2017

Construction of a zero-coupon yield curve for the Nairobi Securities Exchange and its application in pricing derivatives

Lucy Muthoni
Strathmore Institute of Mathematical Sciences (SIMs)
Strathmore University

Construction of a Zero-Coupon Yield Curve for the Nairobi Securities Exchange and its Application in Pricing Derivatives

Lucy Muthoni

Thesis submitted to the School of Graduate Studies and Institute of Mathematical Sciences, Strathmore University, in partial fulfillment of the requirement for a PhD in Financial Mathematics

Strathmore University
Institute of Mathematical Sciences
P. O. Box: 59857-00200,
Nairobi, Kenya

11^{th} May, 2017

Supervisors:
Prof. Michael Ingleby
Prof. Silas Onyango

Declaration

The work to be submitted has not previously been accepted in substance for any degree and is not concurrently submitted in candidature for any degree. This thesis is a result of my own independent work/investigation, except where otherwise stated. Other sources are acknowledged by explicit references.

Signature of Candidate	Date

Principle Supervisor

This thesis is to be submitted with my approval as the principle supervisor.

Name	Signature	Date

Dean School of Graduate Studies:

Name	Signature	Date

Abstract

Yield curves are used to forecast interest rates for different products when their risk parameters are known, to calibrate no-arbitrage term structure models, and (mostly by investors) to detect whether there is arbitrage opportunity. By yield curve information, investors have opportunity of immunizing/hedging their investment portfolios against financial risks if they have to make an investment with some determined time of maturity. Private sector firms look at yields of different maturities and then choose their borrowing strategy. The differences in yields for long maturity and short maturities are an important indicator for central bank to use in monetary policy process. These differences may show the tightness of the government monetary policy and can be monitored to predict recession in coming years. A lot of research has been done in yield curve modeling and as we will see later in the thesis, most of the models developed had one major shortcoming: non differentiability at the interpolating knot points.

The aim of this thesis is to construct a zero coupon yield curve for Nairobi Securities Exchange, and use the risk- free rates to price derivatives, with particular attention given to pricing coffee futures. This study looks into the three methods of constructing yield curves: by use of spline-based models, by interpolation and by using parametric models. We suggest an improvement in the interpolation methods used in the most celebrated spline-based model, monotonicity-preserving interpolation on r(t). We also use operator form of numerical differentiation to estimate the forward rates at the knot points, at which points the spot curve is non-differentiable.

In derivative pricing, dynamical processes (It$\hat{o}$ processes) are reviewed; and geometric Brownian motion is included, together with its properties and applications. Conventional techniques used in estimation of the drift and volatility parameters such as historical techniques are reviewed and discussed. We also use the Hough Transform, an artificial intelligence method, to detect market patterns and estimate the drift and volatility parameters simultaneously. We look at different ways of calculating derivative prices. For option pricing, we use different methods but apply Bellalahs models in calculation of the Coffee Futures prices because they incorporate an incomplete information parameter.

Contents

List of Figures

List of Tables

Acknowledgement

I give sincere gratitude to God, for giving me life, health, brains and strength to undertake this study. I would like to also thank Strathmore University and Institute of Mathematical Sciences (IMS) for making it possible for me to pursue my academic dream. I would like to make a very special mention of Prof. John Odhiambo and Prof. Vitalis Onyango-Otieno. You were my fathers and mentors, my cheerleaders and motivators. Without your encouragement and faith in me, I would not have made it. Thank you.

I extend gratitude to my best friend and husband, Cosmas Kamuyu. You are the greatest husband and friend anyone can ever have. I thank you so much for the unending patience and understanding you accorded me as I studied during wee hours of the night. I would also like to thank you sincerely for the financial support you accorded me as I went to different countries to present papers related to this thesis, with particular mention of my journey to Namibia in November 2015 and the printing, binding and thesis compilation costs. May God bless you.

I cannot find enough words to express my gratitude towards those who assisted me with the technical know-how with regard to this project. Prof. Silas Onyango and Prof. Michael Ingleby; I would like to take this space and time to say thank you for believing in me, encouraging me, critiquing me and guiding me through many aspects of this thesis.

Finally, I would like to make a special mention of my mother, Teresia Muthoni, who encouraged me all through this study, making me believe that I can make it. Thank you mama.

Dedication

I dedicate this work to my husband and best friend, Cosmas Kamuyu, my fathers Prof. John Odhiambo and Prof. Vitalis Onyango-Otieno, and my cheerleaders Prof. Silas Onyango, Prof. Omolo Ongati and Prof. Michael Ingleby.

I offer special dedication in terms of prayers for Prof. Silas N. Onyangos health. May God grant you good health and speedy recovery. This study would not have been possible without your encouragement and understanding.

Chapter 1

STUDY DEFINITION

1.1 INTRODUCTION

The main intention of this study is to develop a zero coupon yield curve (ZCYC) for Nairobi Securities Exchange (NSE) and an associated model for pricing the coffee futures that are to be introduced in the Kenyan market in 2017-2018, as requested by Capital Market Authority of Kenya. This modeling task requires:

- the existence of zero coupon bonds in the market;
- continuous flow of data from the market;
- and historical prices of coffee futures.

Unfortunately, the required data and information is neither adequate nor available in the market. At the NSE, only coupon bonds are traded; zero coupon bonds are not traded and therefore we need to strip the coupon bonds to create hypothetical bonds associated with zero bond or risk-free rates. Secondly, the bonds available for trade have only specific, European-type, maturities: 1, 2, 3, 4, 5, 7, 10, 15, 20, 25 and 30 years, and there is no information on several tenures. Therefore, it is necessary for us to interpolate the data available so that we can estimate the rates for the missing maturities and construct a yield curve. There are many interpolation methods in the literature so careful selection is a significant part of this study.

There are two broadly different ways of constructing a yield curve: spline interpolation and parametric modeling. Of the former, we analyze different interpolation methods, their strengths and weaknesses, and finally develop a method that improves on the seeming best interpolation method: monotone convex interpolation on r(t)t, which is currently applied at the Johannesburg Securities Exchange when tenures are missing from historical data. As for the parametric models, we test the models which many studies have concluded to be the best: the (Nelson, C. R., & Siegel, A. F. , 1987) model, the (Svensson, L. E. , 1994) model and the (Rezende, R. B. , 2011) model, to decide which seems best suited for Kenyan data. Finally, we compare the results from the spline interpolation model with the best parametric model.

Not only does the NSE not trade in coffee futures; coffee is actually not traded there as a commodity at all. The real coffee beans are simply sold in weekly auctions and then exported to other countries from which they are traded in the commodity exchanges or consumed. The only available NSE data is the coffee beans prices from the auctions, we therefore have to look for alternative sources of data for the coffee commodity prices and futures.

To found Nairobi auction prices on a traded commodity framework, we relate them to data from guide countries that grow coffee and at the same time trade in coffee futures. The countries we consider as guides of this nature are: Ethiopia, Cote dIvoire (Ivory Coast) and Brazil, and we relate their economies to that of Kenya by fairly crude interest-rate comparisons. The basis of comparison that we use is correlation analysis and Frechet distance: In correlation analysis, the country which correlates most strongly with the Kenyan economy is taken as the best guide to starting an eventual commodity exchange price for Kenya. In problem formulation section 1.2 of the thesis, we consider instantaneous correlation of bank rates and also some correlations with a time lag because the economic conditions driving auction bidding in Nairobi may well be those prevailing in the guide market at an earlier date. The results, over a range of plausible time-lags, were consistent in showing that data from Ivory Coast provides a better guide than those from the other two economies. Frechet distance is a measure of similarity between two curves. It is defined as the minimum cord-length sufficient to join a point traveling forward along one curve and one traveling forward along the other curve, although the rate of travel for either point may not necessarily be uniform. The shorter the distance, the closely related the curves are. Frechet distance strongly supports Ivorys interest rates being closely related to Kenyas interest rates than the other two countries, Brazil and Ethiopia.

The coffee futures data to be used are available both at Trading Economics data platform (http://www.tradingeconomics.com/ivory-coast/indicators as on 8th August 2016) and Bar Chart Futures Market Overview data platform (http://www.barchart.com/futures/marketoverview as on 8th August 2016). These data are updated daily and will be available to the NSE.

To model option pricing, we begin with the simplest succession dynamical models of asset prices, the Geometric Brownian Motion (GBM) applied by Black, Scholes and Merton to derive the option pricing formula, the Black-Scholes formula, (see (Hull, J. C. , 2006)).

Generally, the option pricing formula assumes that the asset price evolves on a continuous time axis as a continuous Markov chain, as shown by (Hull, J. C. , 2006) : one of the so-called It$\hat{o}$ processes, which is expressed as a stochastic differential equation:

$$dS(t) = \mu S(t)dt + \sigma S(t)dZ(t) \qquad S(0) = S_0 \tag{1.1.1}$$

where parameter μ is the instantaneous linear drift rate or risk-free rate of return, and $\sigma > 0$ is the volatility of the underlying asset. Stochastic quantity $Z(t)$ is the standard Wiener process with $Z(t) \sim N(0, \sqrt{t})$. The volatility of the underlying asset, $\sigma > 0$ and the risk-free rate of return, μ, are assumed to remain constant during the life of the option. The advantage of this method is that the asset price S(t) at time t, is the value of an exponential function and hence is non-negative at all times.

From equation (1.1.1), we have:

$$dln(S(t)) = (\mu - \sigma^2/2)dt + \sigma dz(t) \tag{1.1.2}$$

From equation (1.1.2), we see that the variable $ln(S(t))$ follows a generalized Wiener process. The change in $ln(S(t))$ between time 0 and time T is normally distributed so that:

$$ln(S(t)) - ln(S(0)) \sim \Phi[(\mu - \sigma^2/2)T, \sigma\sqrt{T}]$$

where S(T) is the stock price at a future T; S(0) is the stock price at time 0 and $\Phi(m, s)$ denotes a normal distribution with mean m and standard deviation s. From this it follows that:

$$ln\left(\frac{S(T)}{S(0)}\right) \sim \Phi[(\mu - \sigma^2/2)T, \sigma\sqrt{T}] \tag{1.1.3}$$

and

$$ln(S(T)) \sim [ln(S(0)) + (\mu - \sigma^2/2)T, \sigma\sqrt{T}] \tag{1.1.4}$$

Equation (1.1.4) shows that $ln(S(T))$ is normally distributed so that S(T) has a lognormal distribution.

1.2 PROBLEM FORMULATION

In a simple market without derivatives, the in-house construction of a yield curve might not have a far reaching impact on the market as a whole; but the introduction of derivatives at the Nairobi Securities Exchange makes urgent the need for a well-researched and formulated yield curve. Briefly, yield curve construction using the spline based approach heavily relies on a good interpolation method to facilitate accurate bootstrapping. According to Investopedia, a website devoted to investing information and education based in New York, Bootstrapping is defined as the procedure used to calculate the zero coupon yield curve from market figures. Because T-bills offered by the Kenyan government are not available for every time period, the bootstrapping method is used to fill in the missing figures in order to derive the yield curve. The bootstrap method uses interpolation to determine the yields for treasury zero-coupon with various maturities (http://www.investopedia.com/terms/b/bootstrapping.aspo on 9th August 2016)

1.2.1 Selection of Interpolation Method

We begin by looking into traditional interpolation methods and their weaknesses. We then look at the spline-based methods, moving from the earliest models proposed by McCulloch, Vasicek, and the others, until we get to *monotone preserving interpolation on r(t)t* method proposed by Du Preez in 2013. Spline-based yield curve models typically involve minimizing the following function:

$$\begin{matrix} min \\ h(t) \end{matrix} = \sum_{i=1}^{N} (P_i - \hat{P_i})^2 \tag{1.2.1}$$

where N is the number of securities used as inputs to the model, P_i are the observed security prices, and h(t) is the chosen method of interpolation (the spline function) used to compute the fitted security prices $\hat{P_i}$.

We look at different interpolation methods h(t) used in the spline-based models. The interpolation method currently used at Johannesburg Securities Exchange, the most developed financial market in Africa, is known as *monotone preserving interpolation on r(t)t*, suggested by Du Preez. Unfortunately, this method produces curves that are non-differentiable at the knot points. This makes it hard for one to calculate forward rates (which are important in pricing derivatives) at the knot points, given spot rates.

1.2.2 Estimation of Forward Rates

Forward rates are calculated from the spot rates as follows:

$$f(t) = \frac{d}{dt} r(t)t \tag{1.2.2}$$

There are two ways of dealing with the non-differentiability at the knot points: first, we can use operator form of numerical differentiation to estimate the derivative, or we can develop a method of removing the non-differentiability at the knot points.

Operator form of numerical differentiation is based on functions f(x) sampled at discrete points $f(x+h), f(x+2h), f(x+3h), \ldots, f(x+nh)$.
Let us define the following:

$$Ef(x) = f(x+h) \quad \text{as the shift operator}$$

$$\Delta f(x) = f(x+h) - f(x) \quad \text{as the forward difference operator} \tag{1.2.3}$$

$$\nabla f(x) = f(x) - f(x-h) \quad \text{as the backward difference operator} \tag{1.2.4}$$

$$\delta f(x) = f(x+h/2) - f(x-h/2) \quad \text{as the central difference operator} \tag{1.2.5}$$

$$\mu f(x) = 1/2[f(x+h/2) + f(x-h/2)] \quad \text{as the average operator} \tag{1.2.6}$$

$$Df(x) = f'(x) \quad \text{as the differential operator} \tag{1.2.7}$$

where h is the difference interval. To link the difference operator with the differential operator, we consider Taylors series:

$$f(x+h) = f(x) + hf'(x) + h^2/2!f''(x) + \ldots + h^n/h!f^((n))(x) + \ldots \tag{1.2.8}$$

In operator notation, we can write:

$$Ef(x) = \left[1 + hD + \frac{1}{2!}(hD)^2 + \ldots\right] f(x) \tag{1.2.9}$$

The series in brackets is the expression for the exponential and hence we have (formally):

$$Ef(x) = e^{hD} f(x)$$

$$E = e^{hD} \tag{1.2.10}$$

Substituting these expressions into the central difference approximation for the derivative, which can also be expressed as $f'(x) \approx \frac{f(x+h)-f(x-h)}{2h}$ we obtain:

$$\frac{f(x+h) - f(x-h)}{2h} = \frac{e^{hD} f(x) + e^{-hD} f(x)}{2h}$$

$$= \frac{e^{hD} + e^{-hD}}{2h} f(x) \tag{1.2.11}$$

Using Taylor series expansion for e^{hD} and e^{-hD} for small values of h, we obtain:

$$e^{hD} = 1 + hD + \frac{1}{2}h^2D^2 + \frac{1}{6}h^3D^3 + \ldots + \frac{1}{n!}h^nD^n + \ldots$$

$$e^{-hD} = 1 - hD + \frac{1}{2}h^2D^2 - \frac{1}{6}h^3D^3 + \ldots + \frac{(-1)^n}{n!}h^nD^n + \ldots \tag{1.2.12}$$

Substituting the above equations into the following:

$$\frac{f(x+h) - f(x-h)}{2h} = \frac{e^{hD} + e^{-hD}}{2h} f(x)$$

We have in central difference form:

$$\frac{f(x+h) - f(x-h)}{2h} = \left[D + \frac{1}{6}h^2D^3 + \ldots\right] f(x) \tag{1.2.13}$$

Reorganizing the equation above, we get:

$$f'(x) \left[D + \frac{1}{6}h^2D^3 + \ldots + \frac{h^nD^{n+1}}{(n+1)!}\right] f(x) \tag{1.2.14}$$

thereby estimating the derivative at a point for tabular data.

The second way we deal with the non-differentiability at the knot points is by removing the monotonicity constraint introduced by Hyman, on the FritschButland algorithm in the monotone convex interpolation on r(t)t method.

We could also use parametric models to calculate the zero coupon yield rates; but since many researchers have established three best models for this, in this thesis we investigate which, among the parametric models is the best, and compare it to the performance of the differentiation methods discussed above. This comparison is done to ensure that we settle on the best method for constructing ZCYCs for the NSE. We then apply the generated risk free rates to pricing of the derivatives.

The accuracy of the pricing models for the Futures will depend on the country that we chose as a source of futures prices data.

1.2.3 Choosing the Lead Country

Here, we look at different methods of analysis taking into account the non-linear nature of the interest rates data. The methods used include: Correlation analysis with time lags and Frechet distance.

1.2.3.1 Correlation analysis with time lags

Correlation is a statistical measure that indicates the extent to which two or more variables fluctuate together or are related to each other. A **correlation coefficient** is a statistical measure of the degree to which changes to the value of one variable predict changes to the value of another. When the fluctuation of one variable reliable predicts a similar fluctuation in a similar variable, theres often a tendency to think that it means that the change in one causes the change in the other. However, correlation does not imply **causation**.

For simple correlation, let there be n pairs of observations on two variables x and y, then the usual correlation coefficient (Pearson's coefficient of correlation) is:

$$r = \frac{\sum_i (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_i (x_i - \bar{x})^2 (y_i - \bar{y})^2}} \tag{1.2.15}$$

Similar idea can be used in time series to see whether successive observations are correlated or not. Given N observations $x_1, x_2, \ldots, x_n$ on a discrete time series, we can form (n-1) pairs of observations such as $(x_1, x_2), (x_2, x_3), \ldots, (x_{n-1}, x_n)$. Here, in each pair, the first observation is as one variable (x_t) and the second observation is as second variable (x_{t+1}). So the correlation coefficient between x_t and x_{t+1} is:

$$r_1 = \frac{\sum_{t=1}^{n-1} (x_t - \bar{x}_{(1)})(x_{t+1} - \bar{x}_{(2)})}{\sqrt{\left[\sum_{t=1}^{n-1} (x_t - \bar{x}_{(1)})^2\right] \left[\sum_{t=1}^{n-1} (y_t - \bar{y}_{(1)})^2\right]}} \tag{1.2.16}$$

where

$\bar{x}_{(1)} = \sum_{t=1}^{n-1} \frac{x_t}{n-1}$ is the mean of the first $n-1$ observations.

$\bar{x}_{(2)} = \sum_{t=2}^{n-1} \frac{x_t}{n-1}$ is the mean of the last $n-1$ observations.

Note that: The assumption is that the observations in autocorrelation are equally spaced (equispaced). For large n, r_1 is approximately:

$$r_1 = \frac{\frac{\sum_{t=1}^{n-1} (x_t - \bar{x})(x_{t+1} - \bar{x})}{n-1}}{\frac{\sum_{t=1}^{n} (x_t - \bar{x})^2}{n}} \tag{1.2.17}$$

or

$$r_1 = \frac{\sum_{t=1}^{n-1} (x_t - \bar{x})(x_{t+1} - \bar{x})}{\sum_{t=1}^{n} (x_t - \bar{x})^2} \tag{1.2.18}$$

For k distance apart, i.e., for k lags, then:

$$r_1 = \frac{\sum_{t=1}^{n-k} (x_t - \bar{x})(x_{t+1} - \bar{x})}{\sum_{t=1}^{n} (x_t - \bar{x})^2} \tag{1.2.19}$$

An r_k value of $\frac{\pm 2}{\sqrt{n}}$ denotes a significant difference from zero and signifies an autocorrelation.

The following results of the correlation with time lags (up to lag 7) between the interest rates of Brazil, Ethiopia and Ivory Coast with Kenyas interest rates are shown in the table below. (See appendix B for the R codes and output).

Table 1.1: Correlation of the Interest Rates Data with Time Lags (up to Lag 7)

KENYA	LAGS	AUTO CORRELATION		
		BRAZIL	ETHIOPIA	IVORY
	1	0.883	0.824	0.966
	2	0.777	0.649	0.932
	3	0.658	0.473	0.898
	4	0.545	0.297	0.863
	5	0.464	0.285	0.829
	6	0.388	0.274	0.795
	7	0.303	0.262	0.752

From the analysis, Ivory Coast has the highest correlation with Kenya and this is consistent up to lag 7.

After running these tests at various lag times, Ivory Coast had the highest correlation of 0.9527 compared to Ethiopias 0.9328 and Brazils 0.8703. This led to this study using Ivory Coast Coffee futures historical prices for development of Kenyas Coffee futures pricing model.

1.2.3.2 Frechet Distance

Frechet distance is a measure of the similarity between curves that takes into account the location and ordering of the points along the curves. Let S be a metric space. A curve A in S is a continuous map from the unit interval into S, i.e. $A : [0,1] \to S$. A reparameterization α of $[0,1]$ is a continuous, non-decreasing, surjection $\alpha : [0,1] \to [0,1]$. Let A and B be two given curves in S. Then, the Frechet distance between A and B is defined as the infimum over all reparameterizations α and β of $[0,1]$ of the maximum over all $t \in [0,1]$ of the distance in S between $A(\alpha(t))$ and $B(\beta(t))$. In mathematical notation, the Frechet distance $F(A,B)$ is defined by:

$$F(A,B) = \underset{(\alpha,\beta)}{inf} \quad \underset{(t \in [0,1])}{max} \quad \{d\left(A(\alpha(t)), B(\beta(t))\right)\} \tag{1.2.20}$$

where d is the distance function of S.

The results (the Frechet distance for the interest rates data), were generated using R programming as shown in the table below. The R codes and output can be found in the appendices section.

Table 1.2: Frechet Distance for the Interest Rates Data

KENYA	FRECHET DISTANCE		
	BRAZIL	ETHIOPIA	IVORY COAST
	13.00839	4.113821	1.40443

Ivory Coast is again more related to Kenya, than to the other two lead countries, since it has the shortest distance.

The model we use for pricing is a generalization of Black Scholes Model. The novel item in regard to this study is that the parameters in the model are estimated using an artificial intelligence method, the Hough Transform. However, this thesis still outlines the other parameter estimation methods.

1.3 RESEARCH GOALS AND OBJECTIVES

The broad aims of this project can be translated into the following objectives:

1. Construction of a Zero Coupon Yield Curve for the Nairobi Securities Exchange

2. Development of a model to be used to price Kenyan Coffee Futures

1.4 BASIC CONCEPTS

1.4.1 Yield Curve Construction

A yield curve is a graphical representation as a continuous function of the relationship between present return (yield) of some type of financial instrument and the time remaining to its maturity. We can construct a yield curve using either discrete data sampling returns or from a continuous Markovian model. For discrete data points, interpolation between the sampled points is the most suitable method of constructing the yield curve. When using interpolation, the sampled returns are usually known as 'knot points'. If a Markovian model is used, the mathematical form of the yield curve is usually known as 'the model function' and these usually come in families with parameters to be fixed from data connected to the instrument being modeled.

For yield curve construction, this thesis uses both spline based methods and parametric models. Spline-based methods typically involve minimizing the following type of function

$$\begin{matrix} min \\ h(t) \end{matrix} = \sum_{i=1}^{N} (P_i - \hat{P_i})^2 \tag{1.4.1}$$

where N is the number of knot points, P_i is an observed security price at the i^{th} knot, $h(t)$ is the chosen method of interpolation (the spline function) and $\hat{P_i}$ is its value at the i^{th} knot time.

For parametric models, we use the Nelson-Siegel (NS) class of models. These models were first developed by Charles Nelson and Andrew Siegel from the University of Washington in 1987, improved by Svensson in (1992-1994) who added a second hump and by Rezende and Ferreira in 2011 who introduced the third hump. Their modeling is based on various possible shapes of yield curve: such as flat, hump, and S- shapes, (Nelson, C. R., & Siegel, A. F. , 1987). The model is presented as:

$$f(m) = \beta_0 + \beta_1 exp\left(\frac{-m}{\tau_1}\right) + \beta_2\left(\frac{m}{\tau_1}\right) exp\left(\frac{-m}{\tau_1}\right) + \beta_3\left(\frac{m}{\tau_2}\right) exp\left(\frac{-m}{\tau_2}\right) + \beta_4\left(\frac{m}{\tau_3}\right) exp\left(\frac{-m}{\tau_3}\right) \tag{1.4.2}$$

where f_m is the forward rate of government bond in i where $i = 1, \ldots, n$; n is number of bonds, m is time to maturity. $\tau = \tau_1, \tau_2, \tau_3$ reflects the scale parameter that measures the rate at which the short term and medium term components decay to zero. For example, small value of τ result in rapid decay in the predictor variables and therefore they will be suitable for curvature at low maturities. Corresponding, large volumes of produce slow decay in the predictor variables and will be suitable for curvature over longer maturities, (Christofi, A. C. , 1998).

