \\
&= \tfrac{1}{n^3}(ne^{3c+\frac{9d^2}{2}} + 3(_nP_2)e^{3c+\frac{d^2}{2}(5+4\rho)} + 6\binom{n}{3}e^{3c+\frac{d^2}{2}(3+6\rho)}) \\
\Rightarrow \mathbf{E}[Y_n^3] &= \tfrac{e^{3c+\frac{3d^2}{2}}}{n^3}(ne^{3d^2} + 3(_nP_2)e^{d^2(1+2\rho)} + 6\binom{n}{3}e^{3d^2\rho})
\end{aligned}
\tag{B.38}
$$

Now we substitute this back into the modified equation for $m_3(Y_n)$:

$$
\begin{aligned}
m_3(Y_n) &= \mathbf{E}[Y_n^3] - 3\mu_{Y_n}\sigma_{Y_n}^2 - \mu_{Y_n}^3 \\
m_3(Y_n) &= \tfrac{e^{3c+\frac{3d^2}{2}}}{n^3}(ne^{3d^2} + 3(_nP_2)e^{d^2(1+2\rho)} + 6\binom{n}{3}e^{3d^2\rho}) \\
&\quad -3e^{c+\frac{d^2}{2}}\tfrac{e^{2c+d^2}}{n^2}(ne^{d^2} + 2\binom{n}{2}e^{d^2\rho} - n^2) - (e^{c+\frac{d^2}{2}})^3 \\
&= \tfrac{e^{3c+\frac{3d^2}{2}}}{n^3}[ne^{3d^2} + 3(_nP_2)e^{d^2(1+2\rho)} + 6\binom{n}{3}e^{3d^2\rho} - 3n(ne^{d^2} + 2\binom{n}{2}e^{d^2\rho} - n^2) - n^3] \\
&= \tfrac{e^{3c+\frac{3d^2}{2}}}{n^3}[ne^{3d^2} + 3(_nP_2)e^{d^2(1+2\rho)} + 6\binom{n}{3}e^{3d^2\rho} - 3n^2e^{d^2} - 6n\binom{n}{2}e^{d^2\rho} + 3n^3 - n^3] \\
\Rightarrow m_3(Y_n) &= \tfrac{e^{3c+\frac{3d^2}{2}}}{n^3}[ne^{3d^2} + 3(_nP_2)e^{d^2(1+2\rho)} + 6\binom{n}{3}e^{3d^2\rho} - 3n^2e^{d^2} - 6n\binom{n}{2}e^{d^2\rho} + 2n^3]
\end{aligned}
\tag{B.39}
$$

Finally we can substitute back into our equation for skewness:

$$
\begin{aligned}
skew(Y_n) &= \frac{m_3(Y_n)}{\sigma_{Y_n}^3} \\
&= \frac{\frac{e^{3c+\frac{3d^2}{2}}}{n^3}[ne^{3d^2}+3(_nP_2)e^{d^2(1+2\rho)}+6\binom{n}{3}e^{3d^2\rho}-3n^2e^{d^2}-6n\binom{n}{2}e^{d^2\rho}+2n^3]}{(\frac{e^{2c+d^2}}{n^2}(ne^{d^2}+2\binom{n}{2}e^{d^2\rho}-n^2))^{\frac{3}{2}}} \\
&= \frac{\frac{\cancel{e^{3c+\frac{3d^2}{2}}}}{n^3}[ne^{3d^2}+3(_nP_2)e^{d^2(1+2\rho)}+6\binom{n}{3}e^{3d^2\rho}-3n^2e^{d^2}-6n\binom{n}{2}e^{d^2\rho}+2n^3]}{\frac{\cancel{e^{3c+\frac{3d^2}{2}}}}{n^3}(ne^{d^2}+2\binom{n}{2}e^{d^2\rho}-n^2)^{\frac{3}{2}}} \\
\Rightarrow skew(Y_n) &= \frac{ne^{3d^2}+3(_nP_2)e^{d^2(1+2\rho)}+6\binom{n}{3}e^{3d^2\rho}-3n^2e^{d^2}-6n\binom{n}{2}e^{d^2\rho}+2n^3}{(ne^{d^2}+2\binom{n}{2}e^{d^2\rho}-n^2)^{\frac{3}{2}}} \\
<< OR >> skew(Y_n) &= \frac{\mathbf{E}[Y_n^3]-3\mu_{Y_n}\sigma_{Y_n}^2-\mu_{Y_n}^3}{\sigma_{Y_n}^3}
\end{aligned}
\tag{B.40}
$$

This is in terms of values for which we have previously solved.

B.4.4 Kurtosis

For kurtosis we follow the same method as we did for skewness and variance. We start with $kurt(Y_n) = \frac{m_4(Y_n)}{\sigma_{Y_n}^4}$ and we rewrite $m_4(Y_n)$ in a more useful form:

$$
\begin{aligned}
m_4(Y_n) &= \mathbf{E}[(Y_n - \mu_{Y_n})^4] \\
&= \mathbf{E}[Y_n^4] - 4\mu\mathbf{E}[Y_n^3] + 6\mu_{Y_n}^2\sigma_{Y_n}^2 + 3\mu_{Y_n}^4
\end{aligned}
\tag{B.41}
$$

Now we look at the terms that arise in Y_n^4

$$
\begin{aligned}
Y_n^4 &= \tfrac{1}{n^4}(\sum_{i=1}^n e^{c+dZ_i})^4 \\
&= \tfrac{1}{n^4}(\sum_{i=1}^n (e^{c+dZ_i})^4 + 4\sum_{i=1}^n \sum_{j=1,j\neq i}^n (e^{c+dZ_i})^3 e^{c+dZ_j} \\
&\quad +6\sum_{i=1}^n \sum_{j=i+1}^n (e^{c+dZ_i})^2 (e^{c+dZ_j})^2 \\
&\quad +12\sum_{i=1}^n \sum_{j=1,j\neq i}^n \sum_{k=j+1,k\neq i}^n (e^{c+dZ_i})^2 e^{c+dZ_j} e^{c+dZ_k} \\
&\quad +24\sum_{i=1}^n \sum_{j=i+1}^n \sum_{k=j+1}^n \sum_{\ell=k+1}^n e^{c+dZ_i} e^{c+dZ_j} e^{c+dZ_k} e^{c+dZ_\ell})
\end{aligned}
\tag{B.42}
$$

There are 5 summations is this expansion. Here is a list of the number of terms in each:

$$
\begin{aligned}
1^{st} &: n \\
2^{nd} &: {}_nP_2 \\
3^{rd} &: \binom{n}{2} \\
4^{th} &: {}_nP_3 \div 2 = \binom{n}{3} \times 3 \\
5^{th} &: \binom{n}{4}
\end{aligned}
\tag{B.43}
$$

All of these follow the same logic for when order matters as discussed at length for skewness and should be obvious. We know that the terms of each summation are identical in expectation and from the multivariate normal moment generating function they are:

$$
\begin{aligned}
\mathbf{E}[(e^{c+dZ_i})^4] &= e^{4c+8d^2} \\
\mathbf{E}[(e^{c+dZ_i})^3 e^{c+dZ_j}] &= e^{4c+d^2(5+3\rho)} \\
\mathbf{E}[(e^{c+dZ_i})^2 (e^{c+dZ_j})^2] &= e^{4c+4d^2(1+\rho)} \\
\mathbf{E}[(e^{c+dZ_i})^2 e^{c+dZ_j} e^{c+dZ_k}] &= e^{4c+d^2(3+5\rho)} \\
\mathbf{E}[e^{c+dZ_i} e^{c+dZ_j} e^{c+dZ_k} e^{c+dZ_\ell}] &= e^{4c+2d^2(1+3\rho)}
\end{aligned}
\tag{B.44}
$$

So:

$$
\begin{aligned}
\mathbf{E}[Y_n^4] &= \tfrac{1}{n^4}(\sum_{i=1}^n \mathbf{E}[(e^{c+dZ_i})^4] + 4\sum_{i=1}^n \sum_{j=1,j\neq i}^n \mathbf{E}[(e^{c+dZ_i})^3 e^{c+dZ_j}] \\
&\quad +6\sum_{i=1}^n \sum_{j=i+1}^n \mathbf{E}[(e^{c+dZ_i})^2 (e^{c+dZ_j})^2] \\
&\quad +12\sum_{i=1}^n \sum_{j=1,j\neq i}^n \sum_{k=j+1,k\neq i}^n \mathbf{E}[(e^{c+dZ_i})^2 e^{c+dZ_j} e^{c+dZ_k}] \\
&\quad +24\sum_{i=1}^n \sum_{j=i+1}^n \sum_{k=j+1}^n \sum_{\ell=k+1}^n \mathbf{E}[e^{c+dZ_i} e^{c+dZ_j} e^{c+dZ_k} e^{c+dZ_\ell}]) \\
&= \tfrac{1}{n^4}(\sum_{i=1}^n e^{4c+8d^2} + 4\sum_{i=1}^n \sum_{j=1,j\neq i}^n e^{4c+d^2(5+3\rho)} \\
&\quad +6\sum_{i=1}^n \sum_{j=i+1}^n e^{4c+4d^2(1+\rho)} \\
&\quad +12\sum_{i=1}^n \sum_{j=1,j\neq i}^n \sum_{k=j+1,k\neq i}^n e^{4c+d^2(3+5\rho)} \\
&\quad +24\sum_{i=1}^n \sum_{j=i+1}^n \sum_{k=j+1}^n \sum_{\ell=k+1}^n e^{4c+2d^2(1+3\rho)}) \\
&= \tfrac{1}{n^4}(ne^{4c+8d^2} + 4({}_nP_2)e^{4c+d^2(5+3\rho)} + 6\binom{n}{2}e^{4c+4d^2(1+\rho)} \\
&\quad +12((\tfrac{1}{2}){}_nP_3)e^{4c+d^2(3+5\rho)} + 24\binom{n}{4}e^{4c+2d^2(1+3\rho)}) \\
\Rightarrow \mathbf{E}[Y_n^4] &= \tfrac{e^{4c+2d^2}}{n^4}(ne^{6d^2} + 4({}_nP_2)e^{3d^2(1+\rho)} + 6\binom{n}{2}e^{2d^2(1+2\rho)} \\
&\quad +12((\tfrac{1}{2}){}_nP_3)e^{d^2(1+5\rho)} + 24\binom{n}{4}e^{6d^2\rho})
\end{aligned}
\tag{B.45}
$$

Finally we have now calculated all the values to substitute into our equation for kurtosis:

$$
\begin{aligned}
kurt(Y_n) &= \frac{m_4(Y_n)}{\sigma^4_{Y_n}} \\
kurt(Y_n) &= \frac{\mathbf{E}[Y_n^4]-4\mu\mathbf{E}[Y_n^3]+6\mu^2_{Y_n}\sigma^2_{Y_n}+3\mu^4_{Y_n}}{\sigma^4_{Y_n}}
\end{aligned}
\tag{B.46}
$$

B.5 Moments of Sum of n General Lognormal Random Variables

We will not analytically compute the moments for a sum of n lognormal random variables that are or are not all identically correlated, have the the different parameters, and are not evenly weighted because the resulting expectations are not recombining. This should be done by simulation.

Appendix C

Tails of Sums of Lognormal Random Variables

Here we will give a more in depth analysis of how the tails change as n changes for a few different levels of correlation. For all of the examples seen here $Y_1 \sim \mathcal{LN}(0, 1)$ and each term in $Y_n \sim \mathcal{LN}(c, d)$ with parameters chosen to match the first two moments as explained above.

C.1 Analytic Moments

First we will look at lognormal random variables generated from uncorrelated normal random variables; see $\rho = 0$ moments in table C.1 and figure C.1.

From the table or the graph it can easily be seen that the kurtosis blows up as n increases for $\rho = 0$. Skewness even increases linearly here. This means that for lognormal random variables with $\rho = 0$ have increasingly fatter tails than the comparable lognormal random variable and fatter tails than the sum of one less comparable lognormal random variables. That the skewness does increase as well does hinder our ability to use kurtosis to compare tails as the increasing right skew, with a distribution that has a lower bound, will influence the kurtosis. The kurtosis does not change independent of skewness here. To help overcome this we can also look at a series of QQ plots. We will do this when we look at simulations.

Now we will look at a summation of lognormal random variables that are a little more correlated, $\rho = 0.3$; see table C.1 and figure C.2. This is much more promising. Even though as n initially increases the kurtosis does increase, it does not blow up. It even starts to decrease when $n = 13$ and for large enough n converges to the kurtosis of the single lognormal distribution. Skewness also does flatten out and converge to the single lognormal distribution so it stops interfering with our ability to compare the tails by comparing the kurtosis once it flattens out.

Let's look at a higher value of ρ, $\rho = 0.6$; see table C.3 and figure C.3. This is even better than the last one. Skewness is basically flat and converged to the skewness of the single lognormal distribution beyond $n = 3$. For $n > 4$ kurtosis plummets and converges to the kurtosis of the single lognormal distribution.

Let's finally look at some very highly correlated ones, $\rho = 0.9$, in table C.4 and figure C.4. And better still. Skewness is basically flat for all values of n. Kurtosis does increase at $n = 2$

Table C.1: Theoretical Moments of Sums of Uncorrelated Standard Lognormal Random Variables

n	mean	variance	skewness	kurtosis
1	1.6487	4.6708	6.1849	113.9364
2	1.6487	4.6708	8.4373	310.5621
3	1.6487	4.6708	10.6896	672.6729
4	1.6487	4.6708	12.9420	1252.5722
5	1.6487	4.6708	15.1944	2102.5630
6	1.6487	4.6708	17.4468	3274.9488
7	1.6487	4.6708	19.6992	4822.0328
8	1.6487	4.6708	21.9515	6796.1182
9	1.6487	4.6708	24.2039	9249.5083
10	1.6487	4.6708	26.4563	12234.5065
11	1.6487	4.6708	28.7087	15803.4159
12	1.6487	4.6708	30.9611	20008.5398
13	1.6487	4.6708	33.2134	24902.1816
14	1.6487	4.6708	35.4658	30536.6444
15	1.6487	4.6708	37.7182	36964.2316

Figure C.1: Theoretical Moments of Sums of Uncorrelated Standard Lognormal Random Variables

Table C.2: Theoretical Moments of Sums of Correlated Standard Lognormal Random Variables, $\rho = 0.3$

n	mean	variance	skewness	kurtosis
1	1.6487	4.6708	6.1849	113.9364
2	1.6487	4.6708	6.9728	175.9525
3	1.6487	4.6708	7.3819	224.1908
4	1.6487	4.6708	7.6173	261.9439
5	1.6487	4.6708	7.7569	291.4749
6	1.6487	4.6708	7.8383	314.4694
7	1.6487	4.6708	7.8825	332.2214
8	1.6487	4.6708	7.9017	345.7418
9	1.6487	4.6708	7.9039	355.8300
10	1.6487	4.6708	7.8943	363.1240
11	1.6487	4.6708	7.8764	368.1371
12	1.6487	4.6708	7.8527	371.2851
13	1.6487	4.6708	7.8250	372.9068
14	1.6487	4.6708	7.7946	373.2794
15	1.6487	4.6708	7.7624	372.6310

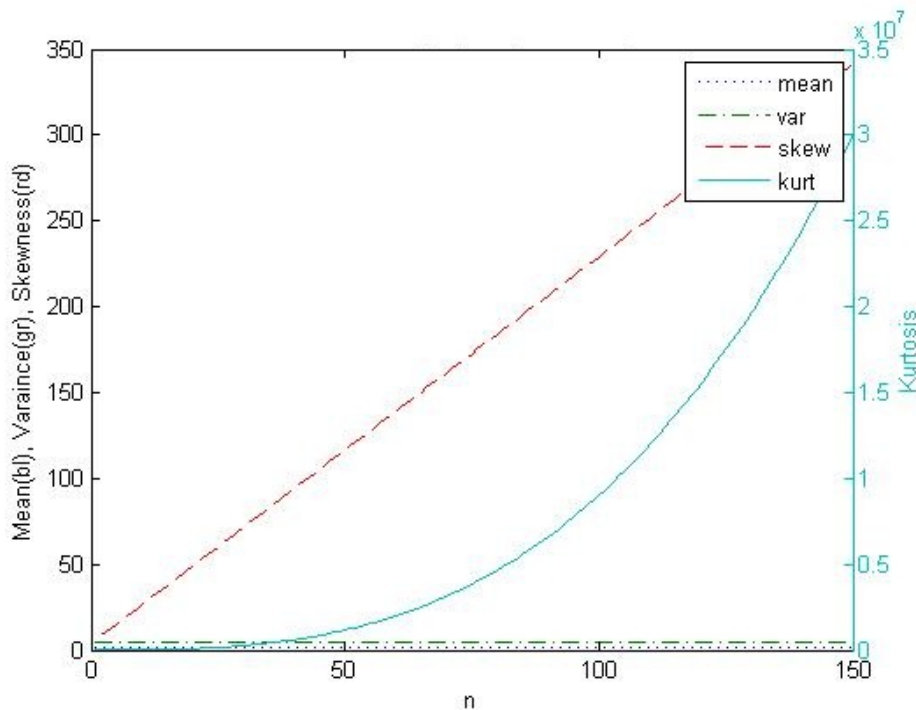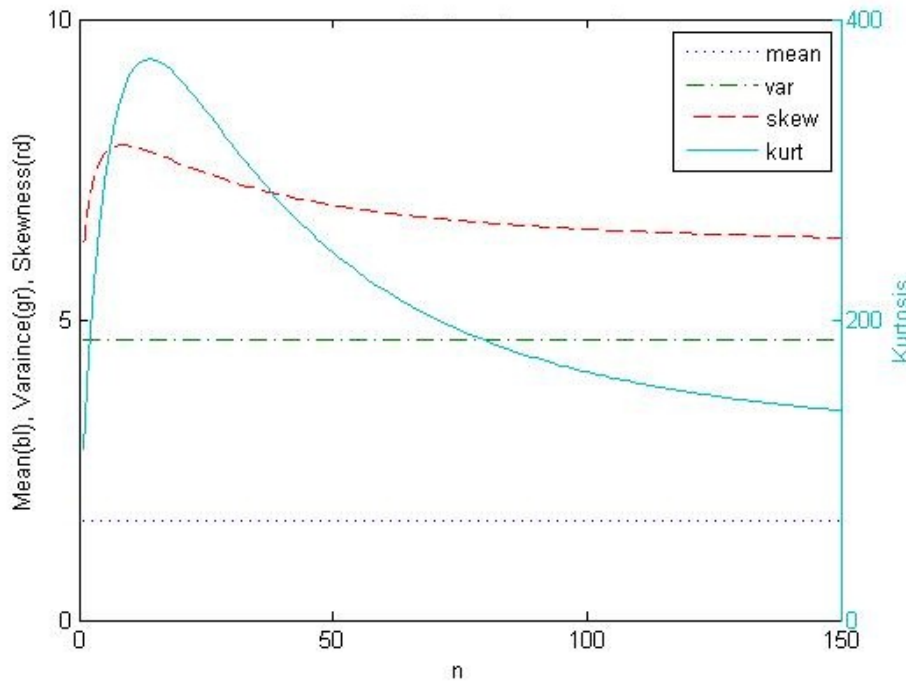Figure C.2: Theoretical Moments of Sums of Correlated Standard Lognormal Random Variables, $\rho = 0.3$

Table C.3: Theoretical Moments of Sums of Correlated Standard Lognormal Random Variables, $\rho = 0.6$

n	mean	variance	skewness	kurtosis
1	1.6487	4.6708	6.1849	113.9364
2	1.6487	4.6708	6.3052	122.2588
3	1.6487	4.6708	6.3036	122.2634
4	1.6487	4.6708	6.2865	121.0494
5	1.6487	4.6708	6.2697	119.8346
6	1.6487	4.6708	6.2558	118.8249
7	1.6487	4.6708	6.2447	118.0206
8	1.6487	4.6708	6.2358	117.3839
9	1.6487	4.6708	6.2286	116.8771
10	1.6487	4.6708	6.2229	116.4698
11	1.6487	4.6708	6.2181	116.1390
12	1.6487	4.6708	6.2142	115.8672
13	1.6487	4.6708	6.2109	115.6418
14	1.6487	4.6708	6.2082	115.4529
15	1.6487	4.6708	6.2058	115.2932

but beyond that it falls faster than before, and converges faster too.

We will look at only one case where $\rho < 0$, because it is just an exaggeration of the bad situation we saw when $\rho = 0$. The case where $\rho = -0.2$ is shown in table C.5 and figure C.5. Since indices do not contain all stocks that are negatively correlated to all the others in the index we will not worry about this undesirable result. It is also irrelevant that the kurtosis here blows up because even though we can calculate this theoretically we cannot simulate this. To simulate correlation we need a Cholesky factorization of the correlation matrix, which must be positive definite to do this. There are some negative correlations that we can compute the Cholesky factorization for, but only for limited n. Some samples of functional pairs of negative correlations and the maximum n that can be handled are displayed in table C.6.

C.2 QQ plots

Here we will show and discuss simulations with sample size of 1,000,000 where $\rho = 0, 0.3, 0.6, 0.9$ and examine the tails by looking at QQ plots for $n = 2, 3, 4, 5, 10, 15$ for $\rho = 0$ and $n = 2, 10, 25, 50, 100, 150$ for $\rho = 0.3, 0.6, 0.9$ against $n = 1$. The parameters are chosen as detailed at the top of this appendix.

Table C.7 shows what we expect to see from our theoretical results for each correlation.

What we will have to consider is that our expected results are from a single measure for both tails, whereas when we read off of a QQ plot we can read both tails separately. Now let's see how our expectations hold up.

First we look at $\rho = 0$. The QQ plots in Figures C.6 confirm the pattern we saw in our theoretical moments by simulation. What we can easily see here is that this is true for the

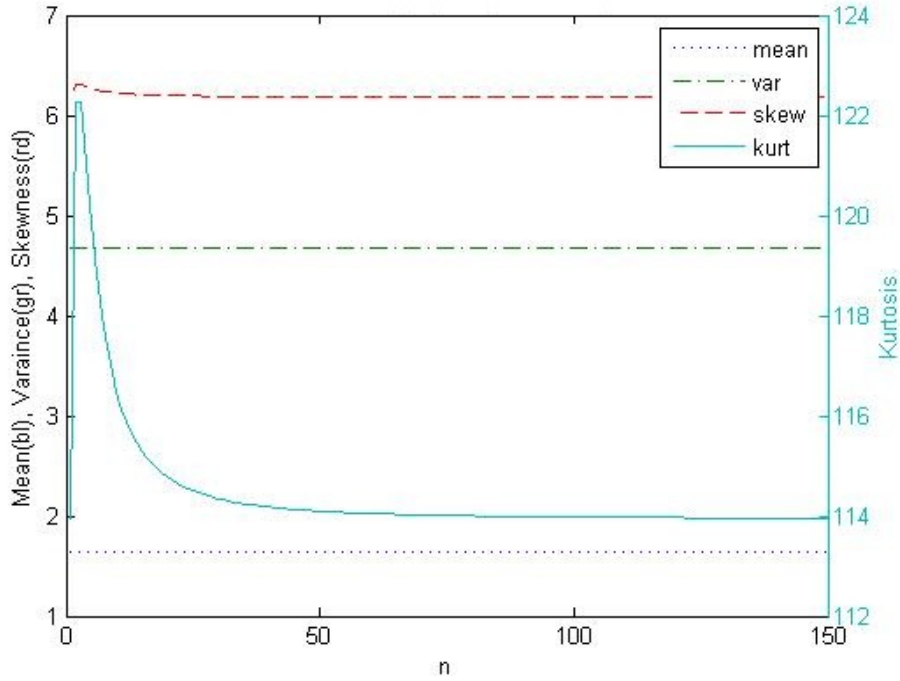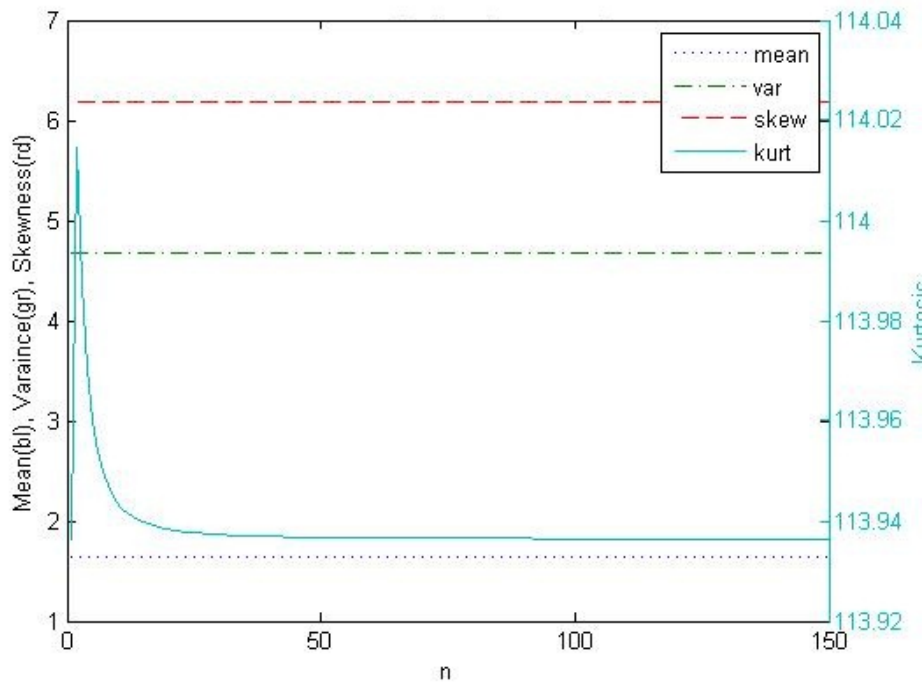Figure C.3: Theoretical Moments of Sums of Correlated Standard Lognormal Random Variables, $\rho = 0.6$

Table C.4: Theoretical Moments of Sums of Correlated Standard Lognormal Random Variables, $\rho = 0.9$

n	$mean$	$variance$	$skewness$	$kurtosis$
1	1.6487	4.6708	6.1849	113.9364
2	1.6487	4.6708	6.1862	114.0147
3	1.6487	4.6708	6.1857	113.9875
4	1.6487	4.6708	6.1854	113.9704
5	1.6487	4.6708	6.1853	113.9603
6	1.6487	4.6708	6.1852	113.9540
7	1.6487	4.6708	6.1851	113.9499
8	1.6487	4.6708	6.1851	113.9471
9	1.6487	4.6708	6.1850	113.9450
10	1.6487	4.6708	6.1850	113.9435
11	1.6487	4.6708	6.1850	113.9424
12	1.6487	4.6708	6.1850	113.9415
13	1.6487	4.6708	6.1850	113.9408
14	1.6487	4.6708	6.1849	113.9402
15	1.6487	4.6708	6.1849	113.9397

Figure C.4: Theoretical Moments of Sums of Correlated Standard Lognormal Random Variables, $\rho = 0.9$

Table C.5: Theoretical Moments of Sums of Correlated Standard Lognormal Random Variables, $\rho = -0.2$

n	$mean$	$variance$	$skewness$	$kurtosis$
1	1.6487	4.6708	6.1849	113.9364
2	1.6487	4.6708	9.6551	432.5010
3	1.6487	4.6708	13.6852	1196.1126
4	1.6487	4.6708	18.0499	2650.9388
5	1.6487	4.6708	22.6453	5066.7076

Table C.6: Maximum Number of Negatively Correltated Lognormal Random Variables

ρ	$max(n)$
−0.1	10
−0.2	5
−0.3	4
−0.4	3
−0.5	2
−0.6	2
−0.7	2

Table C.7: Expected Results of Simulations

ρ	expected results
0	For $n = 2$ we expect the sum to have fatter tails than a single lognormal, and the fatness should increase with n at an increasing rate.
0.3	For $n = 2$ we expect the sum to have fatter tails than a single lognormal, the fatness should increase with n at a decreasing rate until it starts to noticeably flatten out around $n = 8$. Eventually, around $n = 20$, the fatness will start to decrease. For very large n, $n > 80$, the sum will begin to have slightly thinner tails than the single lognormal.
0.6	For $n = 2$ we expect the sum to have fatter tails than a single lognormal, the fatness should increase with n until $n = 4$. Then fatness will start to decrease. Between $n = 7$ and $n = 8$ the tails will be exactly the same as a single lognormal. For larger n, $n > 8$, the sum will have increasingly thinner tails than the single lognormal.
0.9	For $n = 2, 3$ we expect the sum to have the same tails as a single lognormal. At $n = 4$ there will be a jump, the sum will be fatter than the single, then fatness will start to decrease. At $n = 7$ the tails will be the same as a single lognormal. For larger n, $n > 7$, the sum will have increasingly thinner tails than the single lognormal.

Figure C.5: Theoretical Moments of Sums of Correlated Standard Lognormal Random Variables, $\rho = -0.2$
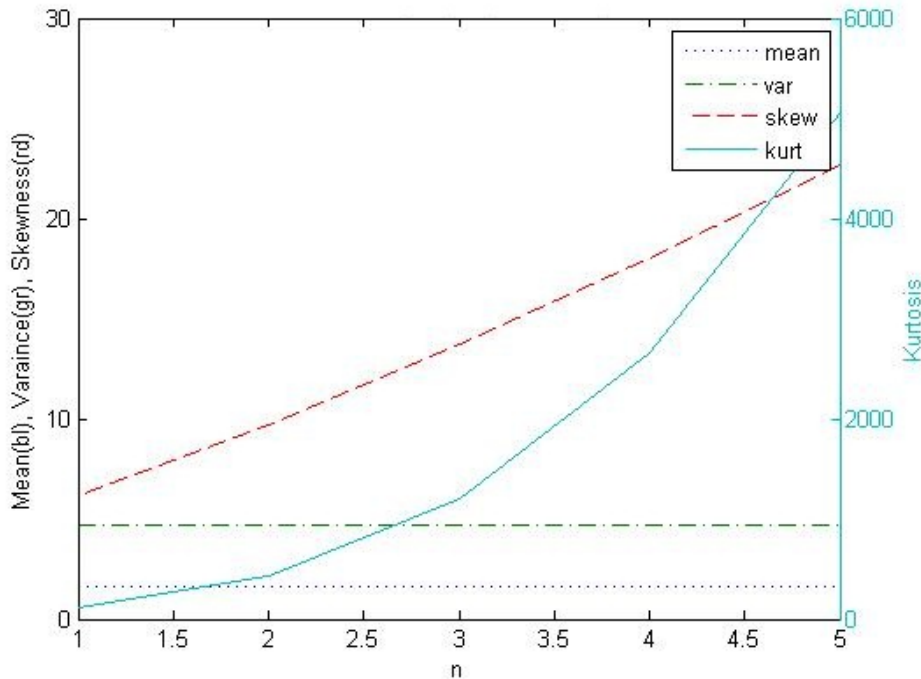
right tails, a summation of two lognormals has a fatter right tail than a single lognormal and the fatness increases as n increases. It is difficult to see the left tails though. This is because lognormal random variables have a lower bound of zero. In figure C.7 we have zoomed in on the left tails. Now we can see that what we expected is also true for the left tails, a summation of two lognormals has a fatter left tail than a single lognormal and the fatness increases as n increases.