β is a linear parameter vector; i.e. $\beta = \beta_0, \beta_1, \beta_2, \beta_3$; in which β_0 is a constant value of forward rate function, β_1 determines the initial value of the curve (short term) in various terms of abbreviations, the curve will be negatively skewed if parameter is positive and vice versa. β_2 determines magnitude and direction of the first extremum of the curve at time τ_1 (if β_2 is positive then the extremum is a hump , if β_2 is negative then U shape will occur at τ_1,) and β_3 determines magnitude and direction of the second local extremum at time, hump or U-shape according to sign. In (Nelson, C. R., & Siegel, A. F. , 1987), $\beta_3 = \beta_4 = 0$ and $\tau_2 = \tau_3 = 0$. In (Svensson, L. E. , 1994), $\beta_4 = 0$ and $\tau_3 = 0$.

1.4.2 Financial Derivatives Pricing

Financial derivatives are contracts used to buy or sell an underlying asset at a future time, at a price, quantity and maturity defined today. The more common examples of derivatives include forwards, futures, call and put options, caps, floors, swaps, dollars and many others [see (Hull, J. C. , 2006) chapters 1 to 5]. The underlying asset classes include interest rates, equity, foreign exchanges (FX), credit, energy and others. Derivatives are traded in organized exchanges as well as over the counter (OTC). The payoff function could be either continuous or discrete. Derivatives are used to hedge the risk of owning underlying assets that are subject to unexpected price fluctuations, (Hull, J. C. , 2006).

In this study, we are going to apply continuous time models on stock prices (coffee futures) from the Ivory Coast, our selected external guide market. Stock prices are stochastic in nature. Any variable whose value changes over time in an uncertain way is said to follow a stochastic process. Stochastic processes are classified as discrete time or continuous time. A discrete-time stochastic process is one where the value of the variable can change only at certain fixed points in time, whereas continuous-time stochastic process is one where changes can take place at any time. Stochastic processes can also be classified as continuous-variable or discrete-variable. In a continuous-variable process, the underlying variable can take any value, whereas in a discrete-variable process only certain discrete values are possible.

In practice, we do not observe stock prices following continuous-variable, continuous-time processes. Stock prices are restricted to discrete values and changes can only be observed when the exchange is open and a bargain is struck. Nevertheless, the continuous-variable continuous-time processes are the ones used most for pricing and modeling purpose. We build the modeling step by step until we get to the pricing model, starting with the Markov processes.

A Markov process is a particular type of stochastic process where only the recent history is relevant for predicting the future. A Markov process is memoryless. It does not matter how it got to the current state or for how long it has been staying there. The more remote history of the variable and the way that the present has emerged from this past are irrelevant. Stock prices are usually assumed to follow Markov process, where predictions for the future are uncertain and must be expressed in terms of probability distributions. The Markov property implies that the probability distribution of the price at any particular future time is not dependent on the particular path followed by the price in the past. The Markov property of stock prices is consistent with the weak form of market efficiency, which states that the present price of a stock impounds all the information contained in a record of stock prices. A Markov chain of higher order is a Markov model with memory, that is, a Markov chain that depends on, not only the current state, but also m-1 states before, where m is the order and is finite. For example, a second order Markov process depends on its current state and also the one just visited state, a third Markov process depends on its current state and the two previous states.

$$Pr(X_n = x_n | X_{n-1} = x_{n-1}, X_{n-2} = x_{n-2}, \ldots, X_{n-m} = x_{n-m})$$

If the weak form of market efficiency were not true, technical analysts could make above-average returns by interpreting charts of the past history of stock prices. It is competition in the market that tends to ensure that weak-form market efficiency holds. There are many investors watching the stock market closely. Trying to make profits from it leads to a situation where a stock price, at any given time, reflects the information in past prices. Suppose that it was discovered that a particular pattern in stock prices always gave a 65% chance of subsequent steep price rises. Investors would attempt to buy a stock as soon as the pattern was observed, and demand for the stock would immediately rise. This would lead to an immediate rise in its price and the observed effect would be eliminated, as would any profitable trading opportunities.

An instance where we see market inefficiency is at New York Stock Exchange, Wall Street, during the liquidity crash in 2007, as explained partly by Michael Lewis in the book The Big Short: Inside the doomsday machine (Gorton, G., Lewis, M., & Zuckerman, G. , 2011). In his book, Lewis speaks of a group of people, who we are going to refer to as the protagonists and follows their financial decisions. The protagonists bet against subprime collateralized debt obligations (CDOs). A CDO is formed by pooling a large number of mortgages and tranching the pool that is, imposing a seniority structure on it. In an ordinary pool of securities, a mutual fund for example, there is no seniority structure and everybody gets an equal share of the revenues. But in CDOs, the super-senior tranche might get 80% of the revenues before anyone else gets anything, the senior tranche gets the next 10%, the mezzanine 5% and the last tranche, the equity tranche, gets whatever is left over, say 5%, if all revenues (mortgage payments) come through. The super-senior tranches and often the senior tranches too were rated AAA by S& P and Moodys. Sometimes, the lower-rated mezzanine tranches were re-pooled and re-tranched into new CDOs and their super-senior tranches were rated AAA too. The protagonists in The Big Short wanted to short the wobbliest of these CDOs.

Initially, there was no way to do it until the Credit Default Swaps (CDSs) were made. CDSs on CDOs were essentially insurance on the CDOs. For a very small premium paid quarterly, one could buy insurance that would make their investment whole if their CDO tranche default, in whole or in part. And one did not even need to own the CDO to buy the insurance. The protagonists bet against subprime CDOs chiefly by purchasing the credit default swaps that insure them. The CDSs values would go up as the CDOs values went down. What actually used to happen is that the CDO seller, who has promised to make good up to the face value of the CDO, can now pool the premiums received on the sell-side of the bunch of the CDSs, tranche them and turn around and sell this to someone else. To the buyer, its virtually indistinguishable from having bought a mortgage CDO. One makes an investment, and gets periodic payments in return. These tranched pools of CDSs that simulated CDOs were called synthetic CDOs. And the more CDSs the protagonists bought, the more the dishonest sellers could create synthetic CDOs. Plenty of investors were convinced they were onto a sure thing and this overconfidence caused market inefficiencies since CDOs and CDSs lack the Markov property.

Next we would like to discuss the Wiener process, denoted by $Z = \{Z(t) : t \geq 0\}$, [Z usually denotes the standard Wiener process with $\sigma = 1$, can be specified by a stochastic differential equation] and is a particular type of Markov stochastic process index by continuous time and taking real values, generated by independent increments. The cumulative output of incremental $Z(t) - Z(s)$ is normally distributed with mean 0 and variance $(t - s)$, for some constant σ and any $0 \leq s < t$. It is sometimes referred to as Brownian motion and is a cornerstone of modern theory of random processes and its application. Wiener process was developed from a concept widely known as the Random Walk.

The term random walk was first used by Karl Pearson in 1905. He proposed a simple model for It$\hat{o}$ infestation in a forest: at each time step, a single It$\hat{o}$ moves a fixed length at a randomly chosen angle. Pearson wanted to know the It$\hat{o}$s distribution after many steps. It is also believed that the theory of random walks was developed a few years before (Bachelier, L. , 1900) in the PhD thesis of a young economist: Louis Bachelier. He proposed the random walk as the fundamental model for financial time series. Bachelier was also the first to draw the connection between discrete random walks and the continuous diffusion equation. In the same year of the paper of (Pearson, K. , 1905), Albert Einstein published his paper on Brownian motion which he modeled as a random walk, driven by collisions with gas molecules. Einstein did not seem to be aware or the related work of Pearson and Bachelier. In 1906, Smoluchowski also published very similar ideas.

A Wiener process is characterized by the following properties:

1. $W_0 = 0$

2. W has independent increments: $W_{t+1} - W_t$ is independent of $\sigma(W_s : s \leq t)$ for $u \geq 0$

3. W has Gaussian increments: $W_{t+1} - W_t$ is normally distributed with mean 0 and variance t: $W_{t+1} - W_t \sim N(0, t)$

4. W has continuous paths: with a probability of 1, W_t is continuous in t

The increments mean that if $0 \leq s_1 < t_1 \leq s_2 < t_2$ then $W_{t_1} - W_{s_1}$ and $W_{t_2} - W_{s_2}$ are independent random variables, and the similar condition holds for n increments. We show some properties of a Wiener process using the random walk example.

The walker starts at position $x = 0$ at the step $t = 0$; at each time-step the walker can go either forward or backward of one position with equal probabilities of 0.5. We need the probability $P(x, t)$, to find the walker at the position x at the time step t.

P(x,t) = no. of sequences that take to x in t steps × prob of any given sequence of t steps

Number of sequence that take to x in t steps= steps taken in the

positive direction × steps taken in the negative direction

$$P(x,t) = \frac{t!}{\left(\frac{t+x}{2}\right)!\left(\frac{t-x}{2}\right)!} \times \left(\frac{1}{2}\right)^t \sim \sqrt{\frac{2}{\pi t}} e^{-\frac{x^2}{2t}} \tag{1.4.3}$$

If we are looking at a stochastic process experiencing normal noise, finite variance and in one dimension, then we have:

$$x(t+1) = x(t) + \eta(t) \tag{1.4.4}$$

According to central limit theorem, the sum of independent identically distributed variables with finite variance will tend to be normally distributed. Therefore, the average distance travelled is calculated as:

$$x(t) \sum_{\tau=1}^{t} \eta(\tau) \tag{1.4.5}$$

$$(x)_t = (\eta)_t = 0 \tag{1.4.6}$$

$$P(x,t) \sim \frac{1}{2\pi t(\eta^2)} e^{-\frac{x^2}{2t(\eta^2)}} \tag{1.4.7}$$

$$\sqrt{((x^2)_t - (x)_t)} = \sqrt{((\eta^2)_t)} \propto \sqrt{t} \tag{1.4.8}$$

A generalized Wiener process for a variable x can be defined as follows:

$$dx = adt + bdz \tag{1.4.9}$$

where a and b are constants and $\Delta z = \varepsilon\sqrt{\Delta t}$ which is the change during a small period of time, Δt.

An Itô Process is a further generalization of the generalized Wiener process where the parameters a and b are functions of the value of the underlying variable, x and time t. It is often implied that a stock price follows a generalized Wiener process; that is, that is has a constant expected drift rate and a constant variance rate. However, this model fails to capture a key aspect of stock prices. This is the expected percentage return required by the investors from a stock is independent of the stocks price. This model also shows that when the variance rate is zero, the stock price grows at a continuously compounded rate of μ per unit of time. In practice, the stock does exhibit volatility.

The model that considers this volatility of stock price behavior is known as Geometric Brownian motion. The discrete-time version of the model is:

$$\frac{\Delta S}{S} = \mu\Delta t + \sigma\varepsilon\sqrt{\Delta t} \tag{1.4.10}$$

or

$$\Delta S = \mu St\sigma t + \sigma S\varepsilon\sqrt{\Delta t} \tag{1.4.11}$$

The variable ΔS is the change in the stock price, S in a small interval of time Δt; and ε is a random drawing from a standardized normal distribution (i.e., a normal distribution with a mean of zero and standard deviation of 1.0). The parameter μ is the expected rate of return per unit of time from the stock and the parameter σ is the volatility of the stock price. Both of these parameters are assumed constant. The $\frac{\Delta S}{S}$ is the return provided by the stock in a short time Δt. The $\mu\Delta t$ is the expected value of this return and the term $\sigma\varepsilon\sqrt{\Delta t}$ is the stochastic component of the return. The variance of the stochastic component is $\sigma^2\Delta t$.

It is from the Brownian motion that the Black-Scholes Pricing Model is derived. This option pricing formula was initially derived in 1973 by Fisher Black and Myron Scholes for asset options and later refined by Black in 1976 for options written on futures. The primary inputs in the formula are the underlying asset price, the strike price, time to expiration, risk-free rate of return, and the standard deviation of the underlying asset return. One of the core assumptions in the derivation of this model is the no-arbitrage principle. In brief, this principle states that in an efficient financial market, it should not be possible to make profits with zero investment and without bearing any market-risk.

One important task with using this formula is the estimation of the key parameters, and those are the drift rate and the volatility rate. In this study, we intend to use an artificial intelligence tool, the Hough transform, to estimate the two parameters.

The Hough Transform (HT) is a popular and robust method of extracting analytic curves from edge-enhanced noisy digital images. The principal concept of HT is to define a mapping between an image space and a parameter space of a class of curves. Each feature point or pixel in an image is mapped to the parameter space to vote for the parameters whose associated curves pass through the data points. The votes from all edge pixels along a curve are accumulated in a histogram and all the peak of the histogram corresponds to the parameters of the curve in the image. The curve path in image space therefore becomes a peak detection problem in the parameter space.

1.5 BACKGROUND INFORMATION

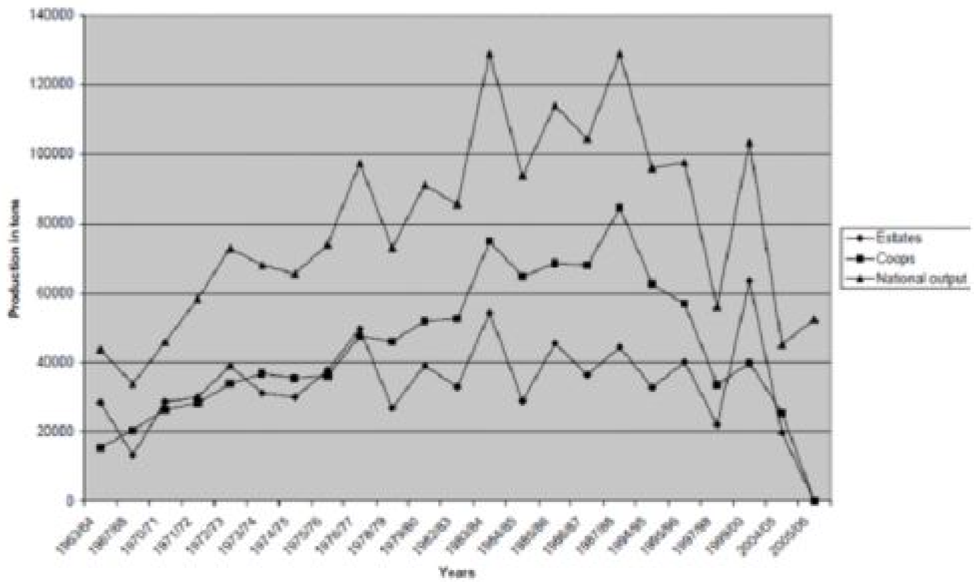
1.5.1 Coffee in Kenya

One of the main intentions of this study is to develop a model for pricing coffee Futures which are to be introduced in the Kenyan market in 2017-2018. The coffee sector in Kenya is an important economic activity in terms of income generation, employment creation, foreign exchange earnings and tax revenue: the coffee sector 'cash cow'. Over the years, the economic performance of coffee has had repercussions on all spheres of life in Kenya; affecting farm input suppliers, the transport sector, savings and investment intermediation, consumption of goods, and households ability to pay for education, health and other services. Even politics is affected by 'gravy trains' derived from the coffee market.

Kenya produces some of the best coffee in the world. Being the more flavourful Coffee Arabica rather than Coffee Canephora (Robusta), the fully washed mild belongs to the top quality group called Colombian milds. Kenya is able to produce the best coffee due and deep red volcanic soils in high altitude regions (1,500-2,000 meters above sea level) where well-distributed rainfall and moderate temperatures (averaging 20 deg centigrade Celsius), couple with characteristically high equatorial ultraviolet sunlight diffusing through thick clouds.

Coffee producers involves about 45 per cent of Kenyas total population (currently about estimated at 42 million [1]). Since some of these people are as much as 40 per cent income-dependent on coffee, their lives revolve around the pricing of coffee. Coffee production increased rapidly in ripples in the two decades after independence. As shown in figure 1.5.1 below; total production for both estates and cooperative sub-sectors rose from 43,778 tons in 1963-64 to 128,941 tons in 1983-84. Since then, however, the coffee industry has been on a downward trend except for a brief spell in 1999-2000. As a result, coffees contribution to incomes, employment creation and foreign exchange earnings has declined.

Figure 1.1: Coffee production trends, 1963-2006



Source: Task Force Report on Coffee Marketing, Ministry of Agriculture August 2003, p.158; Economic Survey, 2006, Government of Kenya; and the Coffee Quarterly, Kenya Coffee Traders Association, No. 2/2006, p.9.

Coffee production undergoes different steps which are normally depicted in a value chain. The value chain in coffee production in Kenya involves the following steps:

1. Farm-level **operations** - nursery operations to produce seedlings, planting, weeding, fertilizing, pruning, spraying, picking/harvesting of red cherry and transportation of cherries to the

[1] Economic survey, Central Bureau of Statistics, 2006, Nairobi, Kenya

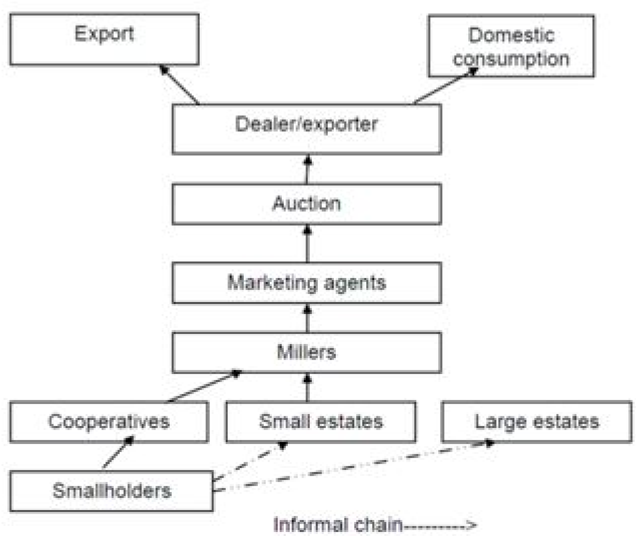
pulpier/coffee factory;

2. Factory primary **processing** - pulping, fermenting, washing and drying to produce parchment coffee, either at a cooperative facility or in a farm-based pulpery; curing for milling (removing parchment/peeling, cleaning and polishing the beans to produce green coffee beans);

3. **Milling** plant operations: hulling, cleaning/polishing, sorting, grading, bagging, e.g. by Kenya Planters Cooperative Union (KPCU) and Thika Coffee Mills;

4. **Auctioning** at the Nairobi Coffee Exchange (NCE) where dealers, roasters, marketers and exporters buy various grades of green coffee Auctioning at the Nairobi Coffee Exchange (NCE) where dealers, roasters, marketers and exporters buy various grades of green coffee. The coffee may be bought by either local consumers (example coffee houses like Nairobi Java House and Dormans Coffee), where they roast, grind, blend and package the coffee for local consumption. It can also be bought by exporters and international dealers who will market and sell it regionally and globally to consumers.

Figure 1.2: Coffee Value Chain



Source: Final Report on Assessment of the Value-adding Opportunities in the Kenyan Coffee Industry, European Commission, April 2004

If exported coffee is not processed and packaged for consumption, it is traded in the international markets; and it is here that the international prices of coffee are determined. There are two types of market: the physical market and the financial market (sometimes known as the futures market).

In the physical market traders buy coffee from exporters. Physical coffee moves between market participants and payments are made for the physical coffee. Often there will be a variety of profit-seeking intermediaries through whom the coffee passes: for example, secondary cooperatives, processors, re-exporters, etc. The intermediaries constitute a supply chain of physical volumes of coffee leading ultimately to the consumer who drinks the coffee.

Futures markets are very different from this physical market both in how they function and in their purpose. Their underlying asset is imported coffee held in a bonded warehouse like a precious metal. There is a holding cost in addition to the Nairobi auction price, transport, insurance, spoilage rate etc. For the purpose of this study, we do not examine such micro-economic matters, assuming the importer can factor them into the price levels at which he chooses to strike futures contracts - his macro-economic concern. Unlike the physical market, in the financial market contracts will only result in physical delivery of coffee at a completion date. These contracts are held for financial purposes and the contracts complete at (for a European option), or prior to (for an American option) the maturity date of the contract. Because such contracts are offset from the physical exchange at a bonded warehouse, they are often referred to as paper contracts. The vast majority of contracts traded on the exchange are traded as a means of providing sellers/importers and buyers and of coffee (in the physical market) with hedging opportunities to manage their exposure to future price risk.

Futures markets are accessed by participants from all over the world, which results in an extremely large number of transactions every day. By having so many market players trading coffee contracts in one location, the demand and supply for these contracts help buyers and sellers determine an aggregate price for coffee which is commonly known as the world price of coffee: in other words, **price discovery.** This price is used by producers, traders, exporters, and roasters around the world as the reference price for coffee on any given day. Of course there are different reference prices for different types and origins of coffee.

Comparing these features of the physical market to those of the financial markets shows that the main differences between the two markets are:

- **Location** the physical market exists in coffee producing countries, with buyers and sellers trading physical or green coffee. There will also be a physical coffee market in importing countries, where physical coffee is traded between importers and coffee roasters. The financial market on the other hand is a global (often electronic) exchange where futures and options (representing coffee for delivery in different months) are exchanged.

- **Activities** in the physical market, the primary activity is the buying and selling of green coffee between businesses that earn money from trading and moving coffee. In the financial markets, coffee contracts are traded with very little expectation of delivery of coffee.

- **Delivery** in physical coffee markets, delivery is usually by shipment from the port of exportation. All contracts on the futures are based on delivery of coffee stored in exchange-licensed warehouses in the US and Europe.

- **Export Terms** Local traders operating in the physical market will have contracts based on Free on Board (FOB) export terms whereas the financial market contracts are priced in store (also called 'ex dock' or delivered licensed warehouse) , meaning that the coffee is presumed to have been bought FOB, shipped from origin and discharged into a licensed or bonded warehouse.

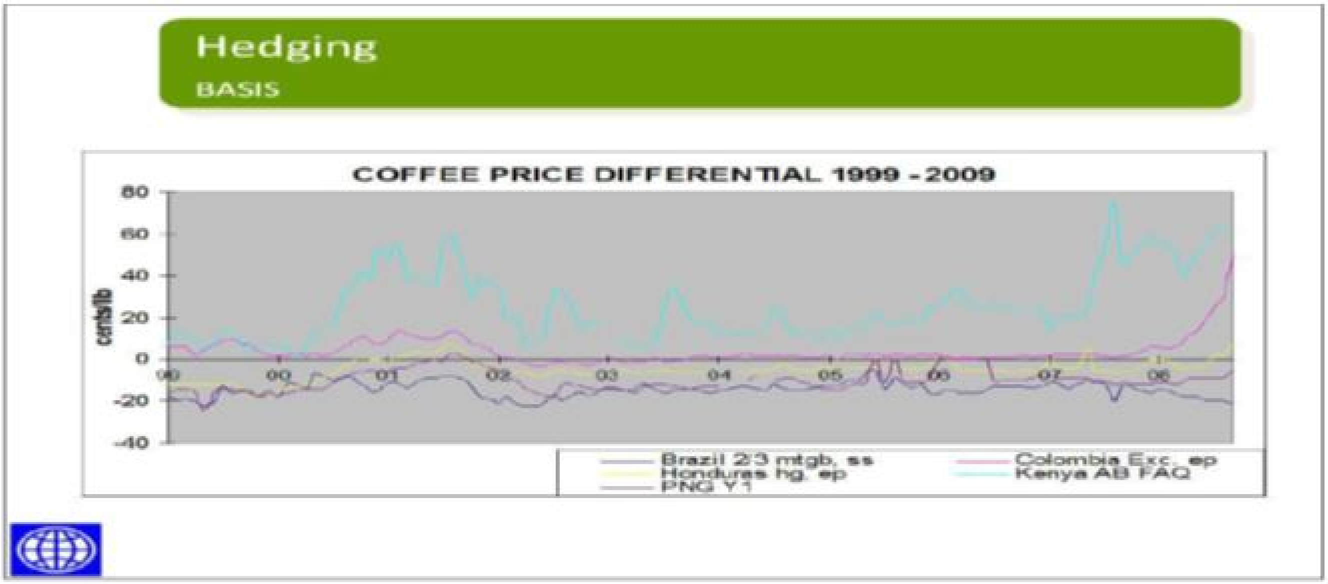
- **Units of Measurement** Local markets utilize their own units of measurement (for example Kilograms for East Africa, Quintiles for Central America, pounds in New York; but futures markets use pounds in the New York Arabica exchange and tons in the London Robusta exchange).

Each of these differences will affect the basis (the differential between FOB and in store prices), as they each involve different costs. Basis refers to the difference between the FOB price (the futures price) and the in store price. Many traders know this more commonly as the differential', and they speak of 'basis risk' or differential risk the two meaning the same thing. This basis is determined by several different factors such as quality differences between one country and another, the costs of transportation, interest and insurance. Basis can be either positive or negative. Positive basis occurs when the local market price for coffee expressed in FOB terms is greater than the in store international price. Basis is readily calculated from unit prices:

FOB unit price in store unit price = Basis

The following figure shows the coffee price differential for five coffee exporting countries. We note that Kenyas differential is very high compared to the other countries sampled, thus the motivation to introduce a coffee futures exchange in Kenya with lower in store price, which it hoped will lead to better prices for the farmers.

Figure 1.3: Coffee Price Differential on New York Stock Exchange



Accordingly, the government of Kenya announced in the June 2010 budget the introduction of a Coffee Futures Exchange. The regulatory framework for this was completed in 2014. The next steps include:

- Infrastructure expansion, for example construction of warehouses for coffee beans storage;
- Charging the Capital Market Authority with the development of pricing tools a ZCYC for the NSE, and a pricing model for coffee future contracts;
- Assistance to the shift from auctioning of coffee at the Nairobi Coffee Exchange to paper trading at the NSE.

This follows the historical precedent set during the evolution of the NSE from an auction-based stock exchange for East African enterprises established early in the nineteenth century and shifted to the continuous paper-trading NSE in the early 1990s. The lessons to be learned from this introduction of yield-curve based trading are outlined below.

1.5.2 Nairobi Securities Exchange and Market Yield Curve

The Nairobi Securities Exchange (NSE) started trading in shares while Kenya was still a British colony in the 1920s, according to International Finance Corporation and Central Bank of Kenya report, (Kenya, C. B. of. , 1986). Share trading was initially conducted in an informal market that would facilitate access to long-term capital by private enterprises and also allow commencement of floating of local registered Government loans. In 1954, the Nairobi Stock Exchange was constituted as a voluntary association of stockbrokers registered under the Societies Act, (Exchange-NSE, N. S. , 1997). The newly established stock exchange was charged with the responsibility of developing the Securities market and regulating trading activities.

The government adopted a new policy with the primary goal of transferring economic and social control to citizens, in 1963-1970. By 1968, the number of listed public sector securities was 66, of which 45% were for Government of Kenya, 23% Government of Tanzania and 11% Government of Uganda, while the rest were sold to limited companies registered in East Africa. During this period, the NSE operated as a regional market in East Africa where a number of the listed industrial shares and public sector securities included issues by the Governments of Tanzania and Uganda (the East African Community). Changing political regimes finally ultimately led to the delisting of companies domiciled in Uganda and Tanzania from the NSE.

The development path of securities in the NSE indicates an evolutionary process where changes in institutional infrastructure and the policy environment were witnessed as efforts were made to facilitate the growth of the securities market. The evolutionary process is characterized by a shift in trading system from a periodic auction system to a continuous trading system. Trading systems define the price discovery process or the transformation of latent demand of investors into realized transactions, according to (Madhavan, A. , 1992). Generally, the evolutionary process of trading systems indicated a shift from manual and decentralized settlement clearing systems to electronic and centralized settlement clearing [2]. Such a trading system enhances efficiency in the price discovery process, provides liquidity at low costs, and has no excess volatility, thus is more desirable for the development of the securities market; (Amihud, Y., Mendelson, H., & Lauterbach, B. , 1997) and (Bessembinder, H., & Kaufman, H. M. , 1997). High liquidity enhances long-term investment by reducing the required rate of return and lowering the cost of capital to the issuers of securities. An efficient price discovery process enhances the role of the market in aggregating and conveying information through price signals, therefore making prices more informative.

The development of the institutional and policy changes which have affected this market can be sum-

[2](Garman, M. B., & Klass, M. J. , 1980) observes an evolutionary pattern in adoption of trading system for the US stock market in response to growth in trading volume; this saw a shift from periodic to continuous trading system. (Amihud, Y., Mendelson, H., & Lauterbach, B. , 1997) note the tendency for emerging markets to shift from periodic to continuous trading in the revitalization process.

marized as an evolutionary process with three stages of development. The initiation stage was mainly characterized by lack of formal rules and regulations and was dominated by foreign investors. In the formalization stage, a self-regulatory system was adopted while attempts were made to increase the participation of local investors by the post independence Government. In the third stage, enhancement stage, various institutional and policy reforms were implemented to enhance the growth of the market.

Two main factors that shaped the development path were first, the political environment both at the national level and in the East African region where changes in the policy and in the composition of market participation had an impact; and second the macroeconomic environment where a demand for locally-mobilized long term capital to enhance economic development became significant.

Considering the developments in various aspects of the market, the following patterns were evident. First, the Securities brokerage industry expanded and diversified with the number of Securities brokers increasing from 6 (in 1950s) to 23 (as at 12th March 2016: source NSE website), and licensing of Securities dealers and investment banks also increased. However, Securities brokerage is yet to be fully negotiable while the role of Securities brokers remains predominately that of an agent. Second, the composition of market participants shows a shift from a market dominated by foreign investors in the initiation stage to increased participation of local investors in the formalization stage (especially in the post-independence period), re-entry of foreign investors, though at limited level and mass education on Securities market operations and assets, in the revitalization stage. Third, the NSE served as a regional market for the East African states in the initial stages but the Securities Exchange lost a significant proportion of its market scope due to the political changes in these states. Attempts are however underway to establish a regional market to facilitate expansion of the NSE market. Fourth, though the diversity of securities traded is still minimal, the reform period saw efforts made to attain market depth with introduction of new instruments such as the planned introduction of derivatives. Fifth, while a discriminative tax policy that penalized share investors heavily was adopted in the post-independence period, an incentive-based tax policy regime was adopted in the reform period to enhance the competitiveness of the financial assets and reduce the barriers to listing new issues. Sixth, in an effort to strengthen the regulatory framework, the regulatory system has witnessed a shift from the non-formal system to self-regulatory and statutory regulatory framework. Finally, the trading system shows a shift from the coffee-house forum to floor trading, and there were attempts to introduce the delivery versus payment system with the introduction of the Central Depository System and automation of the trading cycle. The main aim was to enhance efficiency in the price discovery process and liquidity of the market. These changes mimic development paths of other Securities markets in both the developed and emerging markets.

Healthy development paths in many markets around the globe has led to a number of Futures Exchanges springing up in Latin America, Asia and Africa during the last decade, based on the premise that there is a need to manage price volatility and provide price discovery. In June 2010 Budget speech, the government of Kenya, through a policy pronouncement by the then Deputy Prime Minister and Minister of Finance, announced that steps would be made towards developing institutional and legal frameworks to introduce a Commodities and Futures exchange in Kenya. This has created a need to design tools that will be used by the market to price the derivatives properly. Examples of tools needed are a yield curve and pricing models.

Currently, the NSE does not have a known existing yield curve as indicated by (Ngugi, R., & Agoti, J. , 2007). According to (Ngugi, R., & Agoti, J. , 2007), leading stockbrokers, a group of bank traders and institutional fund managers were working with Reuters Limited in 2006 towards developing an acceptable, credible market yield curve. This was driven by the requirement of International Accounting Standard (IAS) 39, which requires that there should be a standard yield curve to facilitate investors in fixed income instruments to value their portfolios at fair market values. However, the market participants identified limitations in prices reported on the NSE and decided not to use the yield curve created at the time to value their portfolios.

It was then agreed that the yield rates applied to the market were to be derived from a cross-section of key market players. These key market players were required to post their quotes on Reuters assigned pages, out of which aggregate yields are derived. The yield curve, which was primarily available through Reuters, which is a media channel dedicated to providing market data in the international financial market and disseminating that information locally.

In 2011, Cannon Asset Managers (CAM), a Kenyan based Asset Management company, created a yield curve for the NSE using the rates given by the Central Bank of Kenya (CBK), according to CAM (2011). They used logarithmic linear interpolation method to calculate the yield rates at the periods missing from the CBKs bonds. Unfortunately, logarithmic linear interpolation has a tendency of implying discontinuities in the forward rate curve, a weakness depicted by all variations of linear interpolation methods, as shown by (Hagan, P. S., & West, G. , 2006) We intend to construct a yield curve for the NSE by overcoming the limitations experienced by the group working on the yield curve in 2006. The main limitation of their method was using the bonds dirty prices, as quoted at the NSE. To improve on this, we will use raw bond data from CBK, which is the primary issuer of the bonds in Kenya. This way, we will construct a yield curve which is based on clean bond prices (without accruals of interest, as is the case in secondary market). In addition, we intend to use an improvement of *monotone preserving interpolation on r(t)t* method for the scenario where we might need an interpolation model; otherwise, we are going to use numerical approximation methods when needed to determine derivatives of spot rates where we need forward rates. We will use coupon paying bonds data from the Central Bank of Kenya.

1.5.3 Yield Curves

For Kenyas Futures Exchange to take off, among other things, there has to be a yield curve that is going to be accepted and used by the majority of NSEs market players. A yield curve is a graphical representation of relationship between return (yield) of same type of financial instruments and its day to maturity. In other words, in yield curves, all the differences in terms of types, credit risks and liquidity are removed from bonds and just the path of interest rates according to maturity rate is represented.

Yield curves can be grouped into two, in terms of coupons: coupon bearing yield curves and zero coupon yield curve. A coupon bearing yield curve is obtained from observable bonds market at various times to maturity, with the bonds having the same coupon rate. Most of the government securities having long maturity dates usually have coupons. When the bond itself and the coupons traded separately we have what is called STRIPS (Separate Trading of Registered Interest and Principal of Securities), in finance literature.

The zero-coupon yield curve is also known as the term structure of interest rates. It measures the relationships among the yields on default-free securities that differ only in the term to maturity. By offering a complete schedule of interest rates across time, the term structure embodies the markets anticipation of future events. In bond-valuation, the term structure of interest rates refers to the relationship between bond prices of different maturities in general. When interest rates of bonds are plotted against their maturities, this is called the yield curve. The term structure of interest rates and yield curves are used interchangeably in literature.

Yield curves can also be grouped into nominal yield curves, spot yields curves, and forward yield curves. Nominal yield curves take place in primary bond markets, and it is the graph of yields of bonds which are transacted at nominal prices. Spot yield curve is another definition of zero coupon yield curve. Forward yield curve is the curve representing the connection between the forward rate and its corresponding maturity where the forward rate is the interest rate implied by the zero coupon rates for periods of time in the future.

1.5.3.1 The Uses of Yield Curves

The relationship between yield and maturity has critical importance for policy makers, investors and economists. Yield curve can be used for a range of purposes. For example:

- Yield curves are used in forecasting interest rates for different products when their risk parameters are known.
- Yield curves are used mostly by investors to see the differences in yields of different maturities, and to detect if there is arbitrage opportunity.
- By yield curve information, investors can have opportunity of making immunization of their investment portfolios against financial risks if they have to make investment on some determined time of maturity.
- Private sector firms look at yields of different maturities and then choose their borrowing strategy according to information gotten from the yield curve.
- The differences in yields for long maturity and short maturities are an important indicator for central bank to use in monetary policy process as shown by (Akinci, O., Gurcihan, B., Gurkaynak, R., & Ozel, O. , 2006). These differences may show the tightness of the government monetary policy. The differences can be monitored to predict recession coming in next years.
- Yield curves are also used to calibrate no-arbitrage term structure models like the models of (Ho, T. S., & LEE, S.-B. , 1986); (Hull, J., & White, A. , 1990); which are used in pricing different financial products, (Place, J. , 2000).

1.5.3.2 Theories Explaining the Yield Curve

The characteristic shapes of yield curves have been related to attempts to seek rationality in investor behavior - to theorize about their choices of instrument. Three such are described here:

- Expectation theory concerned with expected annual returns on preferred bonds, which empirically are much the same for long and short term maturities [e.g. (Cox, J. C., Ingersoll Jr, J. E., & Ross, S. A. , 1985)].

- The market segmentation hypothesis concerned with the grouping of investors according to their maturity preferences, longer or shorter [e.g. (Munasib, A., & Haurin, D. , 2004)].

- Liquidity preference theory concerned with signs of preference for holding long-term bonds amongst some risk-averse investors (Pacific Investment Management, , 2004).

1.5.3.2.1 Expectation Theory

There are various versions of expectation theories. These theories place predominant emphasis on the expected values of future spot rates or holding-period returns. In its simplest form, the expectation hypothesis postulates that bonds are priced so that the implied forward rates are equal to the expected spot rates. Generally, this approach is characterized by the following propositions: (a) the return on holding a long term bond to maturity is equal to the expected return or repeated investment in a series of the short-term bonds, or (b) the expected rate of return over the next holding period in the same for bonds for all maturities (Cox, J. C., Ingersoll Jr, J. E., & Ross, S. A. , 1985). The key assumption behind this theory is that buyers of bonds do not prefer bonds of one maturity over another, so they will not hold any quantity of a bond if its expected return is less than that of another bond with different maturity. Bonds that have this characteristic are said to be perfect substitutes. Note that what makes long term bonds different from short term bonds are the inflation and interest rate risks. Therefore, this theory essentially assumes away inflation and interest rate risks, (Munasib, A., & Haurin, D. , 2004).

Another version of expectations theory holds that the slope of the yield curve reflects only investors expectations for future short-term interest rates much of the time, (Fisher, M. , 2001). Investors expect interest rates to rise in the future, which accounts for upward slope of the yield curve. If the expectations hypothesis were correct, the slope of term structure could be used to forecast the future path of interest rates. For example, if the yield curve were to slope upward at the short end, it would be because the interest rate is expected to rise. One problem with this version of the expectations hypothesis is that in fact, yield curves slope upward at the short end on average even though interest rates do not rise on the average. One way to explain divergence is to assume that investors are simply wrong on average [3].

[3]Another way to explain the divergence is to assume that investors give some weight to very large increases in the interest rate that have not yet been observed. This is sometimes called a peso-problem. But a good theory should not

The expectations hypothesis can easily be modified to account for this persistent upward slope in a way that does not require systematic errors on the part of investors. Since bond prices do fluctuate over time, there is uncertainty (even for default free bonds) regarding the return from holding a long-term bond over the next period. Moreover, the amount of uncertainty increases with maturity period of the bond. If there were a risk premium associated with uncertainty, then the yield curve could slope upward on average without implying that interest rates increase on average. If the risk premium were constant, the changes, in the slope of the yield curve would forecast changes in the future path of the interest rate. For example, if the slope of the yield curve were to increase, then it would have to be because the path of futures interest rates is expected to be higher. This increase in the slope would imply that future bond yield would be higher.

Another feature of the yield curve that the expectations has difficulty explaining is that the zero-coupon yield curves slopes <u>downward</u> on average at the long end, typically over the range of twenty to thirty years bond. In other words, the yield on a thirty-year zero-coupon bond is typically below the yield on a twenty-year bond. The expectations hypothesis would suggest that that this slope is due to either (1) a persistently incorrect belief that the interest rate will begin to fall about 20 years from or (2) a decrease in the risk premium for bonds with maturities beyond twenty years, even though the uncertainty of the holding-period return for thirty-year bonds. Neither of these reasons is sensible [4].

There is, however, a sensible explanation, for the persistent downward slope for the term structure at the long end. The explanation has to do with uncertainty regarding the future-path of short-term rates. This uncertainty underlies the risk of holding bonds (if there were no uncertainty regarding the future paths, there would be no risk of holding default-free bonds.) Increases in this uncertainty lead to 1) increases in risk <u>premia</u> that increase the slope of the yield curve at the short end and 2) decreases in the slope of the yield curve at the long end via the effect of convexity. Convexity (technically known as Jensen's inequality) arises from the non-linear relation between bond yields and bond prices. As a consequence, a symmetric increase in uncertainty about yield raises the average price of bonds, thereby lowering their current yields. This effect is trivial at the short end of the yield curve where it plays no significant role, because it becomes noticeable and even dominant at the long end. The overall shape of the yield curve involves the trade-off between the competing effects of risk

imply that investors are wrong on average

[4] There is another explanation (not related to the expectations hypothesis) that is sensible. The downward slope at the long end of the yield curve could, in principle, reflect a substantial demand for the longest-maturity (default free) zero coupon bond (for example, to insulate the value of insurance companies' long term liabilities from interest -rate risk).

premia (which cause longer term yields to be lower). Typically, the maximum yield occurs in the fifteen to twenty-five maturity range of the zero coupon yield [5].

1.5.3.2.2 The Market Segmentation hypothesis

Culbertson hypothesized that individuals have strong maturity preferences, and bonds of different maturities trade in separate markets. In its purest form, segmentation theory assumes that markets for different-maturity bonds are completely segmented. The interest rate for each bond with a different maturity is then determined by the supply of and demand for the bond with no effects from the expected returns on other bonds and other maturities. In other words, longer maturity bonds that have associated with the inflation and interest rate risks are completely different assets than the shorter bonds and are used by different types of investor. Thus, bonds of different maturities are not substitutes at all, so the expected returns from a bond of one maturity has no effect on the demand for a bond of another maturity. Because bonds of shorter holding periods have lower inflation and interest rate risks, segmented market theory predicts that yield on longer bonds will generally be higher, which explains why yield curve is usually upward sloping. However, since markets for different-maturities bonds are completely segmented, there is no reason why the short and long yield should move together. And, because of the same reason, the segment market theory also cannot explain why the short-term yield should be more volatile than longer-term yields, as pointed out by (Munasib, A., & Haurin, D. , 2004).

1.5.3.2.3 The Liquidity Preference Theory

This theory is an offshoot of the pure expectations theory and it asserts that long term interest rates not only reflect investors assumptions about future interest rates but also include a premium for holding long term bonds, called the term premium or the liquidity premium. This premium compensates investors for the added risk of having their money tied up for a longer period, including the greater price uncertainty. Because of the term premium, long term bond yields tend to be higher than short-term yields and this gives yield curves an upward slope, (Pacific Investment Management, , 2004).

[5] It should be stressed that the yield curve typically reported in the newspaper is not the zero-coupon yield curve and may display a somewhat different shape owing to a variety of factors (Fisher, M. , 2001).

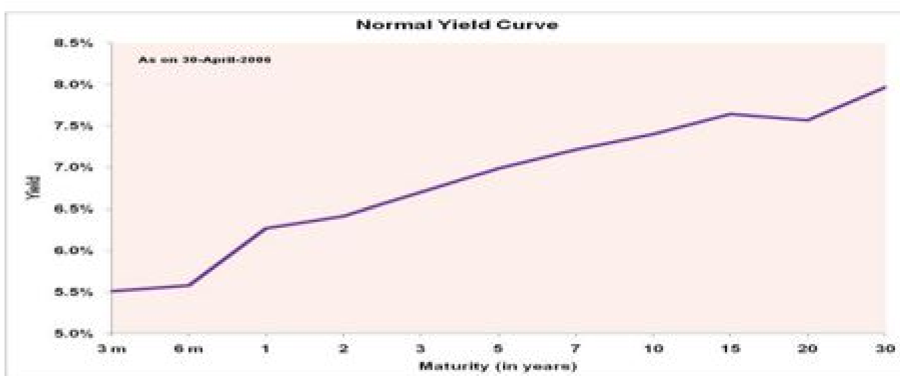
1.5.3.3 Shapes of Yield Curves

As we have seen in the above section, the theories of yield curve try to explain the shapes of the yield curve. Yield curves reflect the market expectations and may take one of the three main patterns: normal, flat or inverted curves. Each shape provides different information for market.

1.5.3.3.1 The Normal Yield Curve

As the name indicates, this is the yield curve shape that forms during normal market conditions. Normal market conditions occur when investors generally believe that there will be no significant changes in the economy, such as in inflation rates, and that the economy will continue to grow at a normal rate. During such conditions, investors expect higher yields for fixed income securities with long-term maturities that occur farther into the future. This is a normal expectation of the market because short-term instruments generally hold less risk than long-term instruments; the farther into the future the bond's maturity, the more time, and therefore, uncertainty, the bondholder faces before being paid back the principal. To invest in one instrument for a longer period of time, an investor needs to be compensated for undertaking the additional risk.

Figure 1.4: A Normal Yield Curve



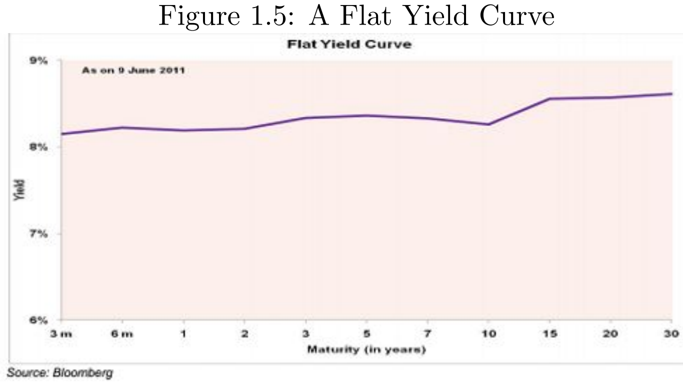
Source: Bloomberg

1.5.3.3.2 The Flat Yield Curve

These curves indicate that the market environment is sending mixed signals to investors, who are interpreting interest rate movements in various ways. During such an environment, it is difficult for the market to determine whether interest rates will move significantly in either direction into the future.

A flat yield curve usually occurs when the market is making a transition that emits different but simultaneous indications of what direction the interest rates will take. In other words, there may be some signals that short term interest rates will raise and other signals that long-term interest rates will fall. This condition will create a curve that is flatter than its normal positive trade off by choosing fixed-income securities with the least risk, or highest credit quality.

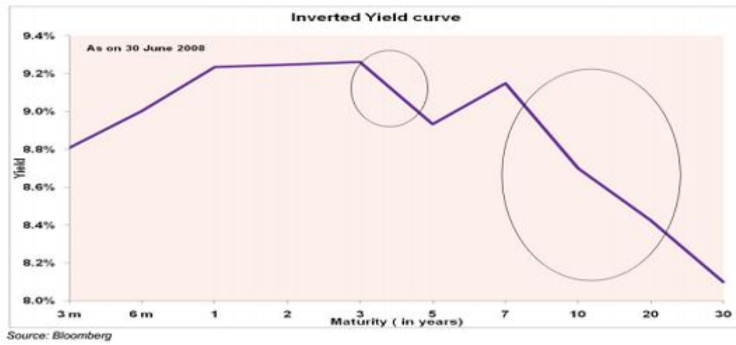
Figure 1.5: A Flat Yield Curve



1.5.3.3.3 The Inverted Yield Curve

An inverted yield curve indicates that investors interpret an inverted curve as an indication that the economy will soon experience a slowdown, which causes future interest rates to give even lower yields before a slowdown, thus creating a need to lock money into long-term investments at present prevailing yields, because it is expected that future yield will be even lower.

Figure 1.6: An Inverted Yield Curve



1.5.3.4 Yield Curves and Bond Prices

Given a set of zero-coupon bonds, spot rates can be derived directly from observed prices. For coupon-bearing bonds usually their yield to maturity or par yield only is quoted. The yield to maturity is its internal rate of return, that is, the constant interest rate r_k, that sets its present value equal to its price, and the equation is expressed as:

$$P_k = \sum_{i}^{n} \frac{CF_i}{(1 + r_k)^{t_i}} \tag{1.5.1}$$

where P_k is the price of bond k which generates n cash-flows (CFs) at periods t_i $(i = 1, 2, \ldots, n)$, r_k are the spot rates applicable on this bond and t_i depicts the maturity dates.

These cash flows consist of the coupon payments and the final repayment of the principal or face value. Yields to maturity on coupon bonds of the same maturity but with different coupon payments are not identical. Nevertheless, if the cash flow structure of a bond trading at the market (at par) is known, it is possible to derive from estimated spot rates, the coupon bonds theoretical yield to maturity, i.e. the rate the bond would require in order to trade at its face value (at par). Drawing on the spot rates $s_{t,m}$ the price equation can be expressed as:

$$P_k = \frac{C}{(1 + r_{t_1})} + \frac{C}{(1 + r_{t_2})^2} + \ldots + \frac{C}{(1 + r_{t_n})^n} + \frac{V}{(1 + r_{t_n})^n} = \sum_{i=1}^{i=n} \frac{C}{(1 + r_{t_i})^i} + \frac{V}{(1 + r_{t_n})^n} \tag{1.5.2}$$

where C represents the coupon payments, V represents the repayments of the principal, t_is and r_is as before. The yield to maturity of a coupon-bearing bond is therefore a geometric average of the spot rates which, in general, varies with the term to maturity.