Next is $\rho = 0.3$; figure C.8 is again hard to read very accurately, especially the left tail, so we will zoom in on the left tail, figure C.9, and examine that first. Now we can see that for small values of n the sum of n lognormals has somewhat of a fatter tail left than a single lognormal. As n increases this moves all the way from having the same left tail to having a slightly thinner tail than a single lognormal for sufficiently large n. When the zoom of this diagram is considered these fatnesses and thinnesses are slight. Since that made it significantly easier to read we will also zoom in on the right tails in figure C.10. We can see easily now that for $n = 2$ the sum of lognormals has a slightly fatter right tail than a single lognormal and as n increases the degree of fatness increases. However, for sufficiently large n the fattening seems to level off, and remain at a constant level.

Now if we combine our left and right to get a combined level of fatness we can compare that to what we expected to see. For $n = 2$ together the tails of the sum of lognormals are fatter than a single lognormal. As n increases the right tail's fatness increases and the left tail's fatness decreases slightly, so for mid-level values of n the fatness does increase at a decreasing rate. Finally for sufficiently large n the right tail levels off and the right tail becomes thinner than a single lognormal, so overall the tail is thinning, though for our largest n shown still fatter

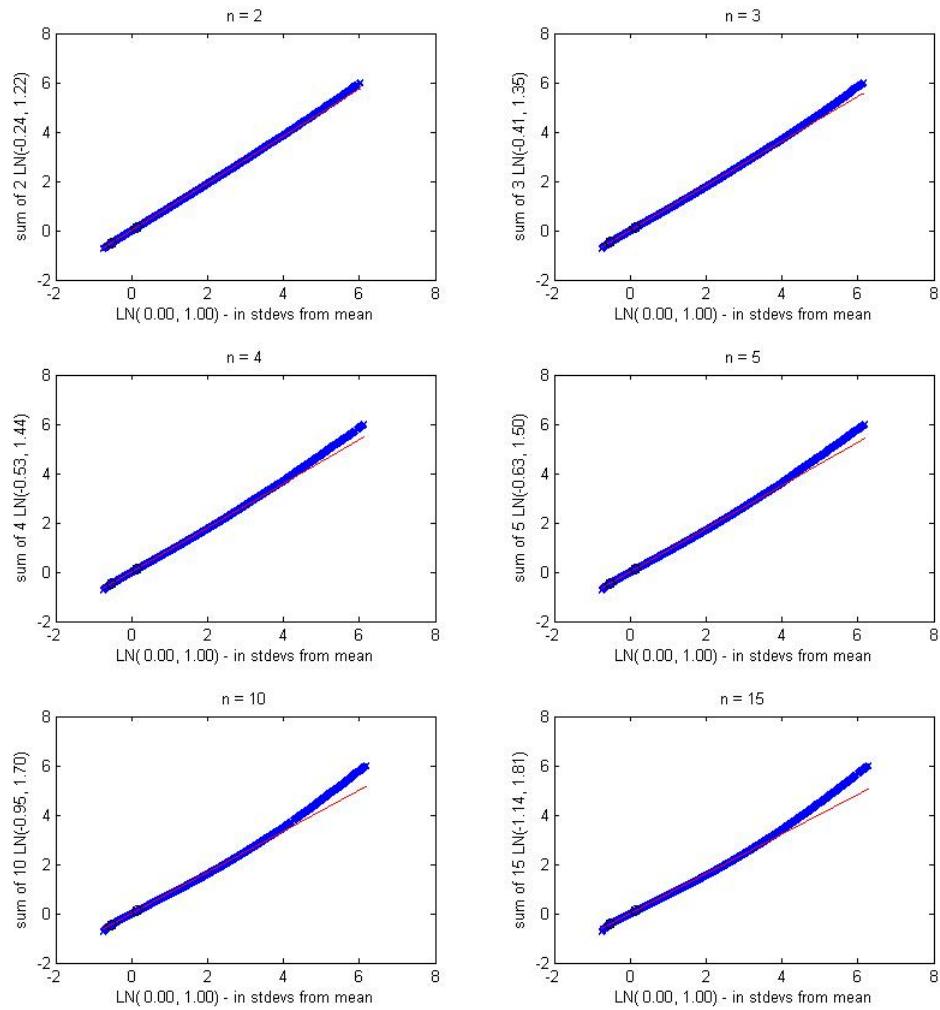Figure C.6: QQ Plot of Sums of Uncorrelated Lognormal Random Variables

Figure C.7: QQ Plot of Sums of Uncorrelated Lognormal Random Variables, Left Tails

Figure C.8: QQ Plot of Sums of Correlated Lognormal Random Variables, $\rho = 0.3$

Figure C.9: QQ Plot of Sums of Correlated Lognormal Random Variables, $\rho = 0.3$ Left Tail
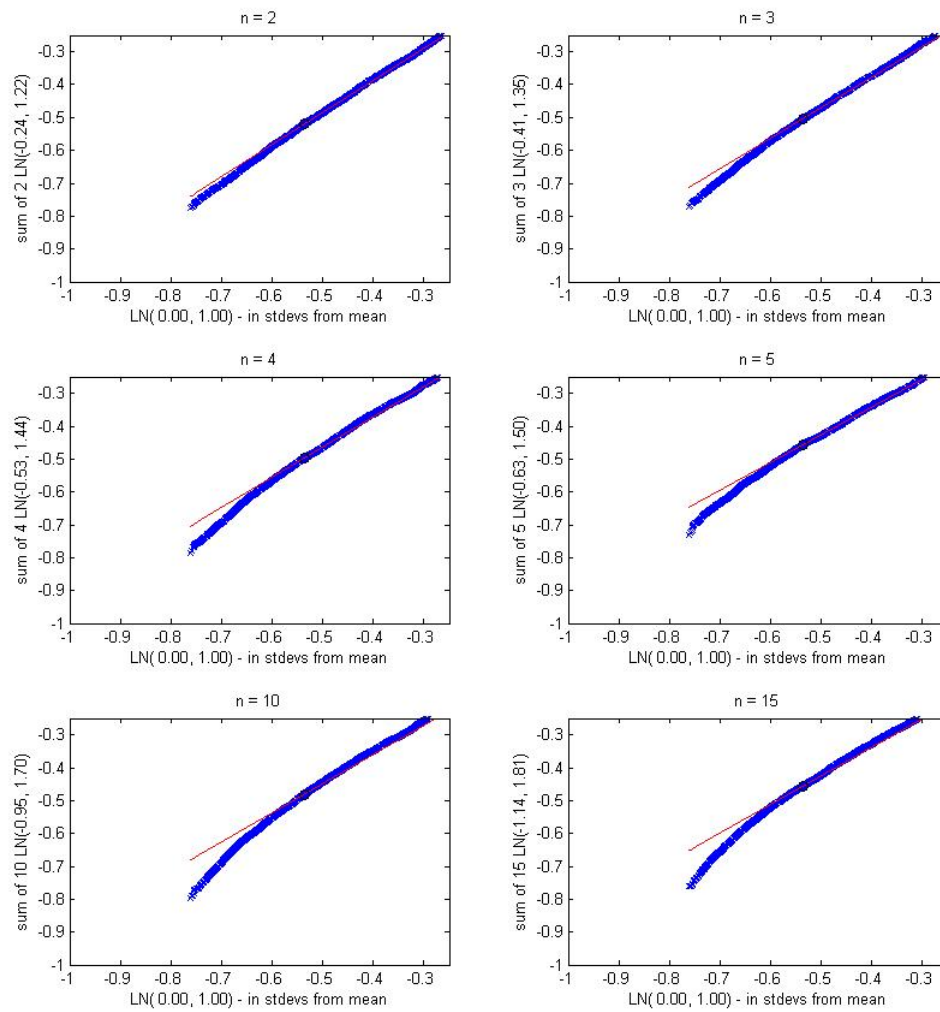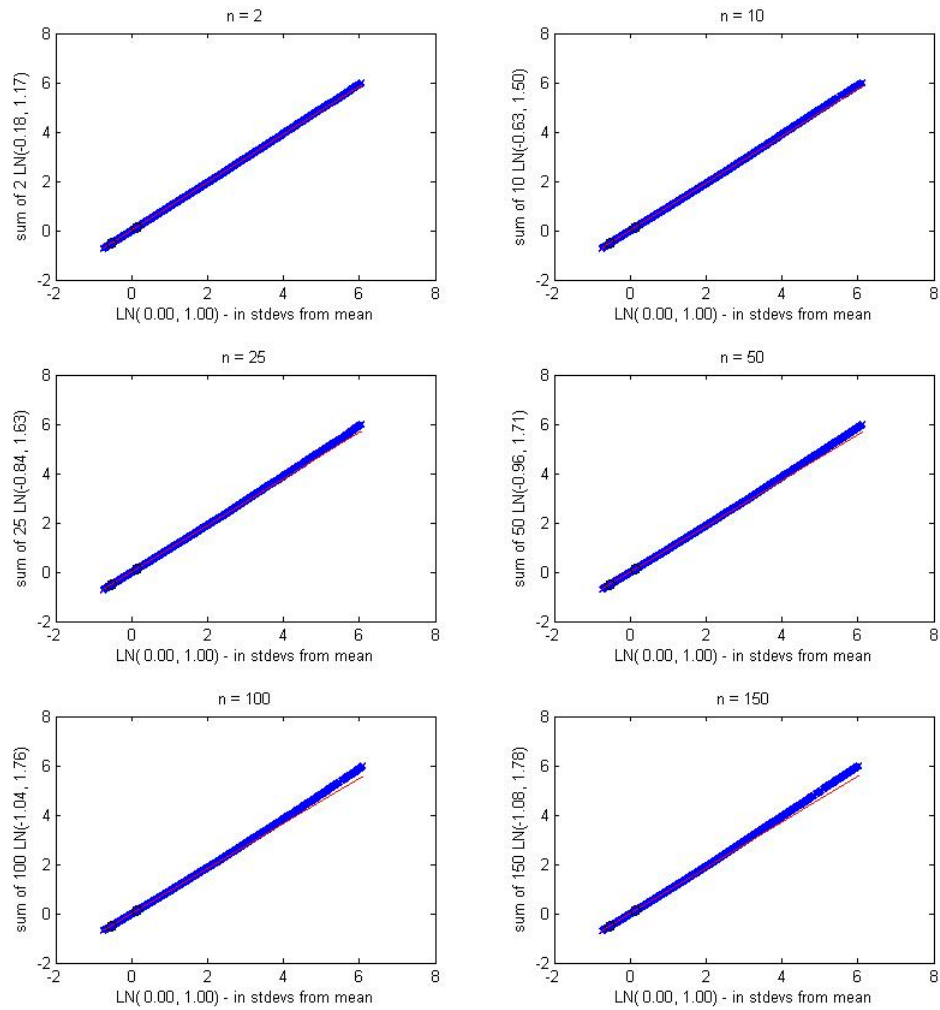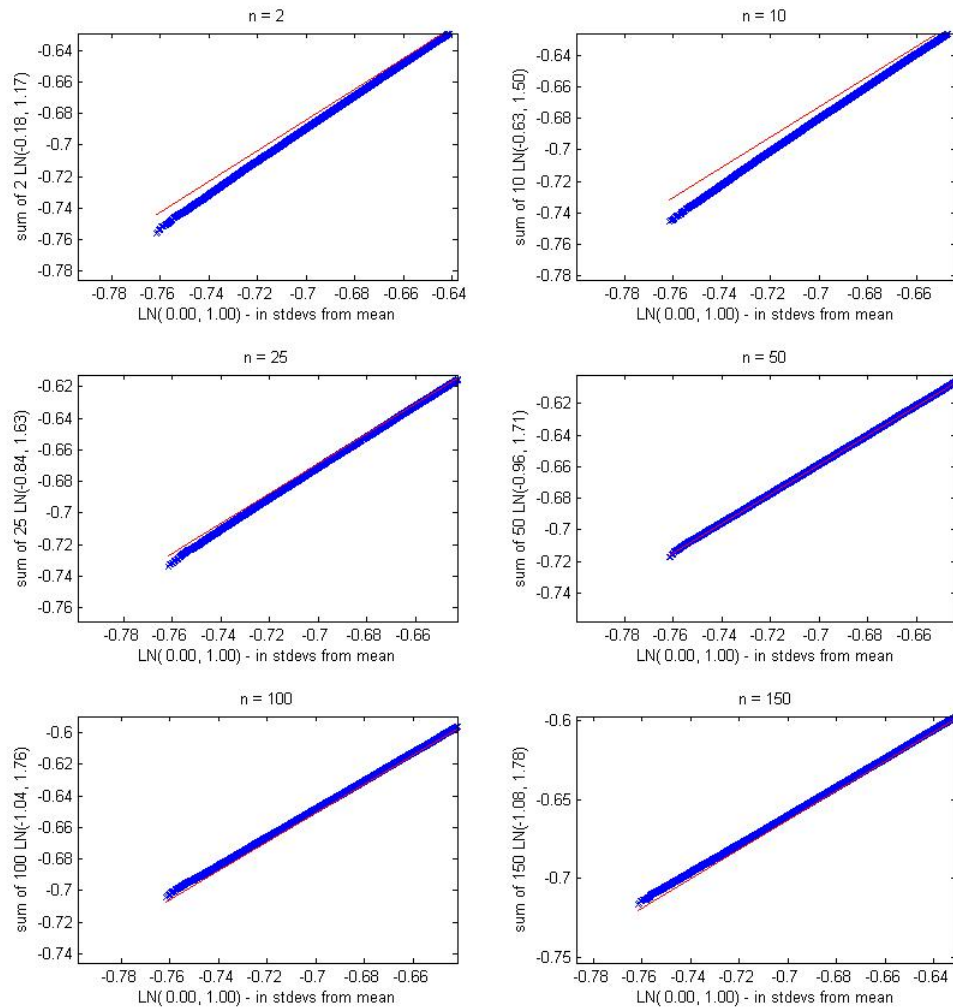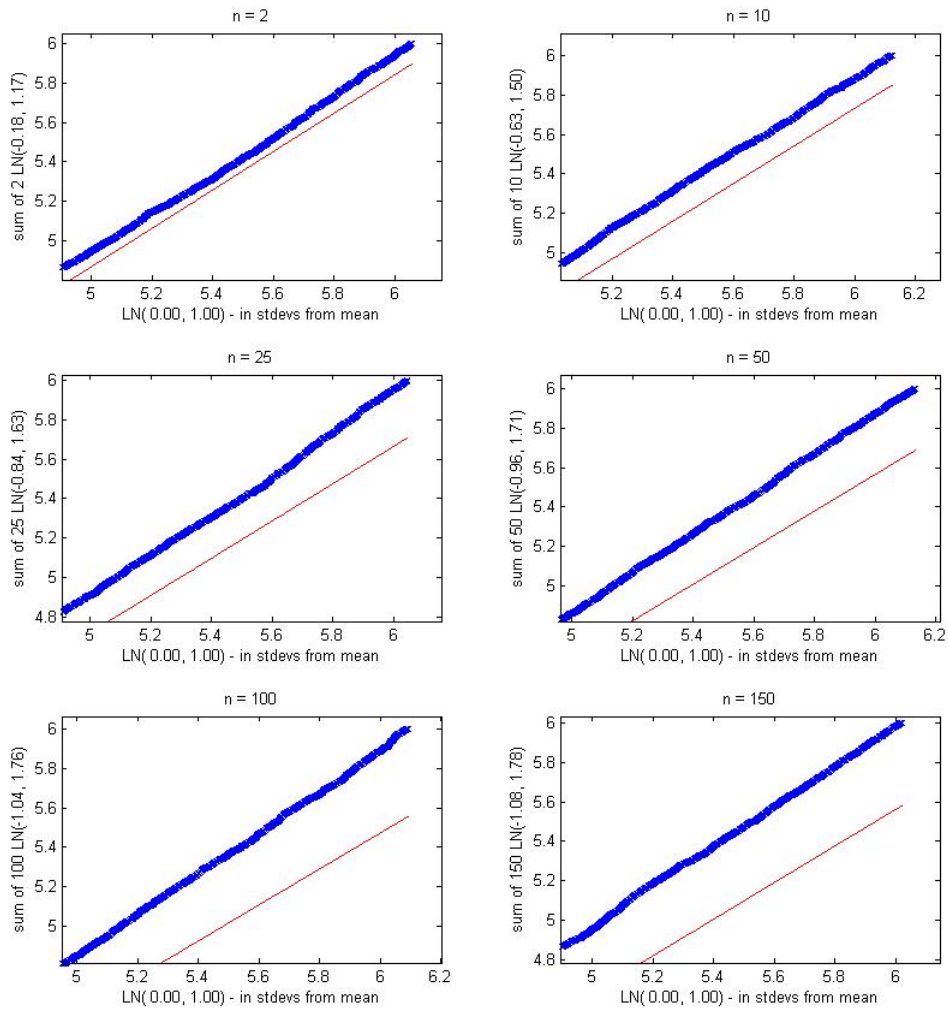
Figure C.10: QQ Plot of Sums of Correlated Lognormal Random Variables, $\rho = 0.3$ Right Tail

than a single lognormal. So compared to our theoretical prediction the pattern we hoped to see was observed but it developed more slowly than expected.