1.5.3.4.1 Extraction of Yield Curve from Bond Prices

To derive the yield curve, the discount function is estimated by applying a (constrained) non-linear optimization procedure to data observed on a trade day. More important than the choice of a particular optimization method (e.g. maximum likelihood, non-linear least squares, generalized method of moments) is the decision whether the yield or price errors should be minimized. If one is primarily interested in minimizing the yield rates errors, they should minimize the deviation between the estimated and observed yields. The estimation proceeds in two stages: the discount function $D(t)$ is used to compute estimated prices and, secondly, estimated yields to maturity are calculated by solving the following equation for each coupon-bearing bond k (Svensson, L. E. , 1994):

$$P_k = \sum_{i=1}^{m} Cexp(-r_k i) + Vexp(-r_k m) \tag{1.5.3}$$

where r_k is the risk free rate. In practice, only a finite set of income securities trade, very few of which are zero-coupon bonds (Anderson, N. , 1996). Practitioners are therefore forced to use yield curves from coupon-paying bonds to extract zero-coupon yield curve.

Knowledge of historical and current values of yield rates is not enough for making sound investment decisions. The ability to transform this data into knowledge of patterns reliable to forecast future values is much more important. The ability to make reliable forecast is determined by the traders experience, intuition and intelligence. However, even a trader with the aforesaid qualities cannot always provide highly accurate forecasts perfectly, and this thesis explores decision-support possibilities.

In Kenya, methods of determining yield rates vary from one company to another. This is because there is no agreed method by the industry to generate similar rates, and the lack of standardization has led to non-uniform valuation of investment and development projects, creating loopholes in accountability whether using public or private resources.

Even in developed markets, we see the use of two distinct methods of generating yield rates: spline-based methods (which are heavily affected by the type of interpolation used) and parametric models. In the latter case, the entire yield curve is modelled using a single parametric function, with the parameters typically estimated through the use of least-squares regression technique. With spline-based models, on the other hand, the yield curve is made up of piecewise polynomials, where the individual segments are joined together continuously at specific points in time (called knot points). A good yield curve should have the following key properties:

1. It should be continuous and in some sense differentiable at all points;

2. It should exhibit monotonicity property in normal economic circumstances;
3. It should produce positive forward rates during normal economic conditions

The decision on whether to use parametric models or spline-based methods to generate the yield curve depends on the intended use of the yield curve. Parametric methods, particularly the (Svensson, L. E. , 1994) model, are very popular amongst Central Banks. This is because Central banks, when determining monetary policy, typically do not require yield curves to prices-back all products exactly. On the other hand, Investment Bankers usually prefer spline-based models because they are able to price-back all financial products.

1.5.3.4.2. Bond Pricing Formula

The prices of the set of trading instruments in the market from which the yield curve is calibrated are related to a discrete set of points along the yield curve. The instruments which are ideal to be used for Kenyan yield curve construction are the government bonds. Before calibration, we need to understand how the prices of these instruments are related to these points along the curve.

Elementary Relations

Let $F(t)$ represent the value that one unit of currency invested at time t_0, would be worth at time t. From elementary calculus it follows that:

$$F(t) = \begin{matrix} lim \\ n \to \infty \end{matrix} \left(1 + \frac{r_t}{n}\right)^{nt} = e^{(1+r_t t)} \tag{1.5.4}$$

Let $D(t)$ represents the value at time t_0 of one unit of currency to be received at time t. $D(t)$ is thus the inverse of$F(t)$ and is referred to as the present value at time t_0, of the zero-coupon bond maturing at time t. It follows from equation (1.5.4) that:

$$D(t) = e^{-r_t t} \tag{1.5.5}$$

Assume that an investor can invest $D(t_1)$ today, in a zero-coupon that pays one unit currency at time t_1. Furthermore, assume an investor can invest $D(t_2)$ today, in a zero-coupon bond that pays one unit of currency at time t_2. From the law of one price it must follow that:

$$D(t_1) \star D(t_0; t_1, t_2) = D(t_2) \tag{1.5.6}$$

where $D(t_0; t_1, t_2)$ represent the price at time t_0, of a zero-coupon bond to be purchased at time t_1, for maturity at t_2.

The discount factor $D(t_0;t_1,t_2)$ is called the forward discount factor from t_1 to t_2. If $f(t_1,t_2)$, represents the continuously compounded rate of interest, as observed at t_0, that an investor can earn from t_1 to t_2 then equation (1.5.5) implies that:

$$D(t_0;t_1,t_2) = e^{-f(t_1,t_2)(t_2-t_1)} \tag{1.5.7}$$

Equation (1.5.6) implies that:

$$D(t_0;t_1,t_2) = \frac{D(t_2)}{D(t_1)} \tag{1.5.8}$$

From equations (1.5.5), (1.5.7) and (1.5.8), it follows that:

$$f(t_1,t_2) = \frac{r_{t_2}t_2 - r_{t_1}t_1}{t_2 - t_1} \tag{1.5.9}$$

The forward rate $f(t_1,t_2)$ is called the discrete forward rate observed at time t_0 applicable to the period from t_1 to t_2. Consider rewriting t_1 and t_2 in equation (1.5.9), as t and $t+\varepsilon$ respectively. We then define $f(t)$, taking the limit as $\varepsilon \to 0$, and obtain that:

$$f(t) = \frac{d}{dt}r_t t, \tag{1.5.10}$$

$$\text{or equivalently:} \quad r_t = \frac{1}{t}\int_0^1 f(\tau)d\tau \tag{1.5.11}$$

The forward rate $f(t)$ is called the instantaneous forward rate observed at time t_0, applicable to time t. Finally, note that if $t \in [t_{i-1}, t_i)$ then it follows from equation (1.5.11) that:

$$r_t t = \int_0^{t_{i-1}} f(\tau)d\tau + \int_{t_{i-1}}^{t} f(\tau)d\tau$$

$$= r_{t_{i-1}}t_{i-1} + \int_{t_{i-1}}^{t} f(\tau)d\tau \tag{1.5.12}$$

This relationship allows us to calculate r_{t_i} given $r_{t_{i-i}}$. This relationship also shows us how to calculate forward rates given the spot rates, and vice versa. The other formula that is important is $D(t) = e^{-r_t t}$ which shows us the relationship between the spot rates and the discount rates, $D(t)$.

In conclusion, given the spot rates, one can calculate the forward rates and the discount rates, and given the forward rate, one can calculate the spot and consequently the discount rates, and finally, given the discount rate one can calculate both the spot and forward rates. We see that for us to be able to calculate the forward rate from the spot rates, we will need to have a spot curve that is differentiable at all points. The curve will assist us in generating the other two rates, which are used in pricing derivatives, and in this case, different options and coffee futures.

1.6 OUTLINE OF THIS THESIS

This thesis is divided into six chapters:

- Chapter 1 contains study definition, problem formulation, research goals and objectives, basic concepts which cover background information on yield curves construction and approaches used in pricing financial derivatives. We also have background information on coffee in Kenya, Nairobi Securities Exchange. Yield curves are discussed and we conclude the chapter with bond pricing formula.

- Chapter 2 introduces the improvement of *monotone preserving interpolation on r(t)t* method which is done by removing the non-differentiability introduced by Fritsch Butland Algorithm at the knot points. We calculate forward rates using numerical differentiation and compare the performance of *monotone preserving interpolation on r(t)t* method with both the numerical differentiation method and our new interpolation function.

- Chapter 3 surveys existing parametric models for yield curve construction and recommend the best among them, by testing the smoothness and accuracy of the models.

- Chapter 4 compares the best parametric model against our suggested improvement of *monotone preserving interpolation on r(t)t* method, and using smoothness and accuracy in order to conclusively suggest the best method for constructing ZCYC for the NSE.

- In Chapter 5, we present the application of the concepts of yield curve construction together with derivative pricing to construct a model for pricing coffee futures for the NSE. We apply the Hough Transform to estimate drift and volatility and the L-BFGS-B algorithm to estimate the other parameters.

- Chapter 6 contains general conclusions of the thesis, and possible directions for further research.

We expound further on the L-BFGS-B algorithm in appendix A. Appendix B contains Data used in the thesis, and the curves generated from different analysis of this data. Appendix C contains a list of papers presented at seminars and conference workshops and derived from this thesis.

All references in the text are listed in the References.

Chapter 2

IMPROVEMENT OF MONOTONE CONVEX INTERPOLATION METHOD ON r(t)t

There is no agreed-upon method used to construct yield curves at the Nairobi Securities Exchange. The existing practice is that each financial company uses in-house methods to construct the yield curves for their pricing and decision making. The most common yield curve used in the market was the one constructed by the Cannon Asset Managers Limited (CAM), a Kenyan company, in 2011. CAM used linear interpolation on the logarithms of the interest rates as their interpolation function. This method is not ideal because studies have shown that all variations of linear interpolations produce discontinuities in the forward rate curve.

To improve on the shortcomings of linear and cubic interpolations by ensuring not only a positive and (mostly) continuous forward rate curve but also a strictly decreasing curve of discount factors, the monotone convex interpolation method was introduced by (Hagan, P. S., & West, G. , 2006). Unfortunately, the model not only depends heavily on an appropriate interpolation algorithm but also produces discontinuity of f(t) under specific conditions. The monotone preserving r(t)t method was introduced to improve on the monotone convex method by ensuring that the knot points are estimated in the manner which ensures positivity and continuity in f(t) besides preserving the geometry of r(t)t. Unfortunately, monotone preserving method has the undesirable characteristic of not being differentiable at the knot-points.

In this chapter, we introduce an improvement on monotone preserving r(t)t interpolation method. We do this by removing the monotonicity constraint introduced by Hyman in the Fritsch Butland algorithm which introduces min/max functions in the interpolation model thus making the knot points nondifferentiable. In our model, we introduce new definitions within the function that ensure that the integrity and validity of the interpolation model are maintained, at the same time improving on the continuity of the resulting forward curve.

2.1 BACKGROUND INFORMATION

The Nairobi Securities Exchange (NSE) started trading in shares while Kenya was still a British colony in the 1920s, according to IFC/CBK. Share trading was initially conducted in an informal market that would facilitate access to long-term capital by private enterprises and also allow commencement of floating of local registered Government loans. In 1954, the NSE was constituted as a voluntary association of Securities brokers registered under the Societies Act (Exchange-NSE, N. S. , 1997). The newly established Securities Exchange was charged with the responsibility of developing the Securities market and regulating trading activities. Despite its long presence, the securities market is yet to make a significant contribution to the countrys development process, perhaps due to its slow development path.

The development path of Securities markets in both the emerging and developed world indicates an evolutionary process where changes in institutional infrastructure and the policy environment are witnessed as efforts are made to facilitate the growth of the Securities market. The evolutionary process is also characterized by a shift in the trading system from a periodic auction system to a continuous trading system.

Trading systems define the price discovery process or the transformation of latent demand of investors into realized transactions (Madhavan, A. , 1992). The evolutionary process of trading systems also indicates a shift from manual and decentralized settlement clearing systems to electronic and centralized settlement clearing [1]. It is argued that such a trading system enhanced efficiency in the price discovery process, provided liquidity at low costs, and had no excess volatility, thus was more desirable for the development of the Securities market ((Amihud, Y., Mendelson, H., & Lauterbach, B. , 1997)(Bessembinder, H., & Kaufman, H. M. , 1997)). High liquidity enhances long-term investment by reducing the required rate of returning and lowering the cost of capital to the issuers of securities. An efficient price discovery process enhances the role of the market in aggregating and conveying information through price signals, therefore making prices more informative.

Since the initiation of the Nairobi Securities Exchange in the 1920s, the development of the institutional and policy changes which have affected the market can be summarized as an evolutionary process with three stages of development. The initiation stage was mainly characterized by no for-

[1](Amihud, Y., Mendelson, H., & Lauterbach, B. , 1997) note the tendency for emerging markets to shift from periodic to continuous trading in the revitalization process.

mal rules and regulations and was dominated by foreign investors. In the formalization stage, a self-regulatory system was adopted while attempts were made to increase the participation of local investors in the post independent Government. In the third stage, various institutional and policy reforms were implemented to enhance the growth of the market.

Two main factors that shaped the development path were first, the political environment both at the local level and in the East African region which saw a change in the policy environment and changed the composition of market participation; and secondly the macroeconomic environment which instigated the demand for locally-mobilized long-term capital to enhance economic development.

Considering the developments in various aspects of the market, the following patterns were evident. First, the Securities brokerage industry expanded and diversified with the number of Securities brokers increasing from 6 to 21 (as at September 2014), and licensing of Securities dealers and investment banks also increased. However, Securities brokerage is yet to be fully negotiable while the role of Securities brokers remains predominately that of an agent.

Second, the composition of market participants shows a shift from a market dominated by foreign investors in the initiation stage to increased participation of local investors in the formalization stage (especially in the post-independence period), re-entry of foreign investors, though at limited level and mass education on Securities market operations and assets, in the revitalization stage.

Third, the NSE served as a regional market for the East African states in the initial stages, but the Securities Exchange lost a significant proportion of its market scope due to the political changes in these states. Attempts are however underway to establish a regional market to facilitate the expansion of the NSE market.

Fourth, though the diversity of securities traded is still minimal, the reform period saw efforts made to attain market depth with the introduction of new instruments. Fifth, while a discriminative tax policy that penalized share investors heavily was adopted in the post-independence period, an incentive-based tax policy regime was adopted in the reform period to enhance the competitiveness of the financial assets and reduce the barriers to listing new issues.

Sixth, to strengthen the regulatory framework, the regulatory system has witnessed a shift from the non-formal system to self-regulatory and statutory regulatory framework. Finally, the trading system shows a shift from the coffee-house forum to floor trading, and there were attempts to introduce the delivery versus payment system with the introduction of the Central Depository System and automation of the trading cycle. The main aim was to enhance efficiency in the price discovery process and liquidity of the market. These changes mimic development paths of other Securities markets in both the developed and emerging markets.

Many securities markets in emerging markets are on a healthy development path. This has led to some Futures Exchanges springing up in Latin America, Asia, and Africa, during the last decade, based on the premise that there is a need in the specific countries for a platform to be used to manage price volatility and provide price discovery. However, despite governments and donor agencies support, 2 out of 3 contracts traded in these markets fail as they have not been designed properly, according to the Capital Market Authority report. In June 2010 Budget speech, the government of Kenya, through a policy pronouncement by the then Deputy Prime Minister and Minister of Finance, announced that steps would be made towards developing institutional and legal frameworks to introduce a Commodities and Futures Exchange in Kenya. This will create a need to design a tool that can be used by the market to price the derivatives properly.

Currently, the NSE does not have a known existing yield curve as indicated by (Ngugi, R., & Agoti, J. , 2007). According to (Ngugi, R., & Agoti, J. , 2007), leading stockbrokers, a group of bank traders and institutional fund managers were working with Reuters Limited (a media channel dedicated to providing market data in the international financial market and disseminating that information locally) in 2006 towards developing an acceptable, credible market yield curve. This was driven by the requirement of International Accounting Standard (IAS) 39, which requires that there should be a standard yield curve to facilitate investors in fixed income instruments to value their portfolios at fair market values. It was then agreed that the yield rates applied to the market were to be derived from a cross-section of key market players. These key market players were required to post their quotes on Reuters assigned pages, out of which aggregate yields are derived. The yield curve was primarily available through Reuters. This method had major limitations, one of which was the use of bonds dirty prices as a component in yield curve construction. The market participants identified limitations in prices reported on the NSE and decided not to use the yield curve created at the time to value their portfolios.

In 2011, Cannon Asset Managers (CAM), a Kenyan based Asset Management company, created a yield curve for the NSE using the rates given by the Central Bank of Kenya (CBK), according to CAM. They used logarithmic linear interpolation method to calculate the yield rates at the periods missing from the CBKs bonds. Unfortunately, logarithmic linear interpolation has a tendency of implying discontinuities in the forward rate curve, a weakness depicted by all variations of linear interpolation methods, as shown by (Hagan, P. S., & West, G. , 2006).

2.2 LITERATURE REVIEW

Interpolation is a method of constructing new data points within the range of a discrete set of known data points (called knot points). The simplest method for interpolating between two points is by connecting them through a straight line. Some variations of linear interpolation are capable of ensuring a strictly decreasing curve of discount factors. However, all the variations of linear interpolation imply discontinuities in the forward rate curve.

To produce continuous forward rates curves, researchers introduced cubic methods of interpolation. An example of cubic interpolation algorithm is the cubic Hermite spline. Under cubic Hermite splines, the derivative of the data of each knot point is assumed to be known, and the interpolation function is required to be differentiable. Often, these derivatives will not be known and will have to be estimated. One method for estimating these derivatives, described by (De Boor, C., & Schwartz, B. , 1977) as the Bessel method involves estimating the derivative though the use of a three-point difference formula.

Unfortunately, all the traditional cubic methods are incapable of ensuring strictly positive forward rates, which are synonymous with non-decreasing discount factors, as shown by (Hagan, P. S., & West, G. , 2006). Furthermore, some cubic methods have an inherent lack of locality in the sense that a local perturbation of curve input data will cause changes in the data far away from the perturbed data point as shown by (Anderson, L. , 2007).

All variations of linear interpolations were seen to produce discontinuities in the forward rate curve, while all variations of cubic interpolations were seen to be incapable of ensuring strictly decreasing discount factors. Non-decreasing discount factors imply arbitrage opportunities, while discontinuous forward rates unacceptable from an economic perspective (unless the discontinuities occur on or around meetings of monetary authorities).

To counter this, a monotone convex interpolation method was developed, which it is claimed to be capable of ensuring a positive and (mostly) continuous forward rate curve (Hagan, P. S., & West, G. , 2006). This method proposed by (Hagan, P. S., & West, G. , 2006), was specifically designed to interpolate yield curve data and involves fitting a set of quadratic polynomials to a discrete set of estimated instantaneous forward rates. The method is designed such that f(t) preserves the shape of the set of discrete forward rates. The monotone convex method was also seen to be capable of ensuring a strictly decreasing curve of discount factors. Unfortunately, the model depends heavily on an appropriate interpolation algorithm. Also, it was discovered that there were specific conditions under which the interpolation function of the monotone convex interpolation would produce discontinuity f(t). (Du Preez, P. F. , 2011)

This led Du Preez to the monotone preserving r(t)t method of interpolation (Du Preez, P. F. , 2011). Essentially, this method involves applying cubic Hermite interpolation to the r(t)t at the knot points thereby ensuring that the values of f(t) at the knot points, are estimated in a manner which ensures positivity in f(t). Constructing an interpolating algorithm capable of preserving the monotonicity of the discount factors, was thus sufficient for ensuring positive forward rates.

(Bingham, N. H., Goldie, C. M., & Teugels, J. L. , 1989) suggested that the convolution of regularly varying probability densities is asymptotic to their sum, and therefore, also being regularly varying. However, (Bingham, N. H., Goldie, C. M., & Omey, E. , 2006) shows that while the Monotone Density Theorem allows us to 'diffrentiate asymptotic relations' and the weakest possible condition relaxing monotonicity is known, conditions of that type are awkward to handle in practice.

Monotone preserving r(t)t method is capable of ensuring a positive and continuous forward rate curve and was designed to preserve the geometry r(t)t. Monotonicity in the discount factors implies monotonicity in the r(t)t which is achieved by applying the work done in the field of shape preserving cubic interpolation, by authors such as (Akima, H. , 1991; De Boor, C., & Schwartz, B. , 1977; Fritsch, F. N., & Carlson, R. E. , 1980). Apart from being an improvement of the monotone convex method where it ensured positive forward rates, the monotone preserving r(t)t method was also capable of ensuring continuity of f(t).

In the study by (Du Preez, P. F. , 2011), they found that the monotone preserving r(t)t method to perform slightly better regarding stability, and continuity of f(t) than the monotone convex method. This suggests that when bootstrapping, the monotone preserving r(t) method could be the ideal method of interpolation. Unfortunately, monotone preserving method had the undesirable characteristic of not being differentiable at the knot-points.

2.3 METHODOLOGY

2.3.1 Extraction of Yield Curve from Bond Prices

To derive the yield curve, the discount function is estimated by applying a (constrained) non-linear optimization procedure to data observed on a trading day. More important than the choice of a particular optimization method (e.g. maximum likelihood, non-linear least squares, generalized method of moments) is the decision whether yield or price errors should be minimized. If one is primarily interested in minimizing the yield rates errors, it is suggested to minimize the deviation between the estimated and observed yields. In this case, the estimation proceeds in two stages: first, the discount function $d_{t,m}$ is used to compute estimated prices and, secondly, estimated yields to maturity are calculated by solving the following equation for each coupon-bearing bond k (Svensson, L. E. , 1994)

$$P_k = \sum_{i=1}^{m} Cexp(-r_k i) + Vexp(-r_k m) \tag{2.3.1}$$

where r_k is the risk-free rate.

In practice, only a finite set of income securities trade, very few of which are zero-coupon bonds (Anderson, L. , 2007). Practitioners are therefore forced to use yield curves from coupon-paying bonds to extract zero-coupon yield curve. Therefore , a nominal yield curve must exist from which the ZCYC can be extracted. This chapter concentrates on the construction of the nominal yield curve, with its novelty being in the improvement of the interpolation method used to the constructing the curve.

2.3.2 Dealing with Non-Differentiability at the Knot Points

In this section, we deal with the issue of non-differentiability first by approximating the derivative of a point on discrete data, using numerical differentiation method; secondly, we deal with this issue by using a mathematical model where the non-differentiability at the knot points created by use of Hyman monotonicity constraint in the monotone preserving interpolation on r(t)t method in (Du Preez, P. F. , 2011), is removed.

2.3.2.1 Numerical Differentiation

Little is usually taught about numerical differentiation. Perhaps that is because the processes should be avoided whenever possible. The reason for this can be seen like polynomials. High-degree polynomials tend to oscillate between the points of constraint. Since the derivative of a polynomial is itself a polynomial, it too will oscillate between the points of constraint, but perhaps not quite so wildly. To minimize this oscillation, one must use low degree polynomials which then tend to reduce the accuracy of the approximation. Another way to see the dangers of numerical differentiation is to consider the nature of the operator itself. Remember that

$$\frac{df(x)}{dx} = \lim_{\Delta x \to 0} \frac{f(x+\Delta x) - f(x)}{\Delta x} \tag{2.3.2}$$

Since there are always computational errors associated with the calculation of $f(x)$ and $f(x + \Delta x)$, they will remain as $\Delta x \to 0$ and the ratio will end up being largely determined by the computational error in $f(x)$. Therefore numerical differentiation should only be done if no other method for the solution of the problem can be found, and then only with considerable circumspection.

With these caveats clearly in mind, let us develop the formalisms for numerically differentiating a function f(x). We have to approximate the continuous operator with a finite operator and the finite difference operators described. We may approximate the derivative of a function f(x) by

$$\frac{df(x)}{dx} = \frac{\Delta f(x)}{\Delta x} \tag{2.3.3}$$

The finite difference operators are linear so that repeated operations with the operator lead to

$$\Delta^n f(x) = \Delta[\Delta^{n-1} f(x)] \tag{2.3.4}$$

This leads to the Fundamental Theorem of the Finite Difference Calculus which is 'The nth difference of a polynomial of degree n is constant ($a_n n! h^n$), and the $(n+1)$ difference is zero.'

The extent to which equation (2.3.4) is satisfied will depend partly on the value of h. Also, the ability to repeat the finite difference operation will depend on the amount of information available. To find a nontrivial nth order finite difference will require that the function is approximated by an nth degree polynomial which has n+1 linearly independent coefficients. Thus one will have to have knowledge of the function for at least n+1 points. For example, if one were to calculate finite differences for the function x^2 at a finite set of points x_i, then one could construct a finite difference table of the form:

Table 2.1: A Typical Finite Forrward Difference Table For $f(X) = x^2$

x_i	$f(x_i)$	$\delta f(x)$	$\Delta^2 f(x)$	$\Delta^3 f(x)$
2	f(2)=4			
		$\Delta f(2) = 5$		
3	f(3)=9		$\Delta^2 f(2) = 2$	
		$\Delta f(3) = 7$		$\Delta^3 f(2) = 0$
4	f(4)=16		$\Delta^2 f(3) = 2$	
		$\Delta f(4) = 9$		$\Delta^3 f(3) = 0$
5	f(5)=25		$\Delta^2 f(4) = 2$	
		$\Delta f(5) = 11$		
6	f(6)=36			

This table demonstrates the fundamental theorem of the finite difference calculus while pointing out an additional problem with repeated differences. While we have chosen f(x) to be a polynomial so that the differences are exact and the fundamental theorem of the finite difference calculus is satisfied exactly, one can imagine the situation that would prevail should f(x) only approximately be a polynomial.

2.3.2.1.1 Approximating derivatives from data

Suppose that a variable y depends on another variable x, i.e. y = f(x), but we only know the values of f at a finite set of points, e.g., as data from an experiment or a simulation: $(x_1, y_1), (x_2, y_2), ..., (x_n, y_n)$. Suppose then that we need information about the derivative of f(x). One obvious idea would be to approximate $f'(x_i)$ by the Forward Difference:

$$f'(x_i) = y_i' \approx \frac{y_{i+1} - y_i}{x_{i+1} - x_i} \tag{2.3.5}$$

This formula follows directly from the definition of the derivative in calculus. An alternative would be to use a Backward Difference

$$f'(x_i) \approx \frac{y_i - y_{i-1}}{x_i - x_{i-1}} \tag{2.3.6}$$

Since the errors for the forward difference and backward difference tend to have opposite signs, it would seem likely that averaging the two methods would give a better result than either alone. If the points are evenly spaced, i.e. $x_{i+1} - x_i = x_i - x_{i-1} = h$, then averaging the forward and backward differences leads to a symmetric expression called the Central Difference

$$f'(x_i) \approx y_i' \approx \frac{y_{i+1} - y_{i-1}}{2h} \tag{2.3.7}$$

2.3.2.1.2 Errors of approximation

We can use Taylor polynomials to derive the accuracy of the forward, backward and central difference formulas. For example, the usual form of the Taylor polynomial with remainder (sometimes called Taylors Theorem) is

$$f(x+h) = f(x) + hf'(x) + \frac{h^2}{2}f''(c) \tag{2.3.8}$$

Where c is some (unknown) number between x and x+h. Letting $x = x_i, x+h = x_{i+1}$ and solving for $f'(x_i)$ leads to

$$f'(x_i) = \frac{f(x_{i+1}) - f(x_i)}{h} - \frac{h}{2}f''(c) \tag{2.3.9}$$

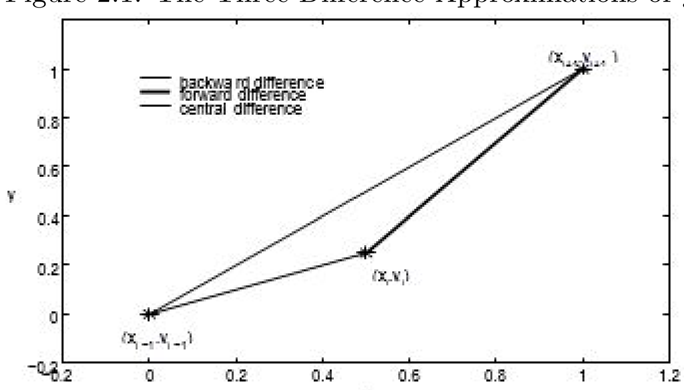
Notice that the quotient in this equation is exactly the forward difference formula. Thus the error of the forward difference is $-\frac{h}{2}f''(c)$ which means it is O(h). Replacing h in the above calculation by -h gives the error for the backward difference formula; it is also O(h). For the central difference, the error can be found from the third degree Taylor polynomials with remainder (sometimes called Taylors Theorem) is

$$f(x_{i+1}) = f(x_i + h) = f(x_i) + hf'(x_i) + \frac{h^2}{2}f''(x_i) + \frac{h^3}{3!}f'''(c_1) \tag{2.3.10}$$

$$f(x_{i-1}) = f(x_i - h) = f(x_i) - hf'(x_i) + \frac{h^2}{2}f''(x_i) - \frac{h^3}{3!}f'''(c_2) \tag{2.3.11}$$

where $x_i \leq c_1 \leq x_{i+1}$ and $x_{i-1} \leq c_2 \leq x_i$. Subtracting these two equations and solving for $f'(x_i)$

Figure 2.1: The Three Difference Approximations of y_i'



leads to

$$f'(x_i) = \frac{f(x_{i+1}) - f(x_{i-1})}{2h} - \frac{h^2}{3!}\frac{f'''(c_1) + f'''(c_2)}{2} \tag{2.3.12}$$

This shows that the error for the central difference formula is $O(h^2)$. Thus, central differences are significantly better and so: It is best to use central differences whenever possible.

2.3.2.2 Using Interpolation Function to Deal with Non-differentiability

In this chapter, we have discussed the improvement of the interpolation method, starting with the simple linear interpolation, moving onto cubic and spline interpolation, monotone convex interpolation and finally to Du Preez's monotone convex interpolation of r(t)t. The latter was considered to be the best interpolation method so far, leading to its application and use at Johannesburg Stock Exchange. This interpolation, however, has a weakness in that the curve generated is not differentiable at the knot points. It is a fact that continuity does not imply differentiability, but differentiability does imply continuity. In this section, we develop a new interpolation method which removes non-differentiability at the knot point. This is done by removing the monotonicity constraint introduced by Hyman and based on the Fritsch Butland algorithm. We concentrate on the forward rate function, whose knot point values can be calculated by simple bootstrapping of spot rates at the knot points which are given by the available tenures of bonds in the Kenyan market.

2.3.2.2.1 Definition of Forward Rate Function

We start with a mesh of data points $\{t_1, t_2, \ldots, t_n\}$ (we will think of these x- values as time points on the x-axis) and the corresponding y values are defined as $\{f_1, f_2, \ldots, f_n\}$ for a generic but unknown function f(t). Cubic splines are defined by the piece-wise cubic polynomial that passes through consecutive points:

$$f(t) = a_i + b_i(t - t_i) + c_i(t - t_i)^2 + d_i(t - t_i)^3 \tag{2.3.13}$$

With $t \in [t_i, t_{i+1}]$ and $i = 1, \ldots, n$. We will use the following definitions

$$h_i = t_{i+1} - t_i \tag{2.3.14}$$

$$m_i = \frac{f_{i+1} - f_i}{h_i} \tag{2.3.15}$$

With $i = 1, \ldots, (n-1)$ The coefficients a_i ,b_i,c_i and d_i depends on the details of the method, and are related to the values of f(t) and its derivatives at the knot points. In general

$$a_i = f(t_i) \equiv f_i, \quad b_i = f'(t_i), \quad c_i = \frac{f''(t_i)}{2} \quad \text{and so on} \tag{2.3.16}$$

2.3.2.2.2 The Derivatives

In the equation (2.3.15) above, the prime denotes the derivative of the interpolating function f(t) w.r.t.

its argument t . Moreover, given a_i and b_i , we can express c_i and d_i as follows:

$$c_i = \frac{3m_i - b_{i+1} - 2m_i}{h_i} \tag{2.3.17}$$

$$d_i = \frac{b_{i+1} - b_i - 2m_i}{h_i^2} \tag{2.3.18}$$

We note that a_i,c_i and d_i are defined using parameters which can easily be picked off tabular data of rates, maturities, and the term b_i. Now, given equation (2.3.16), how do we calculate the derivative of $f(x)$ with respect to the forward rate at point j?

$$f(t) = a_i + b_i(t - t_i) + c_i(t - t_i)^2 + d_i + (t - t_i)^3$$

$$\frac{\partial f(t)}{\partial f_j} = \frac{\partial a_i}{\partial f_j} + \frac{\partial b_i}{\partial f_j}(t - t_i) + \frac{\partial c_i}{\partial f_j}(t - t_i)^2 + \frac{\partial a_i}{\partial f_j}(t - t_i)^3 \tag{2.3.19}$$

Recall that:

$$h_i = t_{i+1} - t_i$$
$$m_i = \frac{f_{i+1} - f_i}{h_i}$$
$$a_i = f(t_i)$$
$$c_i = \frac{3m_i - b_{i+1} - 2m_i}{h_i}$$
$$d_i = \frac{b_{i+1} - b_i - 2m_i}{h_i^2}$$

We have:

$$\frac{\partial a_i}{\partial f_j} = \delta_i^j \tag{2.3.20}$$

$$\frac{\partial m_i}{\partial f_j} = \frac{1}{h_i}\delta_{i+1}^j - \frac{1}{h_i}\delta_i^j \tag{2.3.21}$$

$$\frac{\partial c_i}{\partial f_j} = \frac{1}{h_i}\left(3\frac{\partial m_i}{\partial f_j} - \frac{\partial b_{i+1}}{\partial f_j} - 2\frac{\partial m_i}{\partial f_j}\right) \tag{2.3.22}$$

$$\frac{\partial d_i}{\partial f_j} = \frac{1}{h_i^2}\left(\frac{\partial b_{i+1}}{\partial f_j} - \frac{\partial b_i}{\partial f_j} - 2\frac{\partial m_i}{\partial f_j}\right) \tag{2.3.23}$$

We note that for us to complete this differentiation, we need to compute the element $\frac{\partial b_i}{\partial f_j}$. Here δ_i^j is the Kronecker delta, which is equal to one if $i = j$ and zero otherwise. Once we are able to calculate the derivatives of theb_i coefficients then we will be able to get the derivative of the function.

This calculation is tricky if we use monotone preserving splines (or any other method which enforces monotonicity where b_i are non-differentiable functions of the f_js since they involve the min and max functions. This b_i are defined in such a way that they have min/max functions so as to enforce monotonicity of the forward curve. This definition of this part of the function makes it possible for the forward curve to be calculated using Fritsch Butland algorithm. This constraint was introduced by Hyman and enforced by (Hagan, P. S., & West, G. , 2006) and passed down to Monotone preserving interpolation of $r(t)t$. In this section, we will redefine b_i in a manner that gets rid of this constraint. There is a tradeoff, however, this means that the curve will not produce a decreasing curve, which would indicate the presence of negative interest rates, both spot and forward rates. This trade off, we feel, is realistic given some parts of the globe are already enforcing negative interest rates into their monetary policies, example: Japan, Denmark and European Central Bank.

Let us start by recalling the formulas for the b_is in the monotone preserving cubic spline method as defined in the(Hagan, P. S., & West, G. , 2006). First of all, at the boundaries, we have:

$$b_i = 0, b_n = 0 \tag{2.3.24}$$

For the internal data, if the curve is not a monotone at t_i , i.e $m_{i-1}.m_i \leq 0$, then the boundaries become:

$$b_i = 0 \quad (if m_{i-1}.m_i \leq 0) \tag{2.3.25}$$

So that it will have a turning point there. Instead if the trend is a monotone at t_i , i.e. $m_{i-1}.m_i > 0$, then one defines:

$$\beta_i = \frac{3m_{i-1}.m_i}{max(m_{i-1}, m_i) + 2min(m_{i-1}, m_i)} \tag{2.3.26}$$

and:

$$b_i = \begin{cases} min(max(0, \beta_i), 3min(m_{i-1}, m_i)) & if m_{i-1}.m_i > 0 \\ min(max(0, \beta_i), 3min(m_{i-1}, m_i)) & if m_{i-1}.m_i < 0 \end{cases} \tag{2.3.27}$$

The upper part of the definition of b_i is made when the curve is increasing (positive slopes), while the lower when the curve is decreasing (negative slopes). Equation (2.3.29) represents the monotonicity constraint introduced by Hyman and based on the Fritsch Butland algorithm. Due to the min/max condition in this equation, it was considered non-differentiable.

In this study, we proceed by using the following deductions:

$$\frac{\partial}{\partial f_j} max(0, \beta_i) = \begin{cases} 0 & if \beta_i < 0 \Leftrightarrow m_{i-1}, m_i < 0 \\ \frac{\partial \beta_i}{\partial f_j} & if \beta_i > 0 \Leftrightarrow m_{i-1}, m_i > 0 \end{cases} \tag{2.3.28}$$

$$\frac{\partial}{\partial f_j}min(0,\beta_i) = \begin{cases} 0 & if\beta_i > 0 \Leftrightarrow m_{i-1}, m_i > 0 \\ \frac{\partial \beta}{\partial f_j} & if\beta < 0 \Leftrightarrow m_{i-1}, m_i < 0 \end{cases} \tag{2.3.29}$$

Suppose first that the trend is increasing, i.e. $m_{i-1}, m_i, \beta_i > 0$. Using (2.3.29) we find:

$$\frac{\partial b_i}{\partial f_j} = \frac{\partial}{\partial f_j}\left[min(max(0,\beta), 3 \star min(m_{i-1}, m_i))\right] \tag{2.3.30}$$

$$= \begin{cases} \frac{\partial}{\partial f_j}max(0,\beta) & ifmax(0,\beta) < 3 \star min(m_{i-1}, m_i) \\ 3 \star \frac{\partial}{\partial f_j}min(m_{i-1}, m_i) & ifmax(0,\beta) > 3starmin(m_{i-1}, m_i) \end{cases} \tag{2.3.31}$$

Let us now suppose that the trend is decreasing instead, i.e. $m_{i-1}, m_i, \beta < 0$. By (2.3.31) we have:

$$\frac{\partial b_i}{\partial f_j} = \frac{\partial}{\partial f_j}\left[max(min(0,\beta), 3 \star max(m_{i-1}, m_i)\right] \tag{2.3.32}$$

$$= \begin{cases} \frac{\partial}{\partial f_j}min(0,\beta_i) & ifmin(0,\beta_i) > 3 \star max(m_{i-1}, m_i) \\ 3\frac{\partial}{\partial f_j}max(m_{i-1}, m_i) & ifmin(0,\beta_i) < 3 \star max(m_{i-1}, m_i) \end{cases} \tag{2.3.33}$$

We see that for any i and j

$$\frac{\partial}{\partial f_j}max(m_{i-1}, m_i) = \begin{cases} \frac{\partial m_{i-1}}{\partial f_j} & ifm_{i-1} > m_i \\ \frac{\partial m_i}{\partial f_j} & ifm_{i-1} < m_i \end{cases} \tag{2.3.34}$$

$$\frac{\partial}{\partial f_j}min(m_{i-1}, m_i) = \begin{cases} \frac{\partial m_{i-1}}{\partial f_j} & ifm_{i-1} < m_i \\ \frac{\partial m_i}{\partial f_j} & ifm_{i-1} > m_i \end{cases} \tag{2.3.35}$$

With the above definitions, then we are able to find the derivative of the function as follows:

$$\frac{\partial f(x)}{\partial f_j} = \frac{\partial a_i}{\partial f_j} + \frac{\partial b_i}{\partial f_j}(t-t_i) + \frac{\partial c_i}{\partial f_j}(t-t_i)^2 + \frac{\partial a_i}{\partial f_j}(t-t_i)^3$$

$$= \delta_i^j + \frac{\partial b_i}{\partial f_j} + \frac{1}{h_i}\left(3\frac{\partial m_i}{\partial f_j} - \frac{\partial b_{i+1}}{\partial f_j} - 2\frac{\partial m_i}{\partial f_j}\right) + \frac{1}{h_i^2}\left(\frac{\partial b_{i+1}}{\partial f_j} - \frac{\partial b_i}{\partial f_j} - 2\frac{\partial m_i}{\partial f_j}\right) \tag{2.3.36}$$

$$= \delta_i^j + \frac{\partial b_i}{\partial f_j} + \frac{1}{h_i}\left(3\left(\frac{1}{h_i}\delta_{i+1}^j - \frac{1}{h_i}\delta_i^j\right) - \frac{\partial b_{i+1}}{\partial f_j} - 2\left(\frac{1}{h_i}\delta_{i+1}^j - \frac{1}{h_i}\delta_i^j\right)\right) + \frac{1}{h_i^2}\left(\frac{\partial b_{i+1}}{\partial f_j} - \frac{\partial b_i}{\partial f_j} - 2\left(\frac{1}{h_i}\delta_{i+1}^j - \frac{1}{h_i}\delta^j\right)\right) \tag{2.3.37}$$

$$= \delta_i^j + \frac{\partial b_i}{\partial f_j} - \frac{1}{h_i}\frac{\partial b_{i+1}}{\partial f_j} + \frac{1}{h_i^2}\left[3\delta_{i+1}^j - 3\delta_i^j - 2\delta_{i+1}^j + 2\delta_i^j - \frac{\partial b_i}{\partial f_j} + \frac{\partial b_{i+1}}{\partial f_j}\right] + \frac{1}{h_i^3}\left[2\delta_{i+1}^j - 2\delta_i^j\right] \tag{2.3.38}$$

$$= \delta_i^j + \frac{\partial b_i}{\partial f_j} - \frac{1}{h_i}\frac{\partial b_{i+1}}{\partial f_j} + \frac{1}{h_i^2}\left[\delta_{i+1}^j - \delta_i^j - \frac{\partial b_i}{\partial f_j} + \frac{\partial b_{i+1}}{\partial f_j}\right] + \frac{2}{h_i^3}\left[\delta_{i+1}^j - \delta_i^j\right]$$

To complete the differentiation, we use the definitions of b_i to find its derivative with respect to forward rate at point j. Recall the earlier definition of b_i is as follows: at the boundaries, we have: $b_i = 0$, $b_n = 0$. For internal data at t_i , i.e. $m_{i-1}.m_i > 0$, then one defines:

$$\beta_i = \frac{3m_{i-1}.m_i}{max(m_{i-1}, m_i) + 2min(m_{i-1}, m_i)} \tag{2.3.39}$$

We take into account the monotonicity constraint, where:

$$b_i = \begin{cases} min(max(0, \beta_i), 3min(m_{i-1}, m_i)) & if m_{i-1}.m_i > 0 \\ min(max(0, \beta_i), 3min(m_{i-1}, m_i)) & if m_{i-1}.m_i < 0 \end{cases} \tag{2.3.40}$$

We are going to ignore the boundary values and concentrate on the internal values first, before taking the monotonicity constraint into account. Using both quotient and product rules, we can calculate the derivative as follows:

$$\frac{\partial b_i}{\partial f_j} = \begin{cases} \frac{\left[(max(m_{i-1},m_i)+2min(m_{i-1},m_i))\star\frac{m_{i-1}}{h_i}\left(\delta^j_{i+1}-\delta^j_i\right)+\frac{m_i}{h_{i-1}}\left(\delta^j_i-\delta^j_{i-1}\right)\right]-\left[(m_{i-1}*m_i)\star\frac{m_i}{h_{i-1}}\left(\delta^j_i-\delta^j_{i-1}\right)+2\frac{m_{i-1}}{h_i}\left(\delta^j_{i+1}-\delta^j_i\right)\right]}{(max(m_{i-1},m_i)+2min(m_{i-1},m_i))^2} \\ \frac{\left[(max(m_{i-1},m_i)+2min(m_{i-1},m_i))\star\frac{m_i}{h_i}\left(\delta^j_i-\delta^j_{i-1}\right)+\frac{m_{i-1}}{h_{i-1}}\left(\delta^j_{i+1}-\delta_i-^j\right)\right]-\left[(m_{i-1}*m_i)\star\frac{m_{i-1}}{h_{i-1}}\left(\delta^j_{i+1}-\delta^j_i\right)+2\frac{m_i}{h_i}\left(\delta^j_i-\delta^j_{i-1}\right)\right]}{(max(m_{i-1},m_i)+2min(m_{i-1},m_i))^2} \end{cases}$$

Calculation of the derivative of $b_i = \begin{cases} min(max(0, \beta_i), 3min(m_{i-1}, m_i)) & if m_{i-1}.m_i > 0 \\ min(max(0, \beta_i), 3min(m_{i-1}, m_i)) & if m_{i-1}.m_i < 0 \end{cases}$ is straight forward using the following expressions:

$$\frac{\partial}{\partial f_j} max(m_{i-1}, m_i) = \begin{cases} \frac{\partial m_{i-1}}{\partial f_j} & if m_{i-1} > m_i \\ \frac{\partial m_i}{\partial f_j} & if m_{i-1} < m_i \end{cases}$$

$$\frac{\partial}{\partial f_j} min(m_{i-1}, m_i) = \begin{cases} \frac{\partial m_{i-1}}{\partial f_j} & if m_{i-1} < m_i \\ \frac{\partial m_i}{\partial f_j} & if m_{i-1} > m_i \end{cases}$$

Solving $\frac{\partial b_i}{\partial f_j}$ enables us to solve the derivative of the forward function at the knot points, since it enables to us to calculate the derivative of the other coefficients in the forward function. This solves our problem of non-differentiability found in the monotone preserving convex on $r(t)t$.

2.4 COMPARISON OF THE CURVES IN TERMS OF STABILITY

In this section, we compare the curves generated by the methods discussed in this chapter, in terms of stability. We would like to know by how much the interpolated yield curve values change in other sections if we change the value of an input of at t_i. (Hagan, P. S., & West, G. , 2006) suggest measuring this noise feature on spot and forward rate curves via the following norms:

$$||M(r)|| = \begin{matrix} sup \\ t \end{matrix} \quad \begin{matrix} max \\ i \end{matrix} \left| \frac{\partial r(t)}{\partial r_i} \right| \tag{2.4.1}$$

$$||M(f)|| = \begin{matrix} sup \\ t \end{matrix} \quad \begin{matrix} max \\ i \end{matrix} \left| \frac{\partial f(t)}{\partial f_i} \right| \tag{2.4.2}$$

(Hagan, P. S., & West, G. , 2006) estimate these norms by calculating the maximum difference, in the supremum norm, between the original curve and any of the $2n$ curves obtained by changing any of the nodes up and down by one basis point. These differences can be estimated by testing at discrete points along the entire curve. The estimated norms are then expressed n terms of basis points.

To gauge the stability of the interpolation methods considered in this section, we calculated $||M(r)||$ and $||M(f)||$, for a set of Kenyan bond curves spanning the period from 31 July 2005, to 4 February 2012. $||M(r)||$ and $||M(f)||$ were estimated by testing at discrete points along the entire curve, in steps of one week each. For each of the curves under consideration, we used the same inputs as those that were used to construct the corresponding perfect fit bond curves. The method that has a mean of rates which reflect the market rates and the lowest standard deviation of the rates is deemed the best for both $||M(r)||$ and $||M(f)||$.

Table 2.2: Statistics for $||M(r)||$ obtained by bootstrapping a set of Kenyan bond curves, under various methods of interpolation

Method	**Mean**	**Std Deviation**
Monotone Preserving Interpolation on r(t)t	26.51575	2.85714
Operator Form	19.27461	1.34142
This Thesis new interpolation Model	19.27461	1.34142

Table 2.3: Statistics for $||M(f)||$ obtained by bootstrapping a set of Kenyan bond curves, under various methods of interpolation

Method	**Mean**	**Std Deviation**
Monotone Preserving Interpolation on r(t)t	28.95963	11.93328
Operator Form	23.84276	7.73945
This Thesis new interpolation Model	20.063951	3.81627

Table 2.4: Inputs in the Kenyan Bond Curve

Financial Instrument	**Rate**
Kenyas Central Bank Lending Rate	17.960%
FXD2/2011/2	7.439%
FXD2/2010/10(R1)	9.307%
FXD1/2009/5	9.750%
FXD4/2008/5	9.500%
FXD2/2007/5	9.500%
FXD2/2006/6	11.500%

The tables 2.2, 2.3 and 2.4 show that using the operator form to calculate the forward rates and using the model constructed in this thesis outperform the *monotone preserving interpolation on r(t)t.* Note that in FXD2/2011/2, FXD refers to fixed income securities like bonds, the first 2 implies that it is the second same-tenure bond issued in the year, 2011 is the year that the bond was issued by the CBK and the last 2 is the tenure of the bond. The Central Bank lending rate is expressed as a simple annual rate, whilst the yields on the set of bonds are nominal annual rates, compounded semi-annually. This means, for example, the annualized rate of FXD2/2011/2 is $i^{(2)} = 14.878$.

2.5 CHAPTER CONCLUSION

We introduce two ways of dealing with the non-differentiability at the knot points. We use numerical methods to approximate the derivatives at the knot points and also develop the mathematical function which removes the non-differentiability at the knot point. We then generate the forward rates using the two methods and compare how closely they relate. This result is important for practical as well as conceptual reasons.