Now we will look at $\rho = 0.6$ in the same fashion.

From figure C.11 it looks as though for all of these values of n a sum of n lognormals is distributed the same as a single lognormal. But let's take a closer look at the left tails in figure C.12. Now we can be sure that the left tails of a sum of lognormals are distributed the same as a single lognormal. Let's look at the right tails in figure C.13

From this figure we could say that maybe for $n = 4$ a sum of lognormals has a slightly thinner tail than a single lognormal, but we will not because our data points run parallel to the quantile line rather than curving away from it. Also, we are looking at the sixth standard deviation away from the mean so we will read this as following the same distribution. We will also read all of the others as a sum of lognormals having the same distribution as a single lognormal. So overall we can easily see that this resolves to a sum of lognormals following the same distribution as a single lognormal for $\rho = 0.6$.

Finally we will look at our most extreme case, $\rho = 0.9$ in figure C.14

As before these plots seem to say that a sum of n lognormals follow the same distribution as a single lognormal, but to be sure we will take a closer look at each tail. First the left tail, figure C.15. Now we have verified what we suspected for the left tail. Now to look at the right in figure C.16. There is some slight wiggle seen here but since it does not follow a defined pattern and it is over five standard deviations from the mean we will read this too as before. So overall we can say that for $\rho = 0.9$ a sum of lognormal random variables follows the same distribution as a single lognormal random variable.

C.3 Non-Constant Correlation

It is also interesting to note that for a particular correlation between the normal random variables, $\rho \in [-1, 1]$, from which the lognormal random variables are generated the correlation between the lognormals is not constant as n changes unless $\rho = 0, 1$. This is because correlation is a function of d and ρ, and d is dependent on n:

$$
\begin{aligned}
\rho_{A,B} = corr(A, B) &= \frac{cov(A,B)}{\sigma_A \sigma_B}) \\
cov(A, B) &= \mathbf{E}[(A - \mu_A)(B - \mu_B)] \\
cov(e^{c+dZ_i}, e^{c+dZ_j}) &= \mathbf{E}[(e^{c+dZ_i} - e^{c+\frac{d^2}{2}})(e^{c+dZ_j} - e^{c+\frac{d^2}{2}})] \\
&= e^{c+d^2}(e^{d^2\rho} - 1) \\
\Rightarrow corr(e^{c+dZ_i}, e^{c+dZ_j}) &= \frac{(e^{d^2\rho}-1)}{(e^{d^2}-1)}
\end{aligned}
\tag{C.1}
$$

where $\mathbf{E}[Z_i Z_j] = \rho$. In table C.8 are some sample correlations between a pair of the random variables included in the summation for various values of ρ and n.

Figure C.11: QQ Plot of Sums of Correlated Lognormal Random Variables, $\rho = 0.6$

Figure C.12: QQ Plot of Sums of Correlated Lognormal Random Variables, $\rho = 0.6$ Left Tail

Figure C.13: QQ Plot of Sums of Correlated Lognormal Random Variables, $\rho = 0.6$ Right Tail
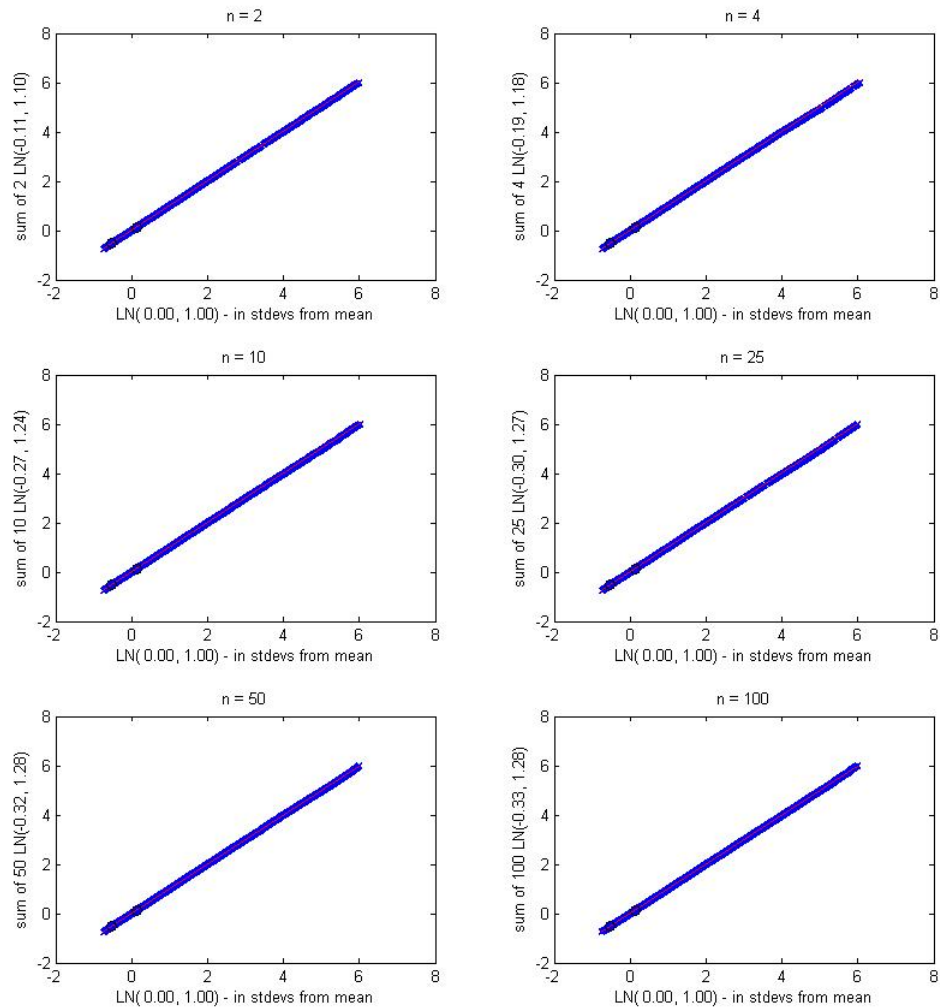
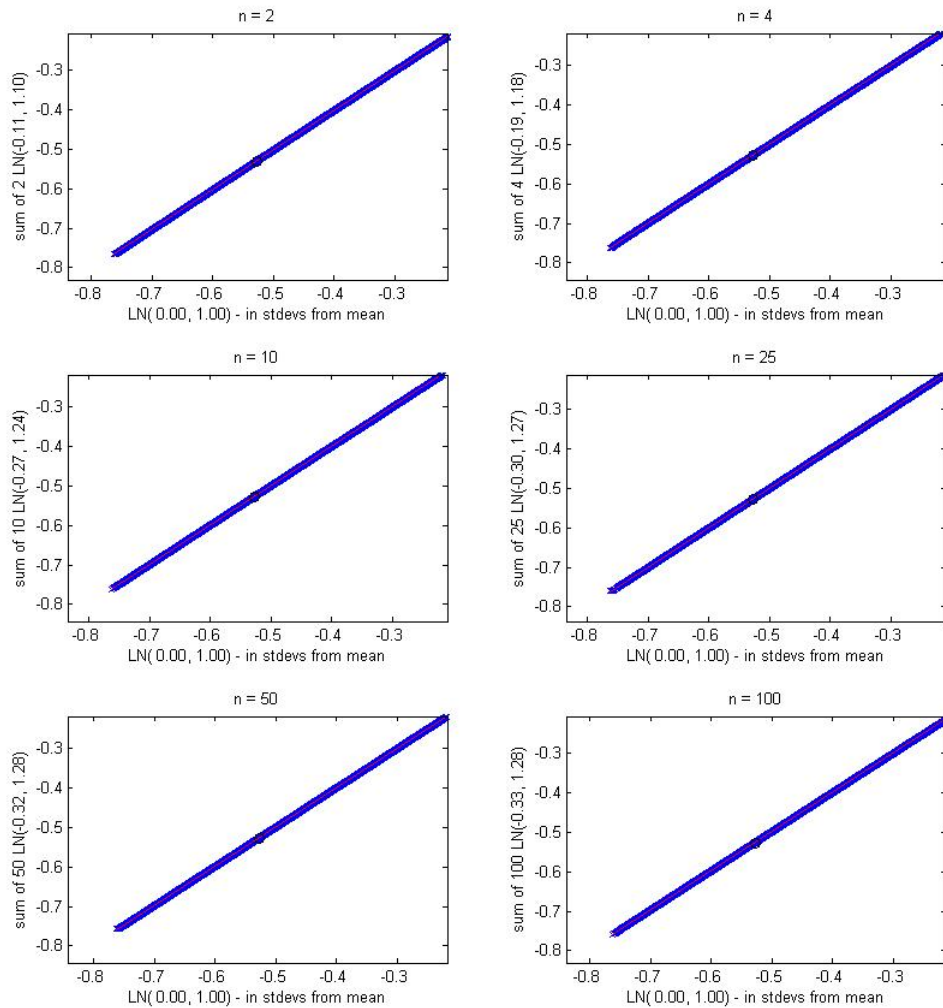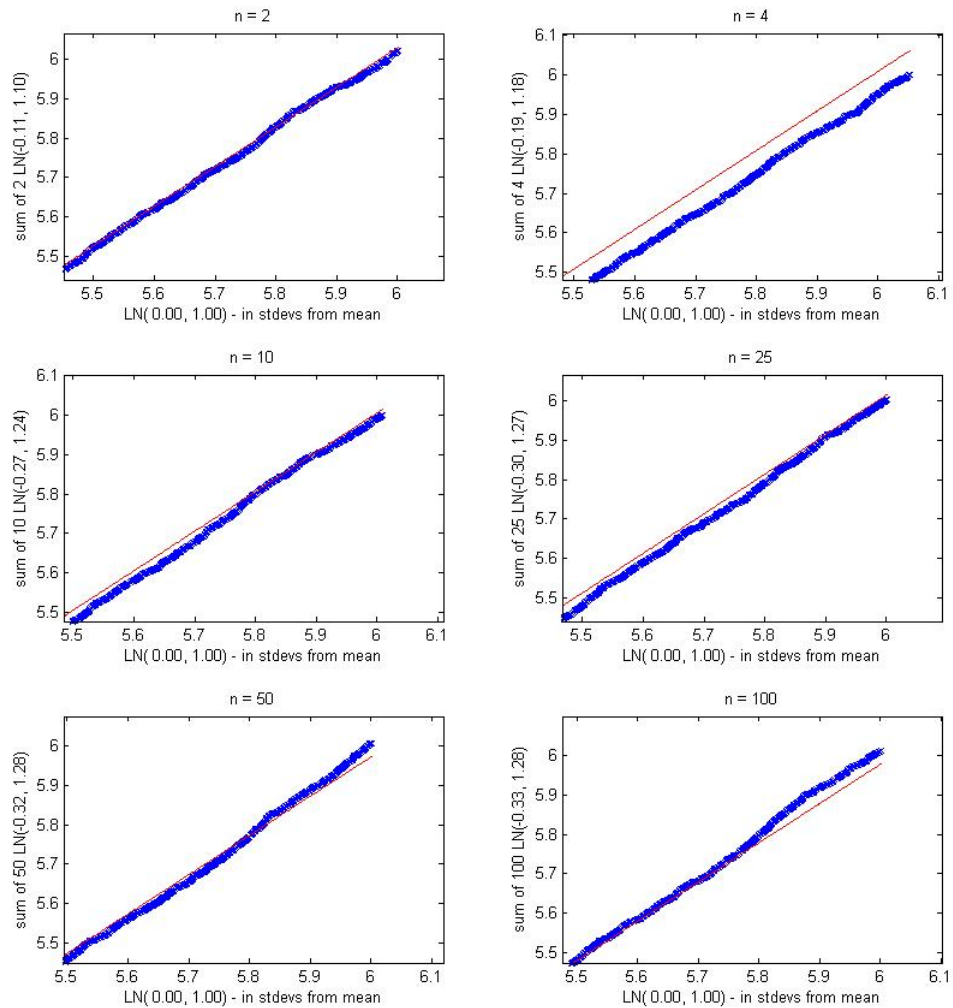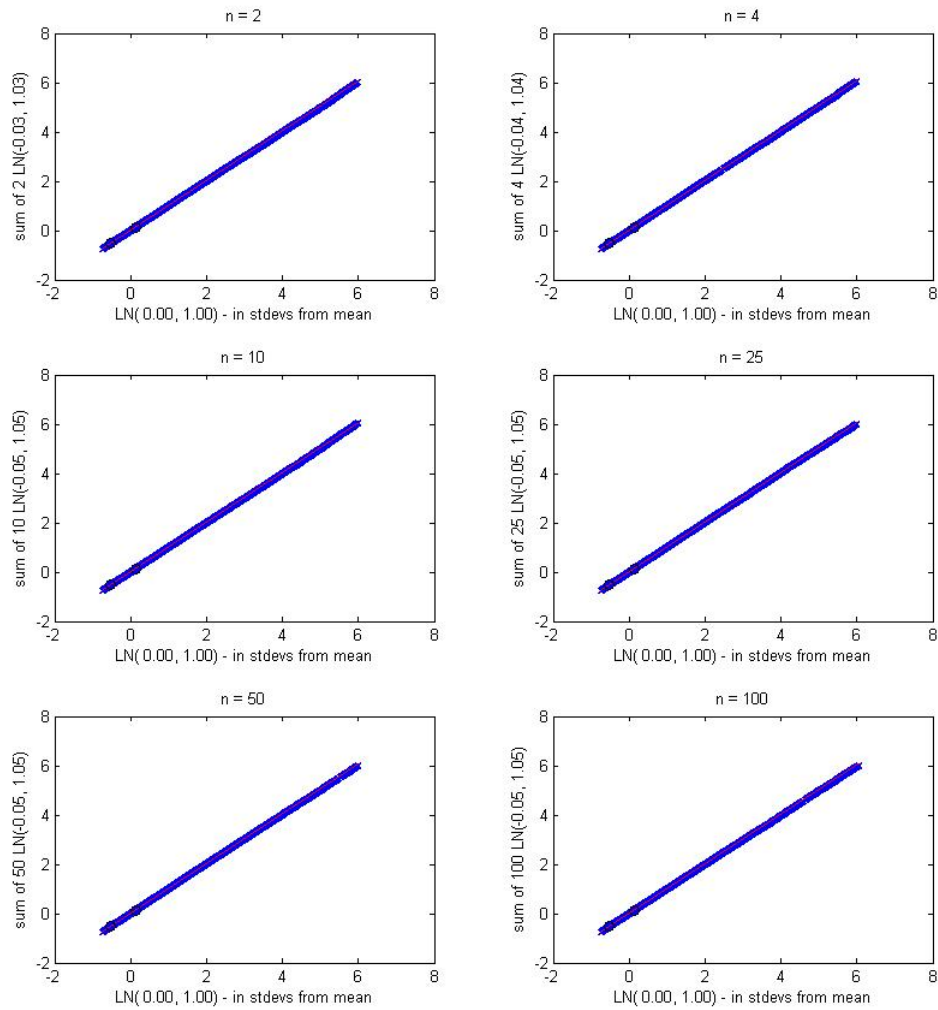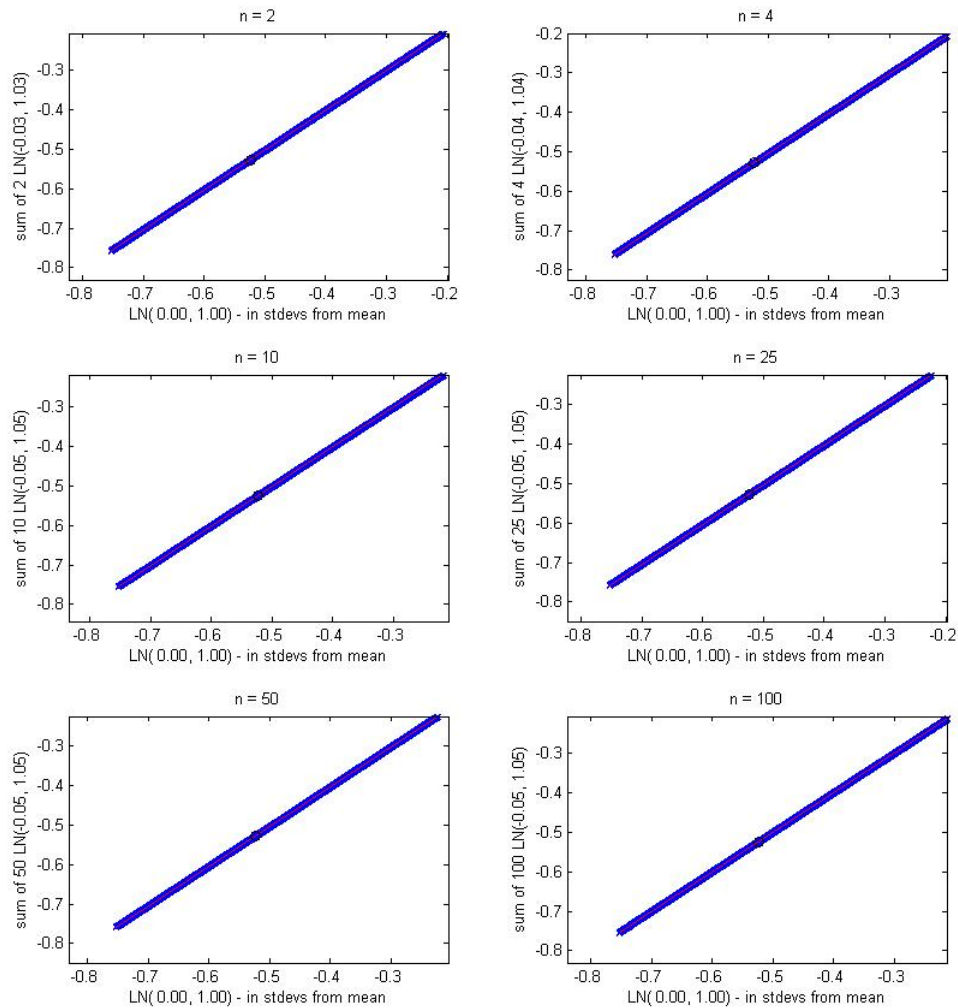Figure C.14: QQ Plot of Sums of Correlated Lognormal Random Variables, $\rho = 0.9$

Figure C.15: QQ Plot of Sums of Correlated Lognormal Random Variables, $\rho = 0.9$ Left Tail
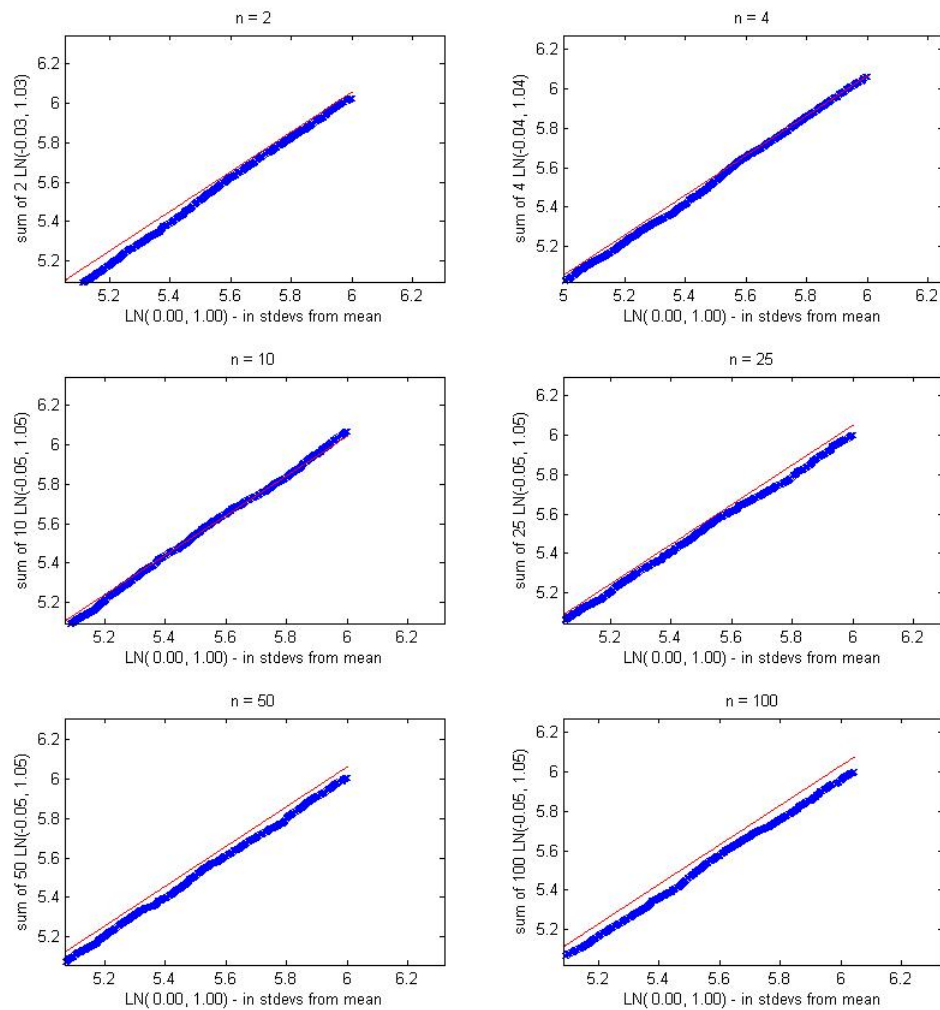
Figure C.16: QQ Plot of Sums of Correlated Lognormal Random Variables, $\rho = 0.9$ Right Tail

Table C.8: Correlations Dependent on ρ and n

$n\backslash\rho$	-1.00	-0.90	-0.80	-0.70	-0.60	-0.50	...
2	-0.1906	-0.1838	-0.1755	-0.1654	-0.1532	-0.1385	...
3	-0.1266	-0.1233	-0.1191	-0.1138	-0.1071	-0.0986	...
4	-0.0944	-0.0925	-0.0899	-0.0866	-0.0823	-0.0765	...
5	-0.0752	-0.0739	-0.0722	-0.0699	-0.0668	-0.0625	...
10	-0.0372	-0.0369	-0.0363	-0.0356	-0.0345	-0.0329	...
15	-0.0247	-0.0245	-0.0243	-0.0239	-0.0233	-0.0224	...

$n\backslash\rho$	-0.40	-0.30	-0.20	-0.10	0.00	0.10	0.20	...
2	-0.1205	-0.0986	-0.0719	-0.0394	0.0000	0.0477	0.1050	...
3	-0.0876	-0.0734	-0.0550	-0.0311	0.0000	0.0402	0.0916	...
4	-0.0688	-0.0586	-0.0447	-0.0258	0.0000	0.0351	0.0825	...
5	-0.0568	-0.0488	-0.0378	-0.0222	0.0000	0.0315	0.0758	...
10	-0.0305	-0.0270	-0.0217	-0.0134	0.0000	0.0219	0.0577	...
15	-0.0210	-0.0188	-0.0154	-0.0097	0.0000	0.0174	0.0489	...

$n\backslash\rho$	0.30	0.40	0.50	0.60	0.70	0.80	0.90	1.00
2	0.1733	0.2536	0.3467	0.4529	0.5720	0.7036	0.8466	1.0000
3	0.1562	0.2357	0.3305	0.4403	0.5639	0.6996	0.8455	1.0000
4	0.1448	0.2239	0.3202	0.4327	0.5593	0.6974	0.8450	1.0000
5	0.1364	0.2154	0.3131	0.4276	0.5562	0.6961	0.8447	1.0000
10	0.1135	0.1931	0.2952	0.4156	0.5495	0.6932	0.8440	1.0000
15	0.1025	0.1830	0.2878	0.4109	0.5470	0.6921	0.8437	1.0000

Appendix D

Implementation Details

D.1 Black-Scholes 2D PDE to Heat Equation

In the main text we presented the short version of this transformation but here is the full, more intuitive version:

Start with the two-dimensional Black-Scholes PDE for $V = V(A, B, t)$:

$$V_t + \frac{1}{2}\sigma_A^2 A^2 V_{AA} + \rho\sigma_A\sigma_B ABV_{AB} + \frac{1}{2}\sigma_B^2 B^2 V_{BB} + rAV_A + rBV_B - rV = 0 \qquad \text{(D.1)}$$

with terminal condition for a Put

$$V(A, B, T) = f(A, B) = (K - A - B)^+ \qquad \text{(D.2)}$$

and boundary conditions from the one-dimensional Black-Scholes Put:

$$V(0, B, t) = g_0(B, t) = Ke^{-r(T-t)}\mathcal{N}(-d_2(B)) - B\mathcal{N}(-d_1(B)) \qquad \text{(D.3a)}$$

$$V(A, 0, t) = h_0(A, t) = Ke^{-r(T-t)}\mathcal{N}(-d_2(A)) - A\mathcal{N}(-d_1(A)) \qquad \text{(D.3b)}$$

$$V(\infty, B, t) = g_\infty(B, t) = 0 \qquad \text{(D.3c)}$$

$$V(A, \infty, t) = h_\infty(A, t) = 0 \qquad \text{(D.3d)}$$

where

$$d_1(*) = \frac{\ln(*/K) + (r + \frac{\sigma_*^2}{2})(T - t)}{\sigma_* \sqrt{T - t}} \qquad \text{(D.3e)}$$

$$d_2(*) = d_1(*) - \sigma_* \sqrt{T - t} = \frac{\ln(*/K) + (r - \frac{\sigma_*^2}{2})(T - t)}{\sigma_* \sqrt{T - t}} \qquad \text{(D.3f)}$$

$$* = A, B$$

This PDE is backward in time, has cross derivatives, first order derivatives, coefficients that vary spatially, and a non-derivative term. We will make substitutions to simplify the PDE before we develop the ADI method to solve it numerically.

D.1.1 τ

The first substitution that we make will change the PDE from forward in time, to backward in time:

$$\tau = T - t \tag{D.4}$$

So $V(A, B, t) = V(A, B, T - \tau) = \bar{V}(A, B, \tau) = \bar{V}$,

$$\frac{\partial}{\partial t} = \frac{\partial \tau}{\partial t}\frac{\partial}{\partial \tau} = -\frac{\partial}{\partial \tau} \tag{D.5}$$

and the spatial derivatives remain the same. The terminal condition becomes an initial condition:

$$\bar{V}(A, B, 0) = f(A, B) = (K - A - B)^+ \tag{D.6}$$

and the boundary conditions become

$$\bar{V}(0, B, \tau) = \bar{g}_0(B, \tau) = Ke^{-r\tau}\mathcal{N}(-\bar{d}_2(B)) - B\mathcal{N}(-\bar{d}_1(B)) \tag{D.7a}$$
$$\bar{V}(A, 0, \tau) = \bar{h}_0(A, \tau) = Ke^{-r\tau}\mathcal{N}(-\bar{d}_2(A)) - A\mathcal{N}(-\bar{d}_1(A)) \tag{D.7b}$$
$$\bar{V}(\infty, B, \tau) = \bar{g}_\infty(B, \tau) = 0 \tag{D.7c}$$
$$\bar{V}(A, \infty, \tau) = \bar{h}_\infty(A, \tau) = 0 \tag{D.7d}$$

where

$$\bar{d}_1(*) = \frac{\ln(*/K) + (r + \frac{\sigma_*^2}{2})\tau}{\sigma_* \sqrt{\tau}} \tag{D.7e}$$

$$\bar{d}_2(*) = \bar{d}_1(*) - \sigma_* \sqrt{\tau} = \frac{\ln(*/K) + (r - \frac{\sigma_*^2}{2})\tau}{\sigma_* \sqrt{\tau}} \tag{D.7f}$$

$$* = A, B$$

So the PDE becomes:

$$\frac{\partial \bar{V}}{\partial \tau} = \frac{1}{2}\sigma_A^2 A^2\frac{\partial^2 \bar{V}}{\partial A^2} + \rho\sigma_A\sigma_B AB\frac{\partial^2 \bar{V}}{\partial A\partial B} + \frac{1}{2}\sigma_B^2 B^2\frac{\partial^2 \bar{V}}{\partial B^2} + rA\frac{\partial \bar{V}}{\partial A} + rB\frac{\partial \bar{V}}{\partial B} - r\bar{V} \tag{D.8}$$