The non-differentiability model is important for those in the financial industry for generating forward curves to price derivatives and risk management of interest rate derivatives, especially if they are not well versed with numerical methods. We compare the models performance with monotone preserving interpolation of r(t)t regarding stability, and the model developed in this thesis produces rates that do not deviate far from the observed market data, even when the perturbation is introduced at a point.

Chapter 3

FINDING THE BEST PARAMETRIC MODEL FOR THE NAIROBI SECURITIES EXCHANGE

The main objective of this study is to construct a zero coupon yield curve (ZCYC) for the Nairobi Securities Exchange (NSE). In this chapter, we use the classical Nelson-Siegel model, Svensson Model, and Rezende-Ferreira models. These models have linear and nonlinear guidelines making them have multiple local minima. This condition causes model estimation more difficult to estimate. We, therefore, use L-BFGS-B method as the optimization approach for estimating the models. Our contribution is twofold: estimation of the parameters using the L-BFGS-B method and using Gauss-Newton numerical method to develop a model for pricing the bonds given the forward and the rate curves equations.

We compare the models' performance regarding continuity and differentiability of the ZCYC, positivity of the forward curve and accuracy in reflecting the observed market bond prices. We use bond data from Central Bank of Kenya (CBK). The best parametric model to be used for the Kenyan securities market and, consequently, the East African Securities markets is chosen if and only if it depicts the aforementioned qualities.

3.1 INTRODUCTION

The definition of yield rate, also called Yield to Maturity (YTM), is the true rate of return an investor would receive if the security were held to maturity. When the YTM is expressed as a function of maturity, then it is known as the term structure of interest rates. A yield curve is the graphical plotting of the yield rate function. The yield curve is one of the most important indicators of the level and changes in interest rates in the economy and hence the interest in studying as well as accurately modeling it.

The Yield to Maturity (YTM) can also be defined as the single discount rate on an investment that equates the sum of the present value of all cash flows to the current price of the investment. However, using a single discount rate at different time periods is problematic because it assumes that all future cash flows from coupon payments will be reinvested at the derived YTM. This assumption neglects the reinvestment risk that creates investment uncertainty over the entire investment horizon. Another shortcoming of YTM is that the yields of bonds on the maturity depend on the patterns of their cash flows, which is often referred to as the coupon effect. As a result, the YTM of a coupon bond is not a good measure of the pure price of time and not the most appropriate yield measure in the term structure analysis.

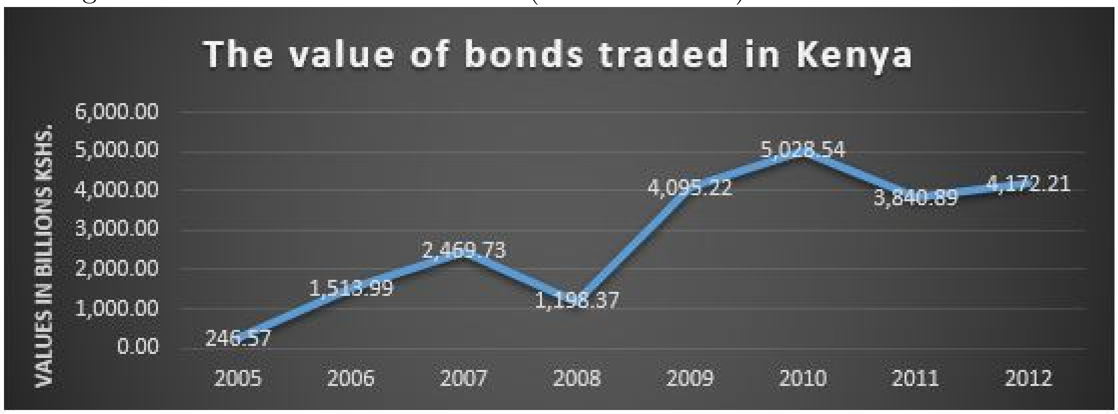
In comparison, zero-coupon securities eliminate the exposure to reinvestment risks because there is no cash flow before maturity to be reinvested. The yields on the zero-coupon securities, called the spot rate, are therefore not affected by the coupon effect since there are no coupon payments. Also, unlike the yield to maturity, securities having the same maturity have theoretically the same spot rates, which provide the pure price of time. As a result, it is preferable to work with zero-coupon yield curves (ZCYC) rather than YTM when analyzing the yield curve.

Various methods exist for estimating zero-coupon yield curves. The most adopted methods are by (Nelson, C. R., & Siegel, A. F. , 1987) method or the extended versions of the same, as suggested by (Svensson, L. E. , 1994) and (Rezende, R. B. , 2011). We are going to illustrate the application of these models in deriving the zero-coupon yield curve for the Nairobi Securities Exchange (NSE).

3.2 BACKGROUND INFORMATION

Kenyan bonds and the T-bill market has a noticeably smaller trading volume and is not liquid. To finance national developments projects, the government issues bonds to investors; this market has traded more and more volume of these securities in both the primary and secondary markets as the year's progress. In 2005, the trading volume of bonds in the secondary market was Kshs. 246.57 billion, as compared to Kshs. 4.172 trillion in 2012, showing that the Kenyan bond market has truly expanded. Figure 3.1 shows the movement of the value of Treasury bonds traded.

Figure 3.1: Movement of bond value (in billions Kshs) traded from 2005 - 2012



Before 2000, only treasury bills were available at the primary market, and virtually no bond market existed. The issuance of securities was not auction based, and there were no market development initiatives. From 2001, the composition of debt portfolio changed to 76:24 the ratio indicating Treasury bills to bonds. The average maturity of debt was at eight months. Then bonds were introduced with the key objective of lengthening the maturity of securities and minimizing refinancing risk. Auction based issuance was adopted to promote price discovery and development of a yield curve. In addition to this, the Market Leaders Forum (MLF) was formed so as to support the development of the bond market in Kenya.

This led to increased trading in bonds after 2013, the composition of debt portfolio reversed to 26:74, T-bills to bonds. The average maturity of all securities moved to about seven years while bonds maturity alone was five years. The longest bond in the market is 30-year, issued in 2011. A 20-year was first issued in 2008, which was followed by a 25-year in 2010. Multi price auction method was introduced which increased bond market activity, thus providing initial pricing for trading.

According to 2013 CBK report, the CBK plans to start what it dubs as Benchmark Bonds Programme.' One of the objectives of the program is to eliminate bond fragmentation at the secondary market and development of a firm, reliable yield curve. This study aims to be one of the tools the CBK will use in meeting this objective by suggesting the best parametric model that should be used in the pricing of the Kenyan bonds.

3.3 LITERATURE SURVEY

Many estimation methods for yield curves have appeared in literature over the years. Generally speaking, there are two distinct approaches to estimate the term structure of interest rates: the equilibrium model and the statistical techniques.

The first approach is formalized by defining state variables characterizing the state of the economy (relevant to the determination of the term structure) which are driven random processes and are related in some way to the prices of the bonds. It then uses no-arbitrage arguments to infer the dynamics of the term structure. Examples of this approach include (Brennan, M. J., & Schwartz, E. S. , 1979; Cox, J. C., Ingersoll Jr, J. E., & Ross, S. A. , 1985; Dothan, L. U. , 1978; Duffie, D., & Kan, R. , 1996; Vasicek, O. , 1977).

Unfortunately, concerning the expedient assumptions about the nature of the random process driving the interest rates, the yield derived by those models have a specific functional form dependent only on a few parameters, and usually, the observed yield curves exhibit more varied shapes than those justified by the equilibrium models.

In contrast to the equilibrium models, statistical techniques focusing on obtaining a continuing yield curve from cross-sectional coupon bond data based on curve fitting techniques can describe a richer variety of yield patterns in reality. The resulting term structure estimated from the statistical techniques can be directly put into the interest rate models such as the, for pricing interest rate contingent claims. Since a coupon bond can be considered as a portfolio of discount bonds with maturities dates consistent with the coupon dates, the discount bond prices can thus be extracted from the actual coupon bond prices by statistical techniques [1]. These techniques can be broadly divided into two categories: the splines and the parsimonious function forms; see (Alper, C. E., Akdemir, A., & Kazimov, K. , 2004). In this paper, we will concentrate on the latter.

Parsimonious models specify a parsimonious parameterizations of the discount function, spot rate or the implied forward rate. Moving from the cubic splines, (Chambers, D. R., Carleton, W. T., & Waldman, D. W. , 1984) introduced the parsimonious function forms by considering an exponential

[1]Once the discount function, P(t), is defined, the spot interest rate (the pure discount bond yield) can be computed by:

$$R(t) = \frac{-lnP(t)}{t}$$

polynomial to model the discount function.(Nelson, C. R., & Siegel, A. F. , 1987) Followed shortly after that by choosing an exponential function with four unknown parameters to model the forward rate of U.S Treasury bills. By considering the three components that make up this function, (Nelson, C. R., & Siegel, A. F. , 1987) illustrated that it could be used to generate a variety of shapes for the forward rate curves and analytically solve for the spot rate. Moreover, the advantage of the classical Nelson-Siegel model is that the three parameters may be interpreted as latent level, slope and curvatures factors. (Diebold, F. X., & Li, C. , 2006; Modena, M. , 2008; Tam, C.-S., & Yu, I.-W. , 2008) employed the Nelson-Siegel interpolant to examine bond pricing with a dynamic latent factor approach and concluded that it was satisfactory.

(Svensson, L. E. , 1994) Increased the flexibility of the original Nelson and Siegel model by adding two extra parameters (Svensson, L. E. , 1994) model which allowed for a second hump in the forward rate curve. Later, (Bliss, R. R. , 1996) introduced the Extended Nelson-Siegel method, which introduced a new appropriating function with five parameters by extending the model developed by (Nelson, C. R., & Siegel, A. F. , 1987). Bliss suggested that a six-parameter model can produce better results for fitting the term structure with longer maturities.

The Nelson-Siegel model class has linear and non-linear parameters depending on the values assumed fixed. Due to this, these models have multiple local minima making model estimation difficult. Previous studies have widely discussed the estimation of Nelson-Siegel model class, and they are: (Bolder, D. J., & Strliski, D. , 1999; Rezende, R. B. , 2011; Gilli, M., Groe, S., & Schumann, E. , 2010; Maria, L. M., Leanez, C., & Moreno, M. , 2009; Rosadi, D. , 2011), among others.

Previous literature indicates that although there are a lot of curve fitting models that have been successfully applied to developed bond markets, comparatively little attention has been paid to emerging markets [2]; (Alper, C. E., Akdemir, A., & Kazimov, K. , 2004). To bridge this gap, a study by (Subramanian, K. V. , 2001) discussed the concept of weighted parameter optimization for the emerging and developed markets. In an illiquid market like India where only about a handful of liquid securities get traded in a day (which is very similar to Kenyan market), illiquid bonds must also be included

[2] The developed bond markets are well established and comprised of relatively liquid securities with short and long maturities. However, in the developing economies with sparse bond market price data, a substantial portion of the secondary market trading is contracted in a handful of bonds that the market perceives liquid; thus it is not meaningful to estimate the term structure based on a small number of liquid securities. (Subramanian, K. V. , 2001) was the pioneer in positing a model for the yield curve estimation based on liquidityweighted objective functions to ensure that liquid bonds in the market are priced with smaller errors than the illiquid bonds.

in the estimation procedure. Hence the estimation methods must incorporate the effect of liquidity premiums on illiquid bonds [3].

3.4 EMPIRICAL METHODOLOGY

3.4.1 Model Selection

3.4.1.1 The Nelson-Siegel (1987) Model

The Nelson-Siegel model sets the instantaneous forward rate at maturity m given by the solution to a second order differential equation with unequal roots as follows:

$$f(m) = \beta_0 + \beta_1 exp\left(\frac{-m}{\tau_1}\right) + \beta_2 \frac{m}{\tau_1} exp\left(\frac{-m}{\tau_1}\right) \tag{3.4.1}$$

where $m > 0$, is the time to maturity of a given bond. Equation (3.4.1) consists of three parts: A constant, an exponential decay functional and Laguerre function. β_0 is independent of m and as much, β_0 is often interpreted as the level of long term interest rates. The exponential decay function approaches zero as m tends to infinity and β_1 as m tends to zero. The effect of β_1 is thus only felt at the short end of the curve. The Laguerre function on the other hand approaches zero as m tends to infinity, and as m tends to zero. The effect of β_2 is thus only felt in the middle section of the curve, which implies that β_2 adds a hump to the yield curve or a U if it is negative. The time constant τ_1 is the scale parameter that measures the rate at which the short term and medium term components decay to zero. For example, small value of τ result in rapid decay in the predictor variables and therefore they will be suitable for curvature at low maturities. Corresponding, large volumes of τ produce slow decay in the predictor variables and will be suitable for curvature over longer maturities, (Christofi, A. C. , 1998).

Following (McCulloch, J. H. , 1971) definition of the yield as an average of the forward rates, the spot interest for maturity m can be derived by integrating equation (3.4.2) from zero to m and dividing by m . The resulting function can be expressed as follows:

$$r(m) = \beta_0 + \beta_1 \frac{\tau_1}{m}\left[1 - exp\left(\frac{-m}{\tau_1}\right)\right] + \beta_2 \frac{\tau_2}{m}\left[1 - exp\left(\frac{-m}{\tau_2}\right)\right] - \beta_2 exp\left(\frac{-m}{\tau_2}\right) \tag{3.4.2}$$

[3] We attempt to estimate the parameters by minimizing the mean absolute deviation between the observed and calculated prices. The weights have been assigned according to the liquidity of individual securities.

From equations (3.4.2) and (3.4.3) it follows that both the spot and forward rate function reduce to $\beta_0 + \beta_1$ as $m \rightarrow 0$. Furthermore, we have $\lim_{m \rightarrow 0} r(m) = \lim_{m \rightarrow 0} f(m) = \beta_0$. Thus, in the absence of arbitrage, we must have that $\beta_0 > 0$ and $\beta_0 + \beta_1 > 0$.

Due to the local-minima problem which makes model estimation difficult in the Nelson-Siegel model and the inadequacy of the calibration methods used so far, we propose Non-Linear Least Squares (NLS) used to find parameter values for non-linear functions- estimation method with L-BFGS-B optimization approach (see Appendix A for further information on L-BFGS-B algorithm). This optimization method is an extension of the limited memory BFGS method (LM-BFGS or L-BFGS) which uses simple boundaries model, according to (Zhu, C., Byrd, R. H., Lu, P., & Nocedal, J. , 1994).

Using L-BGFS-B algorithm, we can estimate the above five parameters: $\psi \equiv \{\beta_0, \beta_1, \beta_2, \tau_1, \tau_2\}$, embedded in the (Nelson, C. R., & Siegel, A. F. , 1987) model, and hence calculate the price of the bond using the following nonlinear constrained optimization estimation procedure based the Gauss-Newton numerical method:

$$P_i = \sum_{m=1}^{T} \left(\frac{CF_{im}}{\left\{1 + \beta_0 + \beta_1 \left(\frac{\tau_1}{m}\right) \left[1 - exp\left(\frac{-m}{\tau_1}\right)\right] + \beta_2 \left(\frac{\tau_2}{m}\right) \left[1 - exp\left(\frac{m}{\tau_2}\right) \left(\frac{m}{\tau_2} + 1\right)\right]\right\}^m} \right) + \varepsilon_i \quad (3.4.3)$$

where P_i is the price of bond i.

3.4.1.2 The Svensson (1994) Model

To increase the flexibility and improve the fitting performance, (Svensson, L. E. , 1994) extends (Nelson, C. R., & Siegel, A. F. , 1987) instantaneous forward rate function by adding a fourth term, a second hump (or trough) $\beta_3 \frac{m}{\tau_2} exp\left(\frac{-m}{\tau_2}\right)$, with two additional parameters β_3 and τ_2. The forward rate function is then set as:

$$f(m) = \beta_0 + \beta_1 exp\left(\frac{-m}{\tau_1}\right) + \beta_2 \left(\frac{m}{\tau_1}\right) exp\left(\frac{-m}{\tau_1}\right) + \beta_3 \left(\frac{m}{\tau_2}\right) exp\left(\frac{-m}{\tau_2}\right) \quad (3.4.4)$$

where the unknown parameters β_0,β_1,β_2 and τ_2 have the same economic interpretation as the Nelson Siegel model and the two additional parameters β_3 and τ_2 denote the same meaning as β_2 and τ_1. The spot rate, derived by integrating the forward rate, is given by:

$$R(m) = \beta_0 + \beta_1 \left(\frac{\tau_1}{m}\right) \left[1 - exp\left(\frac{-m}{\tau_1}\right)\right] + \beta_2 \left(\frac{\tau_1}{m}\right) \left[1 - exp\left(\frac{-m}{\tau_1}\right)\left(\frac{m}{\tau_2} + 1\right)\right] + \beta_3 \left(\frac{\tau_2}{m}\right) \left[1 - exp\left(\frac{-m}{\tau_2}\right)\left(\frac{m}{\tau_2 + 1}\right)\right] \quad (3.4.5)$$

Similarly, using L-BGFS-B algorithm, we can estimate the parameters and calculate the price using the following equation:

$$P_i = \sum_{m=1}^{T} \left(\frac{CF_{im}}{\left\{1 + \beta_0 + \beta_1 \left(\frac{\tau_1}{m}\right) \left[1 - exp\left(\frac{-m}{\tau_1}\right)\right] + \beta_2 \left(\frac{\tau_1}{m}\right) \left[1 - exp\left(\frac{m}{\tau_1}\right)\left(\frac{m}{\tau_1} + 1\right)\right] + \beta_3 \left(\frac{\tau_2}{m}\right) \left[1 - exp\left(\frac{m}{\tau_2}\right)\left(\frac{m}{\tau_2} + 1\right)\right]\right\}^m} \right) + \varepsilon_i \quad (3.4.6)$$

3.4.1.3 The Rezende-Ferreira (2011) Model

(Rezende, R. B. , 2011) decided to increase the accuracy of the (Svensson, L. E. , 1994) by adding a fifth term, a third hump (or trough) $\beta_4 \frac{m}{\tau_3} exp\left(\frac{-m}{\tau_3}\right)$, with two additional parameters β_4 and τ_3. The forward rate function is then set as:

$$f(m) = \beta_0 + \beta_1 exp\left(\frac{-m}{\tau_1}\right) + \beta_2 \left(\frac{m}{\tau_1}\right) exp\left(\frac{-m}{\tau_1}\right) + \beta_3 \left(\frac{m}{\tau_2}\right) exp\left(\frac{-m}{\tau_2}\right) + \beta_4 \left(\frac{m}{\tau_3}\right) exp\left(\frac{-m}{\tau_3}\right) \quad (3.4.7)$$

where the unknown parameters β_0,β_1,β_2 and τ_2 have the same economic interpretation as the Nelson Siegel model and the two additional parameters β_3 and τ_2 denote the same meaning as β_2 and τ_1. The spot rate, derived by integrating the forward rate, is given by:

$$R(m) = \beta_0 + \beta_1 \left(\frac{\tau_1}{m}\right) \left[1 - exp\left(\frac{-m}{\tau_1}\right)\right] + \beta_2 \left(\frac{\tau_1}{m}\right) \left[1 - exp\left(\frac{-m}{\tau_1}\right)\left(\frac{m}{\tau_2} + 1\right)\right] + \beta_3 \left(\frac{\tau_2}{m}\right) \left[1 - exp\left(\frac{-m}{\tau_2}\right)\left(\frac{m}{\tau_2 + 1}\right)\right] + \beta_4 \left(\frac{\tau_3}{m}\right) \left[1 - exp\left(\frac{-m}{\tau_3}\right)\left(\frac{m}{\tau_3 + 1}\right)\right] \quad (3.4.8)$$

Similarly, using L-BGFS-B algorithm, we can estimate the parameters and calculate the price using the following equation:

$$P_i = \sum_{m=1}^{T} \left(\frac{CF_{im}}{\left\{1 + \beta_0 + \beta_1 \left(\frac{\tau_1}{m}\right) \left[1 - exp\left(\frac{-m}{\tau_1}\right)\right] + \beta_2 \left(\frac{\tau_1}{m}\right) \left[1 - exp\left(\frac{m}{\tau_1}\right)\left(\frac{m}{\tau_1} + 1\right)\right] + \beta_3 \left(\frac{\tau_2}{m}\right) \left[1 - exp\left(\frac{m}{\tau_2}\right)\left(\frac{m}{\tau_2} + 1\right)\right] + \beta_4 \left(\frac{\tau_3}{m}\right) \left[1 - exp\left(\frac{-m}{\tau_3}\right)\left(\frac{m}{\tau_3 + 1}\right)\right]\right\}^m} \right) + \varepsilon_i \quad (3.4.9)$$

3.4.2 Liquidity-Weighted Function

In yield curve construction, errors are caused by two reasons: (a) curve fitting mistakes and (b) presence of a liquidity premium. The errors due to curve fitting arise from the calculations and can be avoided. But an error due to the presence of liquidity premium reflects market conditions and cannot be ignored.

Since the reliability of the term structure estimation depends heavily on the precision of the market prices according to (Subramanian, K. V. , 2001), liquid and illiquid securities are a heterogeneous class and including them both in the term structure estimation process poses problems. Illiquid Bonds are traded at a premium to compensate for their undesirable attribute regarding a low price. Assigning equal weights to both types of errors will give undue weight to the kind of error that creeps in due to curve fitting.

(Subramanian, K. V. , 2001) suggests a liquidity weighted objective function, which hypothesizes that a weighted error function (with weights based on liquidity) would lead to better estimation that equal weights to the squared errors of all securities. We, therefore, model the liquidity using a function with two factors: the volume of trade in a security and the number of trades in that security.

The weight of the i^th security W_i is given by:

$$W_i = \frac{\left[\left(1 - e^{\frac{-v_i}{v_{max}}}\right) + \left(1 - e^{\frac{-n_i}{n_{max}}}\right)\right]}{\sum_i W_i} = \left(1 - e^{\frac{-v_i}{v_{max}}}\right) + \left(1 - e^{\frac{-n_i}{n_{max}}}\right) \tag{3.4.10}$$

$$W_i = \frac{\left[\left(tanh\left(\frac{-v_i}{v_{max}}\right)\right) + \left(tanh\left(\frac{-n_i}{n_{max}}\right)\right)\right]}{\sum_i W_i} = \left(tanh\left(\frac{-v_i}{v_{max}}\right)\right) + \left(tanh\left(\frac{-n_i}{n_{max}}\right)\right) \tag{3.4.11}$$

where v_i and n_i are the volume of trade and the number of trades in the i^th security respectively, while v_{max} and n_{max} are the maximum number of trades among all the securities traded for the day respectively.

As given in the equations (3.4.10) and (3.4.11) above, it ensures that the weights of the relative liquid securities would not be significantly different from each other. For the illiquid securities, however the weights would fall quickly as liquidity decreased.

The final error-minimizing function, which should equal to zero, is given by:

$$Min \sum_{i=1}^{n} \left(w_i \left(P_i - \hat{P_i}\right)^2\right) = Min \sum_{i=1}^{n} \left(w_i \varepsilon_i^2\right) = 0 \tag{3.4.12}$$

3.4.3 Test Statistics

In academic literature, there are two distinct approaches used to indicate the term structure fitting performance. One is the flexibility of the curve (accuracy), and the other focuses on smoothness of the yield curve. Although there are numerical methods proposed to estimate the term structure, any method developed has to grapple with deciding the extent of the above trade-off. Hence it becomes a crucial issue to investigate how to reach a compromise between the flexibility and smoothness.

Three simple summary statistics which can be calculated for the flexibility of the estimated yield curve are the coefficient of determination, root mean squared percentage error, and root mean squared error. These are calculated as:

3.4.3.1 The Coefficient of Determination (R^2)

This is the test statistic used to measure the accuracy of the curve, or in other words, how well the model fits the data. The formula is as expressed below:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}((P_i - \hat{P_i})^2)/(n-k)}{\sum_{i=1}^{n}((P_i - \hat{P_i})^2)/(n-1)} \tag{3.4.13}$$

where P_i is the mean average price of all observed bonds, $\hat{P_i}$ is the model price of a bond i, n the number of bonds traded and k is the number of parameters needed to be estimated.

Roughly speaking, with the same analysis in regression, we associate a high value of R^2 with a good fit of the term structure and associate a low R^2 with a poor fit.