D.1.2 ν

To have the $r\bar{V}$ term cancel we will next let

$$\bar{\nu} = \bar{\nu}(A, B, \tau) = \frac{1}{K}e^{r\tau}\bar{V}(A, B, \tau) \tag{D.9a}$$

$$\bar{V}(A, B, \tau) = Ke^{-r\tau}\bar{\nu}(A, B, \tau) \tag{D.9b}$$

The initial condition becomes:

$$\bar{\nu}(A, B, 0) = \frac{1}{K}\bar{V}(A, B, 0) = \frac{1}{K}f(A, B) = \frac{1}{K}(K - A - B)^+ \tag{D.10}$$

and the boundary conditions:

$$\bar{v}(0, B, \tau) = \frac{e^{r\tau}}{K}\bar{V}(0, B, \tau) = \frac{e^{r\tau}}{K}\bar{g}_0(B, \tau) = \mathcal{N}(-d_2(B)) - \frac{B}{K}e^{r\tau}\mathcal{N}(-d_1(B)) \tag{D.11a}$$

$$\bar{v}(A, 0, \tau) = \frac{e^{r\tau}}{K}\bar{V}(A, 0, \tau) = \frac{e^{r\tau}}{K}\bar{h}_0(A, \tau) = \mathcal{N}(-d_2(A)) - \frac{A}{K}e^{r\tau}\mathcal{N}(-d_1(A)) \tag{D.11b}$$

$$\bar{v}(\infty, B, \tau) = \frac{1}{K}e^{r\tau}\bar{V}(\infty, B, \tau) = \frac{1}{K}e^{r\tau}\bar{g}_\infty(B, \tau) = 0 \tag{D.11c}$$

$$\bar{v}(A, \infty, \tau) = \frac{1}{K}e^{r\tau}\bar{V}(A, \infty, \tau) = \frac{1}{K}e^{r\tau}\bar{h}_\infty(A, \tau) = 0 \tag{D.11d}$$

where

$$\bar{d}_1(*) = \frac{\ln(*/K) + (r + \frac{\sigma_*^2}{2})\tau}{\sigma_* \sqrt{\tau}} \tag{D.11e}$$

$$\bar{d}_2(*) = \bar{d}_1(*) - \sigma_* \sqrt{\tau} = \frac{\ln(*/K) + (r - \frac{\sigma_*^2}{2})\tau}{\sigma_* \sqrt{\tau}} \tag{D.11f}$$

$$* = A, B$$

The derivatives change a little too:

$$\frac{\partial \bar{V}}{\partial \tau} = \frac{\partial}{\partial \tau}(Ke^{-r\tau}\bar{v}) = Ke^{-r\tau}\frac{\partial \bar{v}}{\partial \tau} - rKe^{-r\tau}\bar{v} = Ke^{-r\tau}\frac{\partial \bar{v}}{\partial \tau} - r\bar{V} \tag{D.12a}$$

$$\frac{\partial \bar{V}}{\partial *} = Ke^{-r\tau}\frac{\partial \bar{v}}{\partial *} \text{ and } \frac{\partial^2 \bar{V}}{\partial *^2} = Ke^{-r\tau}\frac{\partial^2 \bar{v}}{\partial *^2} \tag{D.12b}$$

$$\text{for } * = A, B$$

so the PDE is now:

$$Ke^{-r\tau}\frac{\partial \bar{v}}{\partial \tau} - r\bar{V} = \frac{1}{2}\sigma_A^2 A^2\frac{\partial^2 \bar{V}}{\partial A^2} + \rho\sigma_A\sigma_B AB\frac{\partial^2 \bar{V}}{\partial A\partial B} + \frac{1}{2}\sigma_B^2 B^2\frac{\partial^2 \bar{V}}{\partial B^2}$$

$$+ rA\frac{\partial \bar{V}}{\partial A} + rB\frac{\partial \bar{V}}{\partial B} - r\bar{V}$$

$$\frac{\cancel{Ke^{-r\tau}}}{\cancel{Ke^{-r\tau}}}\frac{\partial \bar{v}}{\partial \tau} = \frac{1}{2}\sigma_A^2 A^2\frac{\partial^2 \bar{v}}{\partial A^2} + \rho\sigma_A\sigma_B AB\frac{\partial^2 \bar{v}}{\partial A\partial B} + \frac{1}{2}\sigma_B^2 B^2\frac{\partial^2 \bar{v}}{\partial B^2} + rA\frac{\partial \bar{v}}{\partial A} + rB\frac{\partial \bar{v}}{\partial B}$$

$$\frac{\partial \bar{v}}{\partial \tau} = \frac{1}{2}\sigma_A^2 A^2\frac{\partial^2 \bar{v}}{\partial A^2} + \rho\sigma_A\sigma_B AB\frac{\partial^2 \bar{v}}{\partial A\partial B} + \frac{1}{2}\sigma_B^2 B^2\frac{\partial^2 \bar{v}}{\partial B^2} + rA\frac{\partial \bar{v}}{\partial A} + rB\frac{\partial \bar{v}}{\partial B} \tag{D.13}$$

D.1.3 α and β

Now we will make a substitution for A and B that will result in constant parameters:

$$\alpha = \ln(A/K) \text{ and } \beta = \ln(B/K) \tag{D.14a}$$

or

$$A = Ke^{\alpha} \text{ and } B = Ke^{\beta} \tag{D.14b}$$

So $\bar{v}(A, B, \tau) = \bar{v}(Ke^{\alpha}, Ke^{\beta}, \tau) = \tilde{v}(\alpha, \beta, \tau) = \tilde{v}$, and the initial condition becomes:

$$\tilde{v}(\alpha, \beta, 0) = \bar{v}(Ke^{\alpha}, Ke^{\beta}, 0) = \frac{1}{K} f(Ke^{\alpha}, Ke^{\beta}) = (1 - e^{\alpha} - e^{\beta})^{+} \tag{D.15}$$

And the boundary conditions become:

$$\tilde{v}(-\infty, \beta, \tau) = \bar{v}(-\infty, Ke^{\beta}, \tau) = \frac{e^{r\tau}}{K} \bar{g}_0(Ke^{\beta}, \tau)$$

$$= \mathcal{N}(-\tilde{d}_2(\beta)) - e^{\beta} e^{r\tau} \mathcal{N}(-\tilde{d}_1(\beta)) \tag{D.16a}$$

$$\tilde{v}(\alpha, -\infty, \tau) = \bar{v}(Ke^{\alpha}, -\infty, \tau) = \frac{e^{r\tau}}{K} \bar{h}_0(Ke^{\alpha}, \tau)$$

$$= \mathcal{N}(-\tilde{d}_2(\alpha)) - e^{\alpha} e^{r\tau} \mathcal{N}(-\tilde{d}_1(\alpha)) \tag{D.16b}$$

$$\tilde{v}(\infty, \beta, \tau) = \bar{v}(\infty, Ke^{\beta}, \tau) = \frac{1}{K} e^{r\tau} \bar{g}_{\infty}(Ke^{\beta}, \tau) = 0 \tag{D.16c}$$

$$\tilde{v}(\alpha, \infty, \tau) = \bar{v}(Ke^{\alpha}, \infty, \tau) = \frac{1}{K} e^{r\tau} \bar{h}_{\infty}(Ke^{\alpha}, \tau) = 0 \tag{D.16d}$$

where

$$\tilde{d}_1(\star) = \frac{\star + (r + \frac{\sigma_*^2}{2})\tau}{\sigma_* \sqrt{\tau}} \tag{D.16e}$$

$$\tilde{d}_2(\star) = \tilde{d}_1 - \sigma_* \sqrt{\tau} = \frac{\star + (r - \frac{\sigma_*^2}{2})\tau}{\sigma_* \sqrt{\tau}} \tag{D.16f}$$

$$\star = \alpha, \beta \text{ and } * = Ke^{\star} \text{ so } * = A, B$$

And the derivatives change too:

$$\frac{\partial}{\partial A} = \frac{\partial \alpha}{\partial A} \frac{\partial}{\partial \alpha} = \frac{1}{A} \frac{\partial}{\partial \alpha} \tag{D.17a}$$

$$\frac{\partial^2}{\partial A^2} = \frac{\partial}{\partial A}\left(\frac{1}{A} \frac{\partial}{\partial \alpha}\right) = -\frac{1}{A^2} \frac{\partial}{\partial \alpha} + \frac{1}{A} \frac{\partial}{\partial A}\left(\frac{\partial}{\partial \alpha}\right) = -\frac{1}{A^2} \frac{\partial}{\partial \alpha} + \frac{1}{A^2} \frac{\partial^2}{\partial \alpha^2} \tag{D.17b}$$

$$\frac{\partial}{\partial B} = \frac{\partial \beta}{\partial B} \frac{\partial}{\partial \beta} = \frac{1}{B} \frac{\partial}{\partial \beta} \tag{D.17c}$$

$$\frac{\partial^2}{\partial B^2} = \frac{\partial}{\partial B}\left(\frac{1}{B} \frac{\partial}{\partial \beta}\right) = -\frac{1}{B^2} \frac{\partial}{\partial \beta} + \frac{1}{B} \frac{\partial}{\partial B}\left(\frac{\partial}{\partial \beta}\right) = -\frac{1}{B^2} \frac{\partial}{\partial \beta} + \frac{1}{B^2} \frac{\partial^2}{\partial \beta^2} \tag{D.17d}$$

$$\frac{\partial^2}{\partial A \partial B} = \frac{\partial}{\partial A}\left(\frac{1}{B}\frac{\partial}{\partial \beta}\right) = \frac{1}{A}\frac{\partial}{\partial \alpha}\frac{1}{B}\frac{\partial}{\partial \beta} = \frac{1}{AB}\frac{\partial^2}{\partial \alpha \partial \beta} \tag{D.17e}$$

Substituting these into the PDE we get:

$$\frac{\partial \tilde{v}}{\partial \tau} = \frac{1}{2}\sigma_A^2 A^2\left(-\frac{1}{A^2}\frac{\partial \tilde{v}}{\partial \alpha} + \frac{1}{A^2}\frac{\partial^2 \tilde{v}}{\partial \alpha^2}\right) + \rho\sigma_A\sigma_B AB\frac{1}{AB}\frac{\partial^2 \tilde{v}}{\partial \alpha \partial \beta}$$

$$+\frac{1}{2}\sigma_B^2 B^2\left(-\frac{1}{B^2}\frac{\partial \tilde{v}}{\partial \beta} + \frac{1}{B^2}\frac{\partial^2 \tilde{v}}{\partial \beta^2}\right) + rA\frac{1}{A}\frac{\partial \tilde{v}}{\partial \alpha} + rB\frac{1}{B}\frac{\partial \tilde{v}}{\partial \beta}$$

$$\frac{\partial \tilde{v}}{\partial \tau} = \frac{\sigma_A^2}{2}\left(-\frac{\partial \tilde{v}}{\partial \alpha} + \frac{\partial^2 \tilde{v}}{\partial \alpha^2}\right) + \rho\sigma_A\sigma_B\frac{\partial^2 \tilde{v}}{\partial \alpha \partial \beta} + \frac{\sigma_B^2}{2}\left(-\frac{\partial \tilde{v}}{\partial \beta} + \frac{\partial^2 \tilde{v}}{\partial \beta^2}\right) + r\frac{\partial \tilde{v}}{\partial \alpha} + r\frac{\partial \tilde{v}}{\partial \beta}$$

$$\frac{\partial \tilde{v}}{\partial \tau} = \frac{\sigma_A^2}{2}\frac{\partial^2 \tilde{v}}{\partial \alpha^2} + \rho\sigma_A\sigma_B\frac{\partial^2 \tilde{v}}{\partial \alpha \partial \beta} + \frac{\sigma_B^2}{2}\frac{\partial^2 \tilde{v}}{\partial \beta^2} + \left(r - \frac{\sigma_A^2}{2}\right)\frac{\partial \tilde{v}}{\partial \alpha} + \left(r - \frac{\sigma_B^2}{2}\right)\frac{\partial \tilde{v}}{\partial \beta} \tag{D.18}$$

D.1.4 x and y

Finally, to cancel out the first derivative terms, we will let:

$$x = \alpha + \left(r - \frac{\sigma_A^2}{2}\right)\tau \text{ and } y = \beta + \left(r - \frac{\sigma_B^2}{2}\right)\tau \tag{D.19a}$$

or

$$\alpha = x - \left(r - \frac{\sigma_A^2}{2}\right)\tau \text{ and } \beta = y - \left(r - \frac{\sigma_B^2}{2}\right)\tau \tag{D.19b}$$

So $v = v(x,y,\tau) = \tilde{v}(x - (r - \frac{\sigma_A^2}{2})\tau, y - (r - \frac{\sigma_B^2}{2})\tau, \tau) = \tilde{v}(\alpha,\beta,\tau)$. The initial condition becomes:

$$v(x,y,0) = \tilde{v}(x,y,0) = \frac{1}{K}f(Ke^x, Ke^y) = (1 - e^x - e^y)^+ \tag{D.20}$$

And the boundary conditions become:

$$v(-\infty, y, \tau) = \tilde{v}(-\infty, y - (r - \frac{\sigma_B^2}{2})\tau, \tau) = \frac{1}{K}e^{r\tau}\bar{g}_0(Ke^{y-(r-\frac{\sigma_B^2}{2})\tau}, \tau)$$

$$= \mathcal{N}(-d_2^\star(y)) - e^{y+\frac{\sigma_B^2}{2}\tau}\mathcal{N}(-d_1^\star(y)) \tag{D.21a}$$

$$v(x, -\infty, \tau) = \tilde{v}(x - (r - \frac{\sigma_A^2}{2})\tau, -\infty, \tau) = \frac{1}{K}e^{r\tau}\bar{h}_0(Ke^{x-(r-\frac{\sigma_A^2}{2})\tau}, \tau)$$

$$= \mathcal{N}(-d_2^\star(x)) - e^{x+\frac{\sigma_A^2}{2}\tau}e^{r\tau}\mathcal{N}(-d_1^\star(x)) \tag{D.21b}$$

$$v(\infty, y, \tau) = \tilde{v}(\infty, y - (r - \frac{\sigma_B^2}{2})\tau, \tau) = \frac{1}{K}e^{r\tau}\bar{g}_\infty(Ke^{y-(r-\frac{\sigma_B^2}{2})\tau}, \tau) = 0 \tag{D.21c}$$

$$v(x, \infty, \tau) = \tilde{v}(x - (r - \frac{\sigma_A^2}{2})\tau, \infty, \tau) = \frac{1}{K}e^{r\tau}\bar{h}_\infty(Ke^{x-(r-\frac{\sigma_A^2}{2})\tau}, \tau) = 0 \tag{D.21d}$$

where

$$d_1^\star(\bullet) = \frac{\bullet + \sigma_*^2 \tau}{\sigma_* \sqrt{\tau}} \tag{D.21e}$$

$$d_2^\star(\bullet) = d_1^\star(\bullet) - \sigma_* \sqrt{\tau} = \frac{\bullet}{\sigma_* \sqrt{\tau}} \tag{D.21f}$$

$$\bullet = x, y; \text{ and } * = A, B$$

And the derivatives change too. Since $\tilde{v}(\alpha, \beta, \tau) = v(x(\tau), y(\tau), \tau)$ we use the chain rule for the derivative with respect to τ:

$$\frac{\partial}{\partial \tau} = \frac{\partial}{\partial \tau} + \frac{\partial x}{\partial \tau}\frac{\partial}{\partial x} + \frac{\partial y}{\partial \tau}\frac{\partial}{\partial y} = \frac{\partial}{\partial \tau} + (r - \frac{\sigma_A^2}{2})\frac{\partial}{\partial \alpha} + (r - \frac{\sigma_B^2}{2})\frac{\partial}{\partial \beta} \tag{D.22a}$$

$$\frac{\partial}{\partial *} = \frac{\partial}{\partial \star} \tag{D.22b}$$

$$\frac{\partial^2}{\partial *^2} = \frac{\partial^2}{\partial \star^2} \tag{D.22c}$$

for $* = x, y$ and $\star = \alpha, \beta$

So the PDE simplifies :

$$\frac{\partial v}{\partial \tau} + \cancel{(r - \frac{\sigma_A^2}{2})\frac{\partial v}{\partial x}} + \cancel{(r - \frac{\sigma_B^2}{2})\frac{\partial v}{\partial y}} = \frac{\sigma_A^2}{2}\frac{\partial^2 v}{\partial x^2} + \rho\sigma_A\sigma_B\frac{\partial^2 v}{\partial x \partial y} + \frac{\sigma_B^2}{2}\frac{\partial^2 v}{\partial y^2}$$

$$+ \cancel{(r - \frac{\sigma_A^2}{2})\frac{\partial v}{\partial x}} + \cancel{(r - \frac{\sigma_B^2}{2})\frac{\partial v}{\partial y}}$$

$$\frac{\partial v}{\partial \tau} = \frac{\sigma_A^2}{2}\frac{\partial^2 v}{\partial x^2} + \rho\sigma_A\sigma_B\frac{\partial^2 v}{\partial x \partial y} + \frac{\sigma_B^2}{2}\frac{\partial^2 v}{\partial y^2} \tag{D.23}$$

D.2   The Thomas Algorithm

The Thomas Algorithm is a linear complexity solution to tridiagonal matrix equations.