3.4.3.2 Root Mean Squared Error (RMSE) and Mean Squared Percentage Error (RMSPE)

Denoted as the RMSE and RMSPE, these two statistics test the flexibility of a curve. A low value for these measures is assumed to indicate that the model is flexible, on average, and is able to fit the yield curve, and vice versa.

Flexibility is also a measure of accuracy. The formulas for these tests are:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(P_i - \hat{P_i})^2} \tag{3.4.14}$$

$$RMSPE = 100 \star \sqrt{\frac{1}{n}\sum_{i=1}^{n}(P_i - \hat{P_i})^2} \tag{3.4.15}$$

3.4.3.3 Testing for Smoothness

To test the smoothness of the estimated yield curve, we use a modified statistic suggested by (Adams, K. J., & Van Deventer, D. R. , 1994) to reach the maximum smoothness for forward rate curve, and denote the smoothness (Z) for the estimated yield curve as:

$$Z = \sum_{t=1}^{n}\left(\left[f(t) - f(t-\frac{1}{2})\right] - \left[f(t-\frac{1}{2}) - f(t-1)\right]\right)^2 \times \frac{1}{2} \tag{3.4.16}$$

Ideally, the value should equal to zero. In the test for smoothness model above, t stands for time, and we use half-year intervals because the government of Kenyas treasury bonds pay half yearly coupons. The model with the least Z value is deemed to be the best.

3.5 EMPIRICAL RESULTS OF PARAMETRIC MODELS

3.5.1 Data

In Kenya, nearly all bond transactions take place on the OTC market. The data used was supplied by the Central Bank of Kenya. The sample period contains 417 weekly data from January 2005 to December 2012. Weekly prices (every Friday) for 108 Kenyan Government Bonds (KGBs) with original maturity dates ranged from 2 to 30 years are obtained.

3.5.2 Parameter Estimation

3.5.2.1 Nelson-Siegel (1987) Model

Table 3.1 lists the summary statistics of estimated parameters for the (Nelson, C. R., & Siegel, A. F. , 1987) model. It is seen that all estimated values for $\hat{\beta_1}$ and $\hat{\beta_2}$ are negative, which indicates that the yield curves generated by this model are all positively and upward sloping without a visible hump.

Table 3.1: Results for estimated parameters for (Nelson-Siegel, 1987) model

Year	Parameters			
	$\hat{\beta_0}$	$\hat{\beta_1}$	$\hat{\beta_2}$	$\hat{\tau_1}$
2005	0.0587	-0.0115	-0.0040	4.6089
2006	0.0463	-0.0090	-0.0127	3.2148
2007	0.0454	-0.0133	-0.0388	1.8327
2008	0.0356	-0.0183	-0.0425	2.1674
2009	0.0358	-0.0164	-0.0493	1.0232
2010	0.0225	-0.0085	-0.0819	0.6237
2011	0.0225	-0.0085	-0.0819	0.6237
2012	0.0241	-0.091	-0.0213	1.0595

3.5.2.2 Svensson (1994) Model

Table 3.2 reports the summary statistics of estimated parameters for the (Svensson, L. E. , 1994) model. We find that, in the particular years 2007, 2010, 2011 and 2012, the estimated $\hat{\beta_1}$ is negative while the estimated $\hat{\beta_2}$ is positive, showing the yield curves would have a positively upward sloping combined with a slightly humped shape.

Table 3.2: Results for the estimated parameters for (Svensson, 1994) model

Year	Parameters				
	$\hat{\beta_0}$	$\hat{\beta_1}$	$\hat{\beta_2}$	$\hat{\tau_1}$	$\hat{\tau_2}$
2005	0.0590	-0.0138	-0.0010	2.1459	3.3983
2006	0.0473	-0.0121	-0.0068	2.5956	3.8261
2007	0.0425	-0.0182	0.0020	3.3964	5.4206
2008	0.0370	-0.0189	-0.0333	2.5649	8.1488
2009	0.0354	-0.0181	-0.0309	2.8238	7.6419
2010	0.0248	-0.0129	0.0025	8.0086	15.4666
2011	0.0248	-0.0129	0.0025	8.0086	15.4666
2012	0.0250	-0.0114	0.0072	3.3570	4.1933

3.5.2.3 Rezende-Ferreira (2011) Model

Table 3.3 lists the summary statistics of estimated parameters for the (Rezende, R. B. , 2011) model. From the years 2005, 2007 and 2008 the estimated $\hat{\beta_1}$ is negative while estimated $\hat{\beta_2}$ is positive, showing the yield curves have a positively upward sloping combined with a slightly humped shape. And both the estimated parameters $\hat{\beta_1}$ and $\hat{\beta_2}$ are both negative in the years 2006, 2009 to 2012 showing that the yield curves are positively upward sloping.

Table 3.3: Results for the estimated parameters for (Rezende-Ferreira, 2011) model

Year	Parameters					
	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\tau}_1$	$\hat{\tau}_2$
2005	0.0588	-0.0125	0.0316	-0.0365	2.4513	2.3036
2006	0.0477	-0.0152	-0.0034	0.0010	3.4306	3.4287
2007	0.0400	-0.0284	0.0701	-0.0410	3.8971	3.6562
2008	0.0355	-0.0293	0.0070	-0.0189	2.3245	2.1520
2009	0.0359	-0.0308	-0.0068	0.0115	3.2297	2.9008
2010	0.0232	-0.0022	-0.0074	-0.0110	1.1180	1.1183
2011	0.0232	-0.0022	-0.0074	-0.0110	1.1180	1.1183
2012	0.0279	-0.0061	-0.0074	-0.0055	3.7633	3.7498

3.5.2.4 Comparison of Fitting Performance in Terms of Accuracy

A direct comparison of the three models in Table 3.4 appears to favor the (Rezende, R. B. , 2011) yield curve. The (Nelson, C. R., & Siegel, A. F. , 1987) shows the worst fitting performance among the models. Hence, we conclude that in the illiquid bond market, based on a family of Nelson-Siegel yield curve models, it does help to improve the flexibility of the yield curve if we add extra parameters in the parsimonious yield curve model.

Table 3.4: Summary statistics for fitting performance in terms of accuracy

	RMSPE			RMSE			R^2		
	Nelson	Svenson	Rezende	Nelson	Svenson	Rezende	Nelson	Svenson	Rezende
Mean	0.0144	0.0131	0.0122	1.6318	1.4914	1.4043	0.9654	0.9693	0.9738
Std. Dev	0.0066	0.0061	0.0051	0.7752	0.7299	0.6033	0.0357	0.0368	0.0299
Max	0.0413	0.0385	0.0281	4.6311	4.5345	3.3047	0.9973	0.9976	0.9979
Min	0.0050	0.0048	0.0041	0.5311	0.5109	0.4363	0.8015	0.7671	0.8219

3.5.2.5 Comparison for Fitting Performance in Terms of Smoothness

In the academic literature, it has been observed that when comparing alternative methods of term structure fitting models, there is usually a trade-off between flexibility and smoothness. In Table 3.5, the (Rezende, R. B. , 2011) seems to have the best fit in flexibility for fitting the term structure of KGB market. However, as shown in table 3.5 below, the (Nelson, C. R., & Siegel, A. F. , 1987) Model is superior to its counterparts, the (Svensson, L. E. , 1994) Model and the (Rezende, R. B. , 2011) Model, which shows that the (Nelson, C. R., & Siegel, A. F. , 1987) results to a relatively smoother yield curve, compared to the other two models.

Table 3.5: Summary statistics for fitting performance in terms of smoothness

		With liquidity constraint
	R^2	$Smoothness : (Z) \times (10^{-6})$
Nelson-Siegel (1987)	0.9654	4.9822
Svensson (1992)	0.9693	6.3840
Rezende-Ferreira (2011)	0.9738	10.7467

The possible explanation for the observed results in smoothness test is the over-parametization in (Svensson, L. E. , 1994) and (Rezende, R. B. , 2011) models, making them less smooth compared to the (Nelson, C. R., & Siegel, A. F. , 1987) model.

In addition to the tests indicated above, we decided to use two additional test tools to check the adequacy of the (Nelson, C. R., & Siegel, A. F. , 1987) Model. These tools are a) the monotonicity of the discount factors curve and b) the comparison between the observed bonds' dirty prices and the model's dirty prices. The following were the results achieved:

Figure 3.2: NS Discount Graph: Showing That the Discount Function is A Decreasing Function of Time

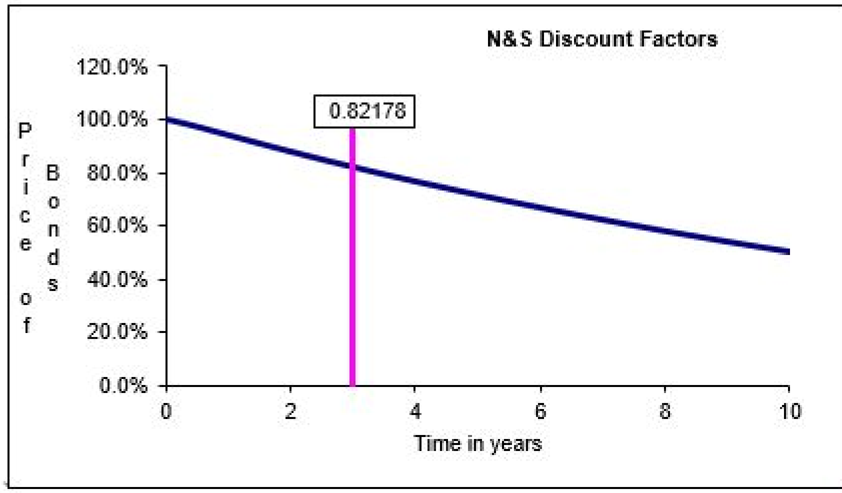
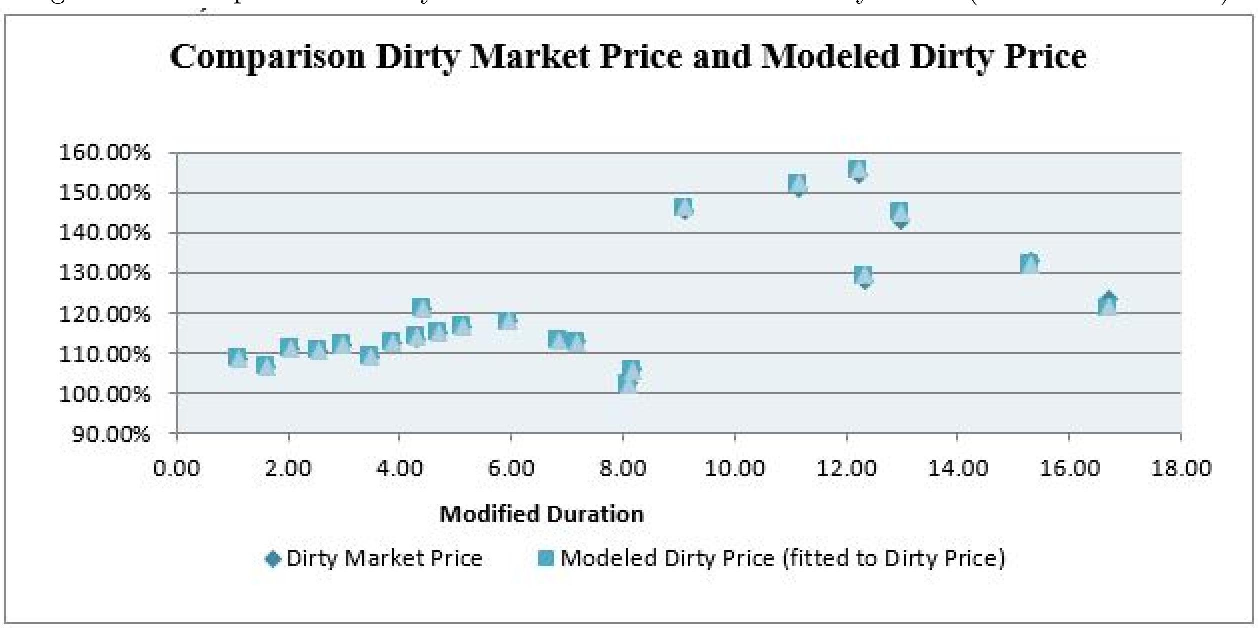


Figure 3.3: Comparison of Dirty Market Price and Modelled Dirty Prices (Data Source: CBK)



We see that the (Nelson, C. R., & Siegel, A. F. , 1987) Model produces a monotonic discount curve; i.e. it produces a decreasing discount curve which points towards the monotonicity of the curve. On the other hand, the fitted price almost perfectly match the observed market prices, again indicating the adequacy of the (Nelson, C. R., & Siegel, A. F. , 1987) Model.

3.6 DISCUSSION OF RESULTS AND CONCLUSION

After investigating the three parametric models of Nelson-Siegel class, (Nelson, C. R., & Siegel, A. F. , 1987) Model gave the best performance in terms of smoothness of the forward curve. However, it performed poorer in terms of accuracy in pricing back the bonds, compared to the other two models.

This ZCYC resulting from this study will be used mostly by the Capital Markets Authority of Kenya to guide on pricing of derivatives in the NSE. The individual investment banks and brokerage firms will load the rate of return according to their policies. Therefore, we will choose smoothness over accuracy in pricing back. In support of this decision, the Bank of International Settlements (2005), which is a technical report on how central banks around the world calculate the ZCYC, reports that central banks around the world do not typically require yield curve models that price back all the inputs exactly, when determining monetary policy. A curve does not have to have 100% accuracy; 95% and above is deemed as adequate. We see that (Nelson, C. R., & Siegel, A. F. , 1987) Model has 96.54% accuracy level, which is above the required 95%.

We, therefore, conclude that the (Nelson, C. R., & Siegel, A. F. , 1987) Model is the best parametric model for the Nairobi Securities Exchange, and in extension, the East African Securities Markets.

Chapter 4

CONSTRUCTION OF ZCYC FOR THE NSE: INTERPOLATION FUNCTION VERSUS PARAMETRIC MODELS

We seek to determine which yield curve construction method produces the best zero coupon yield curve (ZCYC) for the Nairobi Securities Exchange (NSE). The ZCYC should be differentiable at all points and at the same time, should produce a continuous and positive forward curve at all knot points. A decreasing discount curve is also expected from the resulting ZCYC, as an indication of monotonicity. For the interpolation method, we will use an improvement of monotone preserving interpolation method on r(t)t, while the (Nelson, C. R., & Siegel, A. F. , 1987) Model is the parametric model of choice. This is because compared to other interpolation methods, the improvement of monotone preserving interpolation method on r(t)t produces curves with the desirable trait of differentiability at all points, while the (Nelson, C. R., & Siegel, A. F. , 1987) model is shown to produce the best-fit results for Kenyan bond data. We compare the models' performance regarding accuracy in pricing back the fixed-income securities. For this study, we use bond data from Central Bank of Kenya (CBK). The better of the two methods will be used for the Kenyan securities market.

4.1 LITERATURE SURVEY

Many estimation methods for yield curves have appeared in literature over the years. Generally speaking, there are two distinct approaches to estimate the term structure of interest rates: the equilibrium model and the statistical techniques.

The first approach is formalized by defining state variables characterizing the state of the economy (relevant to the determination of the term structure) which are driven random processes and are related in some way to the prices of the bonds. It then uses no-arbitrage arguments to infer the dynamics of the term structure. Examples of this approach include: (Brennan, M. J., & Schwartz, E. S. , 1979; Cox, J. C., Ingersoll Jr, J. E., & Ross, S. A. , 1985; Dothan, L. U. , 1978; Duffie, D., & Kan, R. , 1996; Vasicek, O. , 1977).

Unfortunately, regarding the expedient assumptions about the nature of the random process driving the interest rates, the yield derived by those models have a specific functional form dependent only on a few parameters, and usually, the observed yield curves exhibit more varied shapes than those justified by the equilibrium models.

In contrast to the equilibrium models, statistical techniques focusing on obtaining a continuing yield curve from cross-sectional coupon bond data based on curve fitting techniques can describe a richer variety of yield patterns in reality. The resulting term structure estimated from the statistical techniques can be directly put into the interest rate models such as the (Heath, D., Jarrow, R., & Morton, A. , 1992; Hull, J., & White, A. , 1990) models, for pricing interest rate contingent claims. Since a coupon bond can be considered as a portfolio of discount bonds with maturities dates consistent with the coupon dates, the discount bond prices can thus be extracted from the actual coupon bond prices by statistical techniques.1 These techniques can be broadly divided into two categories: the splines (interpolation methods) and the parsimonious function forms; see (Alper, C. E., Akdemir, A., & Kazimov, K. , 2004).

Interpolation is a method of constructing new data points within the range of a discrete set of known data points (called knot points). The simplest method for interpolating between two points is by connecting them through a straight line. However, all the variations of linear interpolation imply discontinuities in the forward rate curve. To produce continuous forward rates curves, researchers introduced cubic methods of interpolation. Unfortunately, all the traditional cubic methods are incapable of ensuring strictly positive forward rates, which are synonymous with non-decreasing discount factors, as shown by (Hagan, P. S., & West, G. , 2006). Furthermore, some cubic methods have an inherent lack of locality in the sense that a local perturbation of curve input data will cause ringing and cause changes in the data far away from the perturbed data point as shown by (Anderson, L. , 2007). To counter this, a monotone convex interpolation method was developed, which it is claimed to be capable of ensuring a positive and (mostly) continuous forward rate curve (Hagan, P. S., & West, G. , 2006). Unfortunately, the model depends heavily on an appropriate interpolation algorithm. Also, it was discovered by (Du Preez, P. F. , 2011) that there were specific conditions under which the interpolation function of the monotone convex interpolation would produce discontinuity of f(t). This led to the monotone preserving r(t)t method of interpolation, introduced by (Du Preez, P. F. , 2011). Unfortunately, monotone preserving method had the undesirable characteristic of not being differentiable at the knot-points. (Muthoni, L., Onyango, S., & Ongati, N. O. , 2015a) Introduced a new method of interpolation, which is an improvement of monotone preserving r(t)t interpolation method suggested by (Du Preez, P. F. , 2011). This was done by removing the non-differentiability at the knot points in the method above.

Parsimonious models specify parsimonious parameterizations of the discount function, spot rate or the implied forward rate. Moving from the cubic splines, (Chambers, D. R., Carleton, W. T., & Waldman, D. W. , 1984) introduced the parsimonious function forms by considering an exponential polynomial to model the discount function (Nelson, C. R., & Siegel, A. F. , 1987) followed shortly after that by choosing an exponential function with four unknown parameters to model the forward rate of U.S Treasury bills. By considering the three components that make up this function,(Nelson, C. R., & Siegel, A. F. , 1987) illustrated that it could be used to generate a variety of shapes for the forward rate curves and analytically solve for the spot rate. Moreover, the advantage of the classical (Nelson, C. R., & Siegel, A. F. , 1987) model is that the three parameters may be interpreted as latent level, slope and curvatures factors. (Diebold, F. X., & Li, C. , 2006; Modena, M. , 2008; Tam, C.-S., & Yu, I.-W. , 2008) employed the (Nelson, C. R., & Siegel, A. F. , 1987) interpolant to examine bond pricing with a dynamic latent factor approach and concluded that it was satisfactory.

(Muthoni, L., Onyango, S., & Ongati, N. O. , 2015b) Estimated the Kenyan government bonds (KGBs) term structure of interest rates based on the parsimonious functions specifications, i.e. the four parameters (Nelson, C. R., & Siegel, A. F. , 1987) model, the five parameters (Svensson, L. E. , 1994) model and the six parameters (Rezende, R. B. , 2011) model, known as Nelson-Siegel-Svensson model. The reason they chose the Nelson-Siegel family is that these models have the substantial flexibility required to match the changing shape of the yield curve, yet they only use a few parameters. As noted by (Diebold, F. X., & Li, C. , 2006), they can be used to predict the future level, slope, and curvature factors for bond portfolio investments purposes. After comparing the Nelson-Siegel classes of models, (Muthoni, L., Onyango, S., & Ongati, N. O. , 2015b) found (Nelson, C. R., & Siegel, A. F. , 1987) to be the superior model.

4.2 EMPIRICAL METHODOLOGY

4.2.1 The Nelson-Siegel (1987) Model

The (Nelson, C. R., & Siegel, A. F. , 1987) model sets the instantaneous forward rate at maturity m given by the solution to a second order differential equation with unequal roots as follows:

$$f(m) = \beta_0 + \beta_1 exp\left(\frac{-m}{\tau_1}\right) + \beta_2 \frac{m}{\tau_1} exp\left(\frac{-m}{\tau_1}\right) \tag{4.2.1}$$

where $m > 0$, is the time to maturity of a given bond. Equation (4.2.1) consists of three parts: A constant, an exponential decay functional and Laguerre function. β_0 is independent of m and as much, β_0 is often interpreted as the level of long term interest rates. The exponential decay function approaches zero as m tends to infinity and β_1 as m tends to zero. The effect of β_1 is thus only felt at the short end of the curve. The Laguerre function on the other hand approaches zero as m tends to infinity, and as m tends to zero. The effect of β_2 is thus only felt in the middle section of the curve, which implies that β_2 adds a hump to the yield curve.

The spot rate functions under the model of (Nelson, C. R., & Siegel, A. F. , 1987) is as follows:

$$r(m) = \beta_0 + \beta_1 \left(\frac{\tau_1}{m}\right) \left[1 - exp\left(\frac{-m}{\tau_1}\right)\right] + \beta_2 \left(\frac{\tau_2}{m}\right) \left[1 - exp\left(\frac{-m}{\tau_2}\right)\right] - \beta_2 exp\left(\frac{-m}{\tau_2}\right) \tag{4.2.2}$$

$$r(m) = \beta_0 + \beta_1 \left(\frac{\tau_1}{m}\right) \left[1 - exp\left(\frac{-m}{\tau_1}\right)\right] + \beta_2 \left(\frac{\tau_2}{m}\right) \left[1 - exp\left(\frac{-m}{\tau_2}\right)\left(\frac{m}{\tau_2} + 1\right)\right]$$

From equation (4.2.2) it follows that both the spot and forward rate function reduce to $\beta_0 + \beta_1$ as $m \to 0$. Furthermore, we have $\begin{matrix} lim \\ m \to 0 \end{matrix} \; r(m) = \begin{matrix} lim \\ m \to 0 \end{matrix} \; f(m) = \beta_0$. Thus, in the absence of arbitrage, we must have that $\beta_0 > 0$ and $\beta_0 + \beta_1 > 0$.

4.2.2 Improvement of Monotone Convex Interpolation On r(t)t

We start with a mesh of data points $\{t_1, t_2, \ldots, t_n\}$ (we will think of the x-values as time points on the x axis) and the corresponding y values are define as $\{f_1, f_2, \ldots, f_n\}$ for a generic but unknown function $f(t)$. Cubic splines are generally defined by piece-wise cubic polynomial that passes through consecutive points:

$$f(t) = a_i + b_i(t - t_i) + c_i(t - t_i)^2 + d_i(t - t_i)^3 \tag{4.2.3}$$

We use equation (4.2.3) to compute the derivatives (which in our case we will use to construct the forward curve):

$$\frac{\partial f(t)}{\partial f_j} = \frac{\partial a_i}{\partial f_j} = \frac{\partial b_i}{\partial f_j}(t - t_i) + \frac{\partial c_i}{\partial f_j}(t - t_i)^2 + \frac{\partial d_i}{\partial f_j}(t - t_i)^3 \tag{4.2.4}$$

We have:

$$\frac{\partial a_i}{\partial f_j} = \delta_i^j \tag{4.2.5}$$

$$\frac{\partial m_i}{\partial f_j} = \frac{1}{h_i}\delta_{i+1}^j - \frac{1}{h_i}\delta_i^j \tag{4.2.6}$$

$$\frac{\partial c_i}{\partial f_j} = \frac{1}{h_i}\left(3\frac{\partial m_i}{\partial f_j} - \frac{\partial b_{i+1}}{\partial f_j} - 2\frac{\partial m_i}{\partial f_j}\right) \tag{4.2.7}$$

$$\frac{\partial d_i}{\partial f_j}\frac{1}{h_i^2}\left(\frac{\partial b_{i+1}}{\partial f_j}-\frac{\partial b_i}{\partial f_j}-2\frac{\partial m_i}{\partial f_j}\right) \tag{4.2.8}$$

$$\frac{\partial b_i}{\partial f_j}=\begin{cases}\frac{\left[(max(m_{i-1},m_i)+2min(m_{i-1},m_i))\star\frac{m_{i-1}}{h_i}\left(\delta_{i+1}^j-\delta_i^j\right)+\frac{m_i}{h_{i-1}}\left(\delta_i^j-\delta_{i-1}^j\right)\right]-\left[(m_{i-1}*m_i)\star\frac{m_i}{h_{i-1}}\left(\delta_i^j-\delta_{i-1}^j\right)+2\frac{m_{i-1}}{h_i}\left(\delta_{i+1}^j-\delta_i^j\right)\right]}{(max(m_{i-1},m_i)+2min(m_{i-1},m_i))^2}\\ \frac{\left[(max(m_{i-1},m_i)+2min(m_{i-1},m_i))\star\frac{m_i}{h_i}\left(\delta_i^j-\delta_{i-1}^j\right)+\frac{m_{i-1}}{h_{i-1}}\left(\delta_{i+1}^j-\delta_i-^j\right)\right]-\left[(m_{i-1}*m_i)\star\frac{m_{i-1}}{h_{i-1}}\left(\delta_{i+1}^j-\delta_i^j\right)+2\frac{m_i}{h_i}\left(\delta_i^j-\delta_{i-1}^j\right)\right]}{(max(m_{i-1},m_i)+2min(m_{i-1},m_i))^2}\end{cases} \tag{4.2.9}$$

The curves generated are used to price bonds and the results compared with observed market prices.