If we have a tridiagonal matrix equation $Aw = d$ of the form:

$$\begin{bmatrix} b_1 & c_1 & 0 & \dots & \dots & \dots & 0 \\ a_2 & b_2 & c_2 & 0 & \dots & \dots & 0 \\ 0 & a_3 & b_3 & c_3 & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & a_{M-3} & b_{M-3} & c_{M-3} & 0 \\ 0 & 0 & \dots & 0 & a_{M-2} & b_{M-2} & c_{M-2} \\ 0 & 0 & 0 & \dots & 0 & a_{M-1} & b_{M-1} \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ \vdots \\ \vdots \\ w_{M-2} \\ w_{M-1} \end{bmatrix} = \begin{bmatrix} d_1 - a_1 w_0 \\ d_2 \\ \vdots \\ \vdots \\ \vdots \\ d_{M-2} \\ d_{M-1} - c_{M-1} w_M \end{bmatrix} \tag{D.24}$$

This is equivalent to the system of equations that we will start with:

$$a_i w_{i-1} + b_i w_i + c_i w_{i+1} = d_i \text{ for } i = 1, ..., M - 1 \tag{D.25}$$

with boundary conditions:

$$w_0 = \beta_0 \text{ and } w_M = \beta_M \tag{D.26}$$

By Gaussian elimination we get a solution of the form:

$$w_i = p_{i+1} w_{i+1} + q_{i+1} \tag{D.27}$$

But we need to find p and q so we substitute this in to our system for w_{i-1}:

$$a_i(p_i w_i + q_i) + b_i w_i + c_i w_{i+1} = d_i \text{ for } i = 1, ..., M - 1 \tag{D.28}$$

Which we can rearrange to get:

$$w_i = \frac{-c_i}{a_i p_i + b_i} w_{i+1} + \frac{d_i - a_i q_i}{a_i p_i + b_i} \tag{D.29}$$

So we have

$$p_{i+1} = \frac{-c_i}{a_i p_i + b_i} \tag{D.30a}$$

and

$$q_{i+1} = \frac{d_i - a_i q_i}{a_i p_i + b_i} \tag{D.30b}$$

So if we know p_1 and q_1 then we can find all other p_is and q_is, and if we have w_M then we can find all other w_i values too.

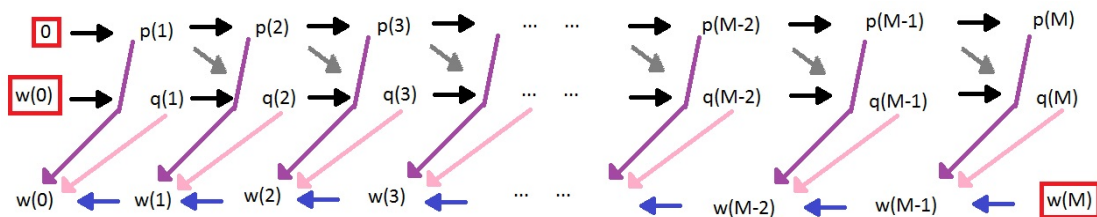From the lower boundary condition, $w_0 = \beta_0$, and

$$w_0 = p_1 w_1 + q_1 \tag{D.31}$$

we can see that $p_1 = 0$ and $q_1 = \beta_0$ follow.

Different types of boundary conditions would result in different p_1 and q_1 values.

To be well conditioned we need $|p_i| \leq 1$

To help understand this process follow the data flow in the following diagram:

We start by knowing $w(0)$ and $w(M)$, in the red boxes, as well as all values of a, b, c, and d. Then the data flows from left to right along the black and grey arrows through p and q. Second the data flows from right to left through w.

Here is some pseudocode for the case where a, b, and c are constant but d is not - this is the case used in this paper.

```
set a, b, c
p(1) = 0, q(1) = beta0
for i = 1 to M-1
    set d
    p(i+1) = -c / ( a*p(i) + b )
    q(i+1) = ( d - a*q(i) ) / ( a*p(i) + b )

w(M) = betaM
for i = M-1 to 0
    w(i) = p(i+1)*w(i+1) + q(i+1)
```

Appendix E

Distributions

In their 1979 paper, "Discretely Adjusted Option Hedges"[10], Boyle and Emanuel showed that Black-Scholes option hedging profits are distributed $\chi^2(1) - 1$ multiplied by a deterministic function based on the model's parameters and the asset value at the begining of the hedging time step. In the body of this paper we showed hedging profits that were obviously not distributed $\chi^2(1) - 1$; they appeared to be normally distributed. This is because of the central limit theorem. What was shown above was the hedging profits for the life of the option and what Boyle and Emanual examined was a single period hedging profit. Since the option price and delta are Markov the time steps are independent and so are the hedging profits at subsequent time steps. Therefore the above hedging profits are distributed as a sum of independent $\chi^2(1) - 1$ random variables; from the central limit theorem for a sufficiently large number for time steps it will be normally distributed.

In our examples we used a time to maturity of 1 year rebalanced daily. In figure E.1 it is obvious that a single time step's hedging profits, or $\chi^2(1)$, will look very different from the normal distribution, but in figure E.2 it can be seen that the hedging profits from a year of hedging rebalanced daily, or $\chi^2(250)$ is distributed very close to the normal distribution. This is the reason that all of our hedging profits appear to be normally distributed whereas hedging profits have been proven to be distributed $\chi^2(1)$ times some model parameters.

By the hedging methods used in this paper we reproduced the results of Boyle and Emanuel. In figures E.3 and E.4 it can be seen that the hedging profits on a vanilla put for a single period appear to be distributed $\chi^2(1)$ shifted to the left. Figure E.5 is a QQ plot of the single period hedging profits from figure E.3 and $\chi^2(1)$, they are identical to 3 standard deviations and very similar to 7, with data very sparse beyond that point.
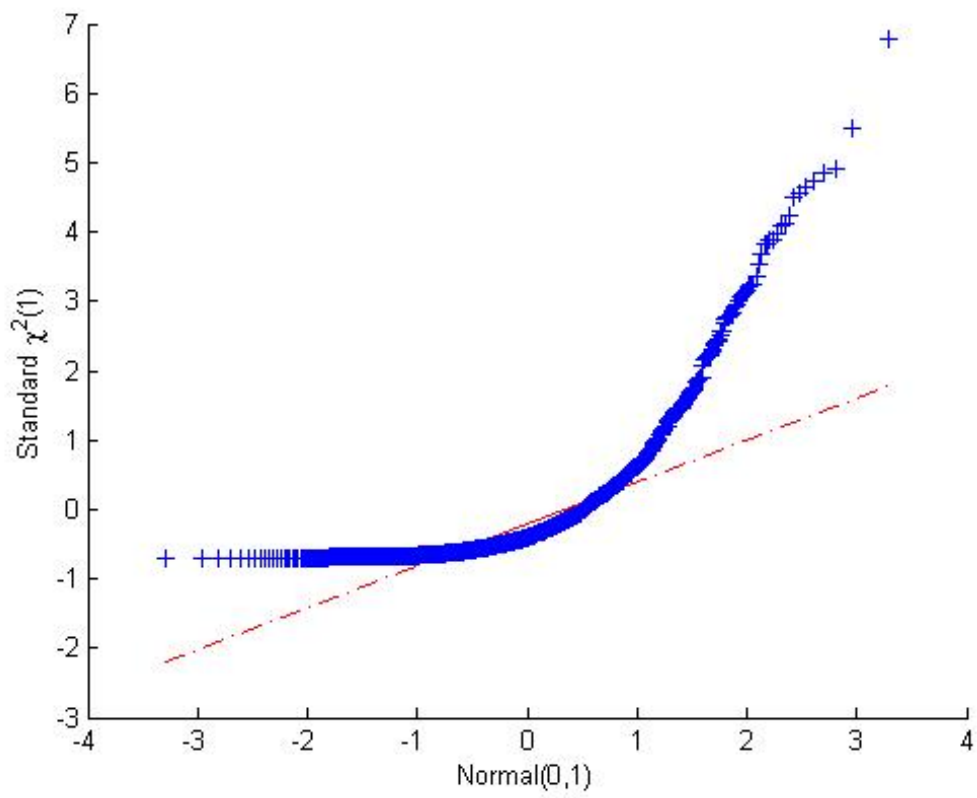
Figure E.1: QQ plot of Normal against $\chi^2(1)$, sample size of 10,000
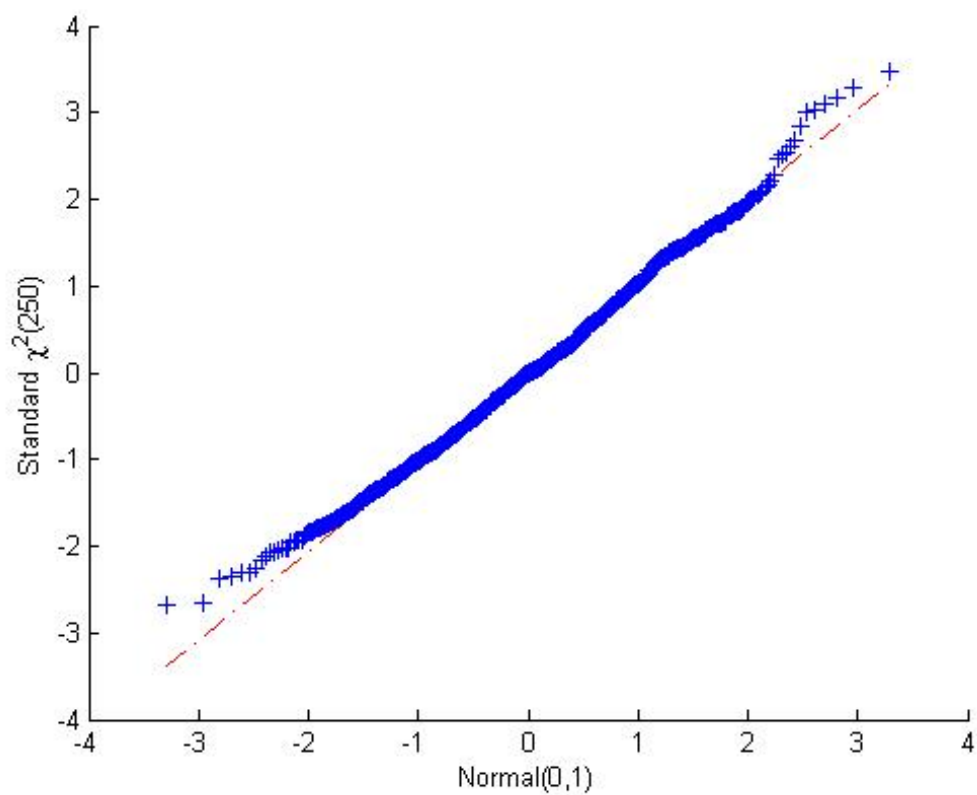
Figure E.2: QQ plot of Normal against $\chi^2(250)$, sample size of 10,000