4.3 EMPIRICAL RESULTS

4.3.1 Parameter Estimation: Nelson-Siegel (1987) Model

Table 4.1 lists the summary statistics of estimated parameters for the (Nelson, C. R., & Siegel, A. F. , 1987) model. It is seen that all estimated values for $\hat{\beta_1}$ and $\hat{\beta_2}$ are negative, which indicates that the yield curves generated by this model are all positively and upward sloping without a visible hump.

Table 4.1: Results for estimated parameters for Nelson-Siegel (1987) model

Year	Parameters			
	$\hat{\beta_0}$	$\hat{\beta_1}$	$\hat{\beta_2}$	$\hat{\tau_1}$
2005	0.0587	-0.0115	-0.0040	4.6089
2006	0.0463	-0.0090	-0.0127	3.2148
2007	0.0454	-0.0133	-0.0388	1.8327
2008	0.0356	-0.0183	-0.0425	2.1674
2009	0.0358	-0.0164	-0.0493	1.0232
2010	0.0225	-0.0085	-0.0819	0.6237
2011	0.0225	-0.0085	-0.0819	0.6237
2012	0.0241	-0.091	-0.0213	1.0595

4.3.2 Parameter Estimation: Thesis Interpolation Model

Table 4.2 lists the summary statistics of estimated parameters for the improved monotone preserving interpolation method on $r(t)t$ Model.

Table 4.2: Results for Estimated Parameters for Thesis Interpolation Model

Year	Parameters			
	a_i	b_i	c_i	d_i
2005	0.0588	-0.0125	0.0316	-0.0365
2006	0.0477	-0.0152	-0.0034	0.0010
2007	0.0400	-0.0284	0.0701	-0.0410
2008	0.0355	-0.0293	0.0070	-0.0189
2009	0.0359	-0.0308	-0.0068	0.0115
2010	0.0232	-0.0022	-0.0074	-0.0110
2011	0.0232	-0.0022	-0.0074	-0.0110
2012	0.0279	-0.0061	-0.0074	-0.0055

4.3.3 Comparison of the Models

4.3.3.1 In terms of Accuracy

A direct comparison of the two models in Table 4.3 appears to favor the (Nelson, C. R., & Siegel, A. F. , 1987) if we consider accuracy and flexibility, but the improved spline method, also referred to as the Thesis Model, has a higher coefficient of determination.

Table 4.3: Summary statistics for fitting performance in terms of accuracy

	RMSPE		RMSE		R_2	
	Nelson-Siegel	SplineMethod	Nelson	Spline-Method	NelsonSiegel	SplineMethod
mean	0.0144	0.0101	1.6318	1.115	0.9654	0.9493
Std. dev.	0.0066	0.0051	0.7752	0.6033	0.0357	0.0299
Max	0.0413	0.0281	4.6311	3.3047	0.9973	0.9979
Min	0.0050	0.0041	0.5311	0.4363	0.8015	0.8219

4.3.3.2 In Terms of Smoothness

When comparing alternative methods of term structure fitting models, there is usually a trade-off between flexibility and smoothness. In Table 4.3, the spline seems to have the best fit in flexibility for fitting the term structure of KGB market (in terms of RMSE and RMSPE). However, as shown in table 4.4 below, the (Nelson, C. R., & Siegel, A. F. , 1987) Model is superior to the interpolation method, where it results to a relatively smoother yield curve.

Table 4.4: Summary statistics for fitting performance in terms of smoothness

		With liquidity constraint
	R^2	$Smoothness:(Z)(\times10^{-}6)$
Nelson-Siegel (1987)	0.9654	4.9822
Spline (interpolation)	0.9493	7.8423

4.4 DISCUSSION OF RESULTS AND CONCLUSION

In this chapter, we set out to compare the performance of the thesis interpolation function and the (Nelson, C. R., & Siegel, A. F. , 1987) model when applied on Kenyan data. We find that In terms of accuracy, the spline method does poorer than the parametric method in that it gives an accuracy of 94.93% compared to an accuracy of 96.54% by the Nelson-Siegel model. When it comes to the test of smoothness, Nelson-Siegel model also performs better with a lower Z value of 4.9822 compared to the splines value of 7.8423.

After considering all the factors at hand, we come to a conclusion that the parametric model (Nelson-Siegel) is the slightly better than the model suggested by this thesis. However, we would like to point out that the decision on whether to use parametric models or spline-based methods to generate the yield curve depends on the intended use of the yield curve. Parametric methods are very popular amongst Central Banks. This is because Central Banks, when determining monetary policy, typically do not require yield curves that price-back all inputs exactly and they find parametric methods easier to implement as compared to the spline models. On the other hand, Investment Bankers usually prefer spline-based models because they can price-back all financial products (inputs) as compared to the parametric models. The other factor is that parametric models might not completely reflect the information on the ground as compared to the spline models. For this reason, we will use the interpolation method suggested in this thesis for the purpose of pricing the coffee futures, which has been shown to perform better than existing interpolation methods.

Chapter 5

PRICING OF FUTURES IN A MARKET WITH INCOMPLETE INFORMATION

The objective of this chapter is to advise on the pricing model to be used in pricing coffee futures to be traded on the Nairobi Securities Exchange (NSE) by the end of the year 2018, according to a report given by the Capital Market Authority of Kenya in 2013. We apply Belallahs three-factor model with modifications. The factors we consider are the rate of interest which is assumed to be mean reverting, the convenience yield which is an adjustment to the pricing formula to reflect constraint in a market, in this case, the cost of information, and the spot price. We include a liquidity constraint in our pricing which reflects the illiquidity of the test market (Ivory Coast) and target market, the NSE. The calibration method used in this study is the L-BGFS-B model which reduces the number of iterations to be undertaken and also the attractive property of having boundary conditions. We use findings by (Onyango, S. N., & Ingleby, M. , 2006) on Ivorian Coffee Futures prices data, the guide country, to calculate grey-scale votes in each cell within the accumulator array used to store process knowledge in the Hough formalism. This enables us to retrieve the drift and volatility of the price data. We then compare the models prices with the observed market prices using correlation with time lags and Fretchet distance to assess the suitability of the model to be used in pricing the futures.

5.1 THE GENERAL DERIVATIVE PRICING

In this section, we develop a general derivative pricing formula. In general, the underlying asset S might not be treated. Interest rates, for example, are not traded assets while stocks and bonds are. Suppose S evolve according to the It$\hat{o}$ process given by

$$\frac{dS}{S} = \mu dt + \sigma dZ \tag{5.1.1}$$

where μ and σ may depend on S and t. Let $f_1(S,t)$ and $f_2(S,t)$ be prices of two derivatives with dynamics:

$$\frac{df_i}{f_i} = \mu_i dt + \sigma_i dZ, \qquad i = 1, 2 \tag{5.1.2}$$

For simplicity we assume that the derivatives share the same Wiener process, Z, as S. A portfolio consisting of $\sigma_1 f_1$units of the first derivative and $\sigma_2 f_2$ units of the second derivative is instantaneously riskless because

$$\sigma_2 f_2 df_1 - \sigma_1 f_1 df_2 = \sigma_2 f_2 f_1(\mu_1 dt + \sigma_1 dZ) - \sigma_1 f_1 f_2(\mu_2 dt + \sigma_2 dZ)$$

$$= (\sigma_2 f_2 f_1 \mu_1 - \sigma_1 f_1 f_2 \mu_2)dt \tag{5.1.3}$$

Equation (5.1.3) does not contain the Wiener process Z. Therefore,

$$(\sigma_2 f_2 f_1 \mu_1 - \sigma_1 f_1 f_2 \mu_2)dt = r(\sigma_2 f_2 f_1 - \sigma_1 f_1 f_2)dt \tag{5.1.4}$$

or by simplifying we have

$$\frac{\mu_1 - r}{\sigma_1} = \frac{\mu_2 - r}{\sigma_2} \equiv \lambda \quad \text{for some} \quad \lambda \tag{5.1.5}$$

Where r is the risk-less interest rate. The constant λ is referred to as the market price of risk, which is independent of the specifics of the derivative. Generally, for any derivative whose value depends on S and t and evolves according to the It$\hat{o}$ process

$$\frac{df}{f} = pdt + qdZ \tag{5.1.6}$$

must thus satisfy

$$\frac{p - r}{q} = \lambda, \text{or equivalently}, p = r + \lambda q \tag{5.1.7}$$

Using It$\hat{o}$ lemma on equation (5.1.2) and putting $f = f(S, t)$ be a derivative that depends on S and t, we have

$$p = \frac{1}{f}\left(\frac{\partial f}{\partial t} + \mu S \frac{\partial f}{\partial S} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 f}{\partial S^2}\right), \qquad q = \frac{\sigma}{f}\frac{\partial f}{\partial S} \tag{5.1.8}$$

If we substitute equation (5.1.8) in equation (5.1.7), we get

$$\frac{\partial f}{\partial t} + (\mu - \lambda\sigma)S\frac{\partial f}{\partial S} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 f}{\partial S^2} - rf = 0 \tag{5.1.9}$$

In general, suppose that assets $S_1, S_2, \ldots, S_N$ pay no dividends and evolve according to the It$\hat{o}$ process $\frac{dS_i}{S_i} = \mu_i dt + \sigma_i dZ$. Let ρ_{ij} be the correlation between dZ_i and dZ_j and $f = f(S_1, S_2, \ldots, S_N, t)$ be a derivative that depends on S_i and t. Then equation (5.12.9) becomes

$$\frac{\partial f}{\partial t} + \sum_{i=1}^{N}(\mu_i - \lambda_i\sigma_i)S_i\frac{\partial f_i}{\partial S_i} + \frac{1}{2}\sum_{i=1}^{N}\rho_{ij}\sigma_i\sigma_j S_i S_j \frac{\partial^2 f}{\partial S_i \partial S_j} - rf = 0 \tag{5.1.10}$$

We refer, for example, to ((Wilmott, P. , 1998), Ch.23; (Hull, J., & White, A. , 1990), Ch.19; (Lyuu, Y.-D. , 2001), Ch.15) for further analysis.

5.2 ESTIMATION OF DRIFT AND VOLATILITY IN PRICING FORMULAS

Volatility has been used in financial markets in the assessment of risk. Another measure of risk that is now popular amongst investment managers is the Value-at-Risk (VaR). This measures the risk that a portfolio of assets reaches a market value below a contracted target: the amount below target is the value in VaR and is calculated using a probability distribution for future prices of assets in the portfolio.

Traditionally, volatility of an underlying asset as used in Black-Scholes world is assumed to be constant throughout the duration of a derivative contract. In such a case the volatility is estimated by using historical data in the form of logarithms of the asset returns. This measure of volatility is usually referred to as historical volatility. It is unconditional: that is, it does not recognize that there are interesting patterns in asset volatility that develop in response to trading conditions. A large number of models have been proposed to address some of the shortcomings of the classical Black-Scholes model (Black, F., & Scholes, M. , 1973; Merton, R. C. , 1973); and for traded options, the response of volatility to the trading of the underlying asset has been widely explored.

Recent research provides abundant evidence that implied volatility might differ considerably from historical volatility and contains information about subsequently released volatility ((Xu, X., & Taylor, S. J. , 1995; Fleming, J. , 1998; Blair, B. J., Poon, S.-H., & Taylor, S. J. , 2010; Malz, A. M. , 2000); and others). In other more theoretical studies, volatility is assumed to be time-varying but may be taken to vary deterministically or stochastically. The case of stochastic volatilities has been investigated by, among others ((Hull, J., & White, A. , 1987; Melino, A., & Turnbull, S. M. , 1990; Heston, S. L. , 1993)), jump models of volatility by (Bates, D. S. , 1996) and Autoregressive Conditional Heteroskedasticity (ARCH) / Generalised Autoregressive Conditional Heteroskedasticity (GARCH) family models by ((Engle, R. F. , 1982; Bollerslev, T. , 1986; Duan, J.-C. , 1995; Heston, S. L., & Nandi, S. , 2000)). The latter approach to the variation of volatility has become the most dominant one, and indeed was the reason for Engle's share in the Nobel Prize for Economics in 2003.

5.2.1 General stochastic volatility models

These models assume that the asset prices evolve according to the geometric Brownian motion:

$$\frac{dS}{S} = \mu dt + \sigma dZ_1 \tag{5.2.1}$$

In which the volatility of the underlying asset evolves according to the It$\hat{o}$ process given as:

$$d\sigma = p(S, \sigma, t)dt + q(S, \sigma, t)dZ_2 \tag{5.2.2}$$

Where increments dZ_1 and dZ_2 are unit Wiener processes $\sigma_1 = \sigma_2 = 1$, $Z_1 \sim N(0, \sqrt{t})$, $Z_2 \sim N(0, \sqrt{t})$. The correlation of these processes remains an unknown a parameter ρ, implicit in the theory, to be fitted to data in practice.

Let the value of the option with stochastic volatility be given as $V(S, \sigma, t)$, i.e. V is a function of three variables. It should be noted that although volatility is not a traded asset, one can hedge an option with two other contracts, one being the underlying traded asset, and the other the volatility risk. To illustrate this, we consider a portfolio which contains one option of value $C(S, \sigma, t)$, a quantity- Δ of the underlying asset and quantity-Δ_1 of another shadow option whose value is denoted as $C_1(S, \sigma, t)$. Here Δ is taken as a coefficient and not as a finite difference operator. The hedge portfolio has value

$$\pi = C(S, \sigma, t) - S\Delta - \Delta_1 C_1(S, \sigma, t) \tag{5.2.3}$$

The change of value of the portfolio from time t to $t + dt$ is given as

$$d\pi = dC(S, \sigma, t) - dS\Delta - \Delta_1 dC_1(S, \sigma, t) \tag{5.2.4}$$

Thus by It$\hat{o}$'s lemma we have

$$\begin{aligned} d\pi = &\left(\frac{\partial C}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 C}{\partial S^2} + \rho\sigma q S \frac{\partial^2 C}{\partial S \partial \sigma} + \frac{1}{2}q^2 \frac{\partial^2 C}{\partial \sigma^2}\right) dt - \\ &\Delta_1 \left(\frac{\partial C_1}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 C_1}{\partial S^2} + \rho\sigma q S \frac{\partial^2 C_1}{\partial S \partial \sigma} + \frac{1}{2}q^2 \frac{\partial^2 C_1}{\partial \sigma^2}\right) dt + \\ &\left(\frac{\partial C}{\partial S} - \Delta_1 \frac{\partial C_1}{\partial S} - \Delta\right) dS + \left(\frac{\partial C}{\partial \sigma} - \Delta_1 \frac{\partial C_1}{\partial \sigma}\right) d\sigma \end{aligned} \tag{5.2.5}$$

To eliminate randomness in (5.2.5), we choose

$$\frac{\partial C}{\partial S} - \Delta_1 \frac{\partial C_1}{\partial S} - \Delta = 0 \qquad \text{and} \qquad \frac{\partial C}{\partial \sigma} - \Delta_1 \frac{\partial C_1}{\partial \sigma} = 0 \tag{5.2.6}$$

The result is then given as:

$$\begin{aligned} d\pi = &\left(\frac{\partial C}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 C}{\partial S^2} + \rho\sigma q S \frac{\partial^2 C}{\partial S \partial \sigma} + \frac{1}{2}q^2 \frac{\partial^2 C}{\partial \sigma^2}\right) dt - \\ &\Delta_1 \left(\frac{\partial C_1}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 C_1}{\partial S^2} + \rho\sigma q S \frac{\partial^2 C_1}{\partial S \partial \sigma} + \frac{1}{2}q^2 \frac{\partial^2 C_1}{\partial \sigma^2}\right) dt \end{aligned} \tag{5.2.7}$$

By arbitrage argument, we have the returns of the portfolio equal to the risk-free rate, i.e.

$$d\pi = r\pi dt = r(C - S\Delta - \Delta_1 C_1) \tag{5.2.8}$$

That is

$$\left(\frac{\partial C}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 C}{\partial S^2} + \rho\sigma q S \frac{\partial^2 C}{\partial S \partial \sigma} + \frac{1}{2}q^2 \frac{\partial^2 C}{\partial \sigma^2}\right) dt - \Delta_1 \left(\frac{\partial C_1}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 C_1}{\partial S^2} + \rho\sigma q S \frac{\partial^2 C_1}{\partial S \partial \sigma} + \frac{1}{2}q^2 \frac{\partial^2 C_1}{\partial \sigma^2}\right) dt = r(C - S\Delta - \Delta_1 C_1)dt \tag{5.2.9}$$

To separate variables, we collect all the C terms on one side and all the C_1 terms on the other

$$\frac{\frac{\partial C}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 C}{\partial S^2} + \rho\sigma q S \frac{\partial^2 C}{\partial S \partial \sigma} + \frac{1}{2}q^2 \frac{\partial^2 C}{\partial \sigma^2} + rS\frac{\partial C}{\partial S} - rC}{\frac{\partial C}{\partial \sigma}}$$

$$= \frac{\frac{\partial C_1}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 C_1}{\partial S^2} + \rho\sigma q S \frac{\partial^2 C_1}{\partial S \partial \sigma} + \frac{1}{2}q^2 \frac{\partial^2 C_1}{\partial \sigma^2} rS\frac{\partial C_1}{\partial S} - rC_1}{\frac{\partial C_1}{\partial \sigma}} \tag{5.2.10}$$

The left-hand side of (5.2.10) is in terms of C only and can be expressed as a function of independent variables S,σ,t ((Wilmott, P. , 1998), 300-301). Thus we have

$$\frac{\partial C}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 C}{\partial S^2} + \rho\sigma q S \frac{\partial^2 C}{\partial S \partial \sigma} + \frac{1}{2}q^2 \frac{\partial^2 C}{\partial \sigma^2} + rS\frac{\partial C}{\partial S} + (p - \lambda q)\frac{\partial C}{\partial \sigma} - rC = 0 \tag{5.2.11}$$

where the separation constant $\lambda(S, \sigma, t)$ is known as the market price of (volatility) risk. In particular, for an underlying asset, if μ is the growth rate of the tradable asset, then $(\mu - r)/\sigma$ is the excess rate of return (above the risk-free rate) per unit risk- thus it is known as market price of risk and is also referred to as Shapiro ratio (see (Lyuu, Y.-D. , 2001), 220; (Hull, J., & White, A. , 1990), 498; (Wilmott, P. , 1998), 301). Under the simplifying assumption that Wiener process Z_1 and Z_2 are not correlated then (5.2.11) becomes

$$\frac{\partial C}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 C}{\partial S^2} + \frac{1}{2}q^2 \frac{\partial^2 C}{\partial \sigma^2} + rS\frac{\partial C}{\partial S} + (p - \lambda q)\frac{\partial C}{\partial \sigma} - rC = 0 \tag{5.2.12}$$

This is a partial differential equation that is analogous to Black-Scholes PDE, but accounts through λ for the shadow option price C_1.

From the perspective of pattern recognition for processes the PDEs are candidates for fitting a real price history in which the volatility risk through λ has exerted an influence on market prices, and has to be estimated by appropriate methods.

5.2.2 Using Hough Transforms to Estimate Drift and Volatility

The Hough Transform (HT) and (Duda, R. O., & Hart, P. E. , 1972; Leavers, V. F. , 1992; Toft, P. A., & Srensen, J. A. , 1996) is a standard tool for extraction of geometrical primitives such as line-segments and arcs from noisy digital images. The original Hough transform (Hough, P. V. , 1962) is commonly used to detect straight lines in edge-enhanced images. The transform was generalized by Duda and Hart (Duda, R. O., & Hart, P. E. , 1972; Leavers, V. F. , 1992; Ballard, D. H. , 1981; Illingworth, J., & Kittler, J. , 1988) to detect circles, ellipses, and even irregular shapes. Hough transforms to change the mode of presentation of data set to ease detection of a specific geometric form being sought. For example, one may want to use Hough Transforms to detect lines, circular arcs of a specific diameter or any other shape of interest. Hough transforms usually use some parametric representation to characterize the form or pattern to be detected.

The characteristic relation of the sought-for-feature is back-projected in the space of pattern parameters $(a_i, a_2, \ldots, a_n)$ a pixel (x,y) that lies on a parametric curve of the characteristic relation that is inconsistent may be represented in the parametric space, if

$$f\left((x,y),(\alpha_s)\right) = 0 \tag{5.2.13}$$

Holds; the idea of the Hough transform is to convert a pixel position (x,y) into a relation between pattern parameters α_s by fixing (x,y) in (5.2.13) above.

The behaviour of Hough transforms is mathematically predicable when a pixel (x,y) is subjected to similarity and affine transformations, and thus is especially useful in image processing, where images suffer a geometrical transformation whenever the vision system moves with respect to its sensed environment (Ser, P.-K., & Siu, W.-C. , 1995; Yip, R. K., Tam, P. K., & Leung, D. N. , 1995; Yuen, S. Y., & Ma, C. H. , 1997; Aguado, A. S., Montiel, E., & Nixon, M. S. , 2000)). The simplest Hough Transforms for line segment recognition works by transforming pixels lying in a tomographic slice through an image into votes that are replaced in cells of a histogram. The histogram accumulates that a given slice of the image contains intensity data, and is called an accumulator array.

We adapt the Hough transforms in (5.2.13) to stochastic dynamics by replacing pixel coordinates (x,y) with sequence of $\left\{log\left(\frac{S(t_i)}{S(t_0)}\right), t_i\right\}$ in a window. Quantity $S(t_i)$ is the price of security at time t_i and $S(t_0)$ is the initial price of the security at t_0. This yields (see (Onyango, S. N., & Ingleby, M. , 2006)):

$$\sigma^2 = f(\mu_j) = \frac{1}{\chi^2}(C_k + B_k\mu_j + A_k\mu_j^2) \tag{5.2.14}$$

The moments of the current returns in the historical price window with n terms, are represented by A,B,C, where

$$C_k = \sum_{i=i(k)}^{i(k)+T-1} \left((log((S(t_i))/(S(t_0)))/\sqrt{(t_i - t_{i-1})}\right)^2$$

$$B_k = -2 \sum_{i=i(k)}^{i(k)+T-1} log((S(t_i))/(S(t_0))$$

$$A_k = \sum_{i=i(k)}^{i(k)+T-1} (t_i - t_{i-1}) = t_n - t_0 \quad (5.2.15)$$

where $i(k)$is the index of the start of ith window and χ^2 is a statistical variate with the standard distribution of Fischers χ^2 with N degrees of freedom (Weatherburn, C. E. , 1949). We begin the adaptation by replacing $\aleph^2$ by its mean value n, for the sake of simplicity.

(Onyango, S. N., & Ingleby, M. , 2006) demonstrate on the effect of the trend in stock prizes. Stock bargains are response to randomly matched buyers and sellers. Unmatched buyers can raise their bid prizes to get a match, and unmatched sellers can lower their offer (asking price). An excess of buyers drives price up, while an excess of sellers drives the prize down. Raising and lowering of bids and offers are carried out randomly, so the prices of stock fluctuate randomly. Stock price movements are random and adjust rapidly to new information as it comes available. During this adjustment the price moves up and down around some trend