Figure E.3: Histogram of a Single Period Hedging Profits from Black-Scholes Pricing of a Vanilla Put, sample size of 10,000
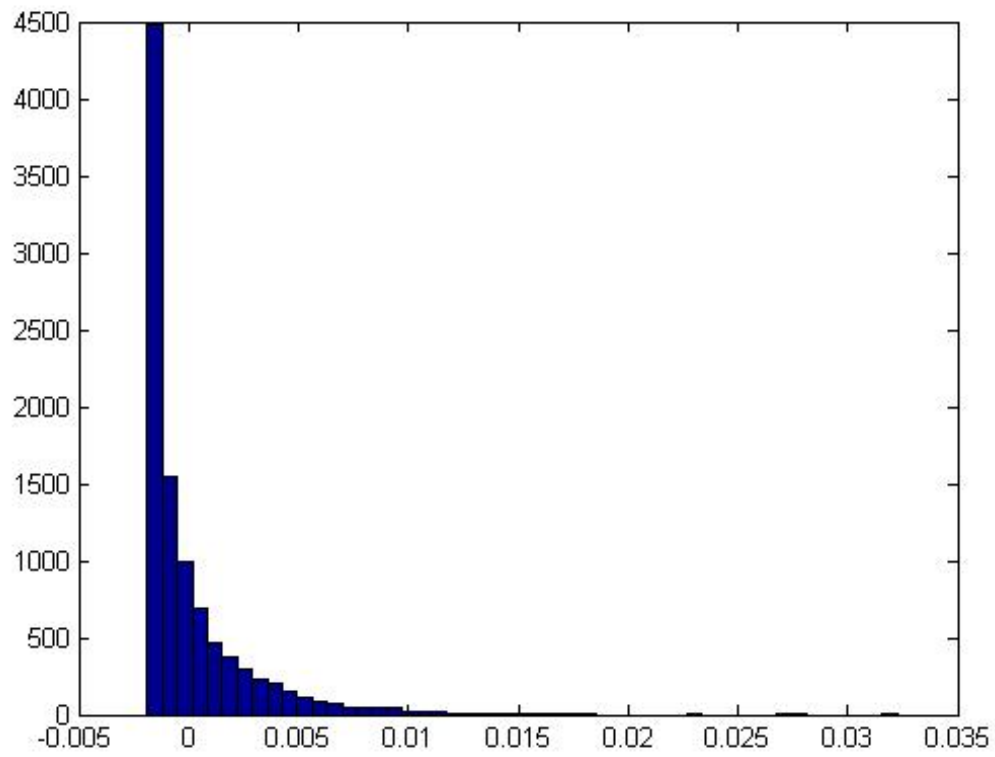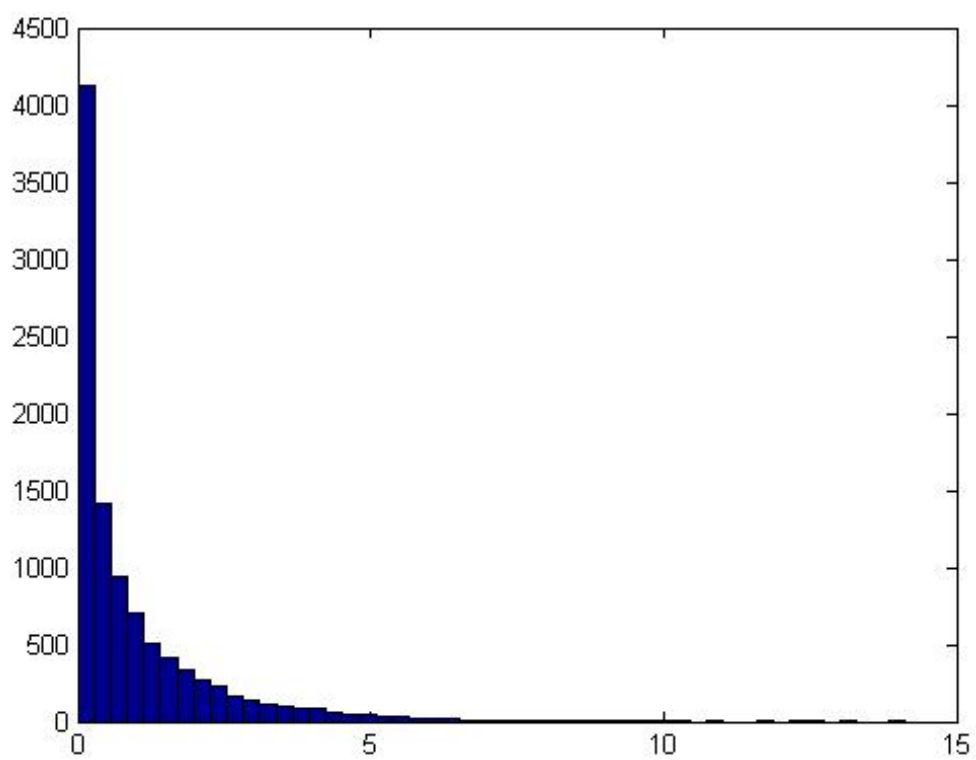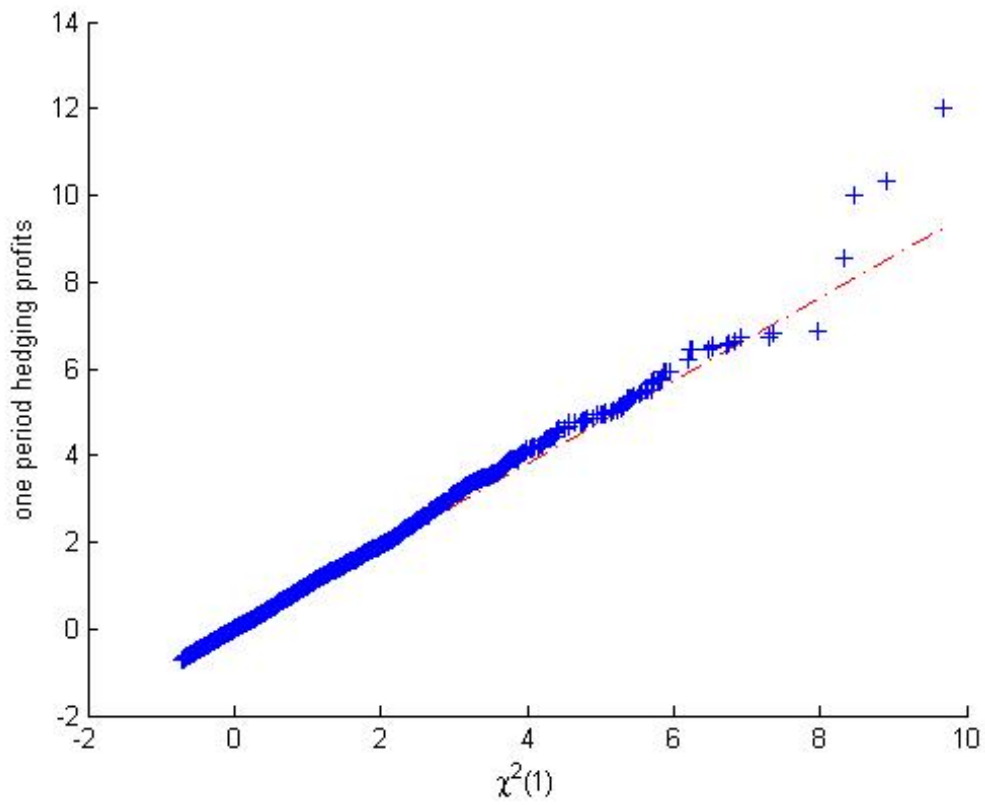
Figure E.4: Histogram of $\chi^2(1)$, sample size of 10,000

Figure E.5: QQ plot of $\chi^2(1)$ and Single Period Hedging Profits of a Vanilla Put, sample size of 10,000

Appendix F

Sample Codes

The parameters required for all of the pricing surface are:

$$
\begin{array}{rcl}
sA & - & \text{sigma for A, } \sigma_A \\
sB & - & \text{sigma for B, } \sigma_A \\
rho & - & \text{correlation between drivers of A and B, } \rho \\
r & - & \text{risk free rate of return} \\
K & - & \text{strike price} \\
T & - & \text{time to maturity} \\
f & - & \text{size of price steps in A} \\
g & - & \text{size of price steps in B} \\
h & - & \text{size of time steps}
\end{array}
$$

F.1   ADI Method

Here we show a code sample to implement the ADI numerical method from the simplified heat equation. Given the inputs:

We can work through the ADI method to solve for the pricing surface of a put:

```
% Define false upper boundary, 3 standard deviations
Amax = max(ceil(K*exp(-(r-sA*sA/2)*T+3*sA*sqrt(T))),K+1);
Bmax = max(ceil(K*exp(-(r-sB*sB/2)*T+3*sB*sqrt(T))),K+1);

% Number of steps in the A and B directions
M = Amax/f;
N = Bmax/g;

% Initialize variables
A = 0:f:Amax;
B = 0:g:Bmax;
V = zeros(T/h+1,M+1,N+1);
vv = zeros(M+1,N+1);  % v_tilde
uu = zeros(M+1,N+1);  % u_bar
```

```
% Fill in terminal condition
tau = 0;
x = log(A/K) + (r-sA*sA/2)*tau;
y = log(B/K) + (r-sB*sB/2)*tau;
for i = 1:1+M
    for j = 1:1+N
        V(1,i,j) = TC(x(i),y(j),K);
    end
end


% PRIMER STEP ---------------------------------------------------
tau = h;
x = log(A/K) + (r-sA*sA/2)*tau;
y = log(B/K) + (r-sB*sB/2)*tau;

% B = 0 border
for i = 1:M
    V(2,i,1) = BS(x(i),K,tau,sA,r);
end
% A = 0 border
for j = 2:N
    V(2,1,j) = BS(y(j),K,tau,sB,r);
end
% B = Bmax border
for i = 1:1+M
    V(2,i,N+1) = 0;
end
% A = Amax border
for j = 1:N
    V(2,M+1,j) = 0;
end
% internal points - from FTCS scheme
for i = 2:M
    for j = 2:N
        temp = (1-((h*sA*sA)/(f*f))-((h*sB*sB)/(g*g)))*V(1,i,j);
        temp = temp + (h*sA*sA/(2*f*f))*(V(1,i+1,j)+V(1,i-1,j));
        temp = temp +  (h*sB*sB/(2*g*g))*(V(1,i,j+1)+V(1,i,j-1));
        V(2,i,j) = temp + (h*rho*sA*sB/(2*f*g))*(V(1,i+1,j+1)...
                            -V(1,i-1,j+1)-V(1,i+1,j-1)+V(1,i-1,j-1));
    end
end
```

```
% WORK BACK IN TIME (forward in tau) --------------------------------------
for n = 2:T/h
    tau = n*h;
    x = log(A/K) + (r-sA*sA/2)*tau;
    y = log(B/K) + (r-sB*sB/2)*tau;

    % B = 0 border
    for i = 1:M
        V(n+1,i,1) = BS(x(i),K,tau,sA,r);
    end
    % A = 0 border
    for j = 2:N
        V(n+1,1,j) = BS(y(j),K,tau,sB,r);
    end
    % B = Bmax border
    for i = 1:1+M
        V(n+1,i,N+1) = 0;
    end
    % A = Amax border
    for j = 1:N
        V(n+1,M+1,j) = 0;
    end

    % u_bar (aka uu) - for mixed derivative
    for i = 1:1+M
        for j = 1:1+N
            uu(i,j) = 3*V(n,i,j)/2-V(n-1,i,j)/2;
        end
    end

    % intermediate step's boundary condition
    j = 2:N;
    vv(1,j) = squeeze(V(n+1,1,j)- (h*sB*sB/(4*g*g))...
                        *(V(n+1,1,j+1) -2*V(n+1,1,j) + V(n+1,1,j-1)));
    vv(1,j) = vv(1,j) + squeeze(V(n,1,j)+ (h*sB*sB/(4*g*g))...
                        *(V(n,1,j+1) -2*V(n,1,j) + V(n,1,j-1)))';
    vv(1,j) = vv(1,j)/2;
    vv(M+1,j) = squeeze(V(n+1,M+1,j)- (h*sB*sB/(4*g*g))...
                        *(V(n+1,M+1,j+1) -2*V(n+1,M+1,j) + V(n+1,M+1,j-1)));
    vv(M+1,j) = vv(M+1,j) + squeeze(V(n,M+1,j)+ (h*sB*sB/(4*g*g))...
                        *(V(n,M+1,j+1) -2*V(n,M+1,j) + V(n,M+1,j-1)))';
    vv(M+1,j) = vv(M+1,j)/2;

    % internal points for intermediat step (vv or v_tilde)
    X = (diag(ones(1,M-2),-1)+diag(ones(1,M-2),1))*(-h*sA*sA/(4*f*f))...
```

```
                        +diag(ones(1,M-1))*(1+h*sA*sA/(2*f*f)));
    i = 2:M;
    for j = 2:N
        % set up c and solve [ X*vv(:,j) = c(Vn) ]
        c = squeeze(V(n,i,j)*(1-h*sB*sB/(2*g*g)) ...
                    + (V(n,i,j-1)+V(n,i,j+1))*(h*sB*sB/(4*g*g)));
        c = c+ squeeze((rho*h*sA*sB/(8*f*g))...
                    *(uu(i+1,j+1)-uu(i+1,j-1)-uu(i-1,j+1)+uu(i-1,j-1)))';
        c(1) = c(1) + (h*sA*sA/(4*f*f))*vv(1,j);
        c(M-1) = c(M-1) + (h*sA*sA/(4*f*f))*vv(M+1,j);
        vv(i,j) = matrixEqn(X,c');
    end

    Y = (diag(ones(1,N-2),-1)+diag(ones(1,N-2),1))*(-h*sB*sB/(4*g*g))...
                    +diag(ones(1,N-1)*(1+h*sB*sB/(2*g*g)));
    j = 2:N;
    for i = 2:M
        % set up d and solve [ Y*V(n+1,:,j) = d(vv) ]
        d = squeeze(vv(i,j)*(1-h*sA*sA/(2*f*f)) ...
                    + (vv(i-1,j)+vv(i+1,j))*(h*sA*sA/(4*f*f)));
        d = d+ squeeze((rho*h*sA*sB/(8*f*g))...
                    *( uu(i+1,j+1)-uu(i+1,j-1)-uu(i-1,j+1)+uu(i-1,j-1)));
        d(1) = d(1) + (h*sB*sB/(4*g*g))*V(n+1,i,1);
        d(N-1) = d(N-1) + (h*sB*sB/(4*g*g))*V(n+1,i,N+1);
        V(n+1,i,j) = matrixEqn(Y,d');
    end
end

% transform back from heat equation to option solution (from nu to V)
for n = 0:T/h
    tau = n*h;
    V(n+1,:,:) = K*exp(-r*tau)*V(n+1,:,:); % plot against A&B
end
```

In the above code TC refers to:

```
function v = TC(x,y,K)
    v = max(1-exp(x)-exp(y),0);
end
```

and BS refers to the one-dimensional Black-Scholes solution used on the lower boundaries after modification by the transformations to change the two-dimensional Black-Scholes PDE to the heat equation:

```
function v = BS(xy,K,tau,sig,r)
```

```
    if xy<-1/eps
        v = 1;
    else
        d1 = (xy+sig*sig*tau)/(sig*sqrt(tau));
        d2 = d1 - sig*sqrt(tau);
        v = normcdf(-d2) - exp(xy+sig*sig*tau/2)*normcdf(-d1);
    end
end
```

For details on solving the matrix equations efficiently see appendix D.2 for details on the Thomas Algorithm.

F.1.1 Monte Carlo Simulation

Here we show how the Monte Carlo method is used. Given any value of *A* and *B* and value for the parameter *n*, the number of samples to use in the Monte Carlo simulation, we can simulate the value of V by the below sample code. To create the pricing surface this would be done in two nested loops.

```
    cv = [1,rho;rho,1];
    mu=[0,0];
    XY =  mvnrnd(mu,cv,n);
    Amat = A*exp((r - sA*sA/2)*tau+sA*sqrt(tau)*XY(:,1));
    Bmat = B*exp((r - sB*sB/2)*tau+sB*sqrt(tau)*XY(:,2));
    V = exp(-r*tau)*sum(max(K-Amat-Bmat,0))/n;
```

The process is similar to create the values for the deltas.

F.2 "gamma" Approximation

This is the implementation that was used for the "gamma" approximation:

```
% Define false upper boundary, 3 standard deviations
Amax = max(ceil(K*exp(-(r-sA*sA/2)*T+3*sA*sqrt(T))),K+1);
Bmax = max(ceil(K*exp(-(r-sB*sB/2)*T+3*sB*sqrt(T))),K+1);

% Number of steps in the A and B directions
M = Amax/f;
N = Bmax/g;

% Initialize variables
A = 0:f:Amax;
B = 0:g:Bmax;
V = zeros(T/h+1,M+1,N+1);
```

```matlab
% fill in terminal condition
for i = 1:M+1
    for j = 1:N+1
        V(1,i,j) = TC(A(i)+B(j),K);
    end
end

% work back in time
for n = 1:T/h
    tau = n*h;
    % B = 0 border
    for i = 1:M+1
        V(n+1,i,1) = BS(A(i),K,tau,sA,r);
    end
    % A = 0 border
    for j = 1:N+1
        V(n+1,1,j) = BS(B(j),K,tau,sB,r);
    end

    % Internal points
    for i = 2:M+1
        for j = 2:N+1
            alpha = A(i) / (A(i)+B(j));
            sig = sqrt(sA*sA* alpha*alpha + 2*rho*sA*sB*alpha*(1-alpha)...
+sB*sB*(1-alpha)*(1-alpha));
            V(n+1,i,j) = BS(A(i)+B(j),K,tau,sig,r);
        end
    end
end
```

Curriculum Vitae

Name:	Helen Cheyne
Post-Secondary Education and Degrees:	Wilfrid Laurier Unversity, Waterloo, ON 2007 - 2011, B.Sc. University of Western Ontario, London, ON 2011 - 2013, M.Sc.
Honours and Awards:	Ken McDowell Scholarship, Wilfrid Laurier University 2011 Sun Life Financial Bright Future Award, Wilfrid Laurier University 2009, 2010 Western Graduate Research Scholarship, University of Western Ontario 2011 - 2013
Related Work Experience:	Summer Research Assistant, Bank of Canada 2013 Teaching Assistant, University of Western Ontario 2011 - 2013 Co-op Student, Ontario Teachers' Pension Plan 2009, 2010 and 2011 Instructional Assistant, Wilfrid Laurier Unversity 2008 - 2011

Talks: This work has been accepted to be presented at Applied Mathematics, Modeling and Computational Science (AMMCS) conference in August 2013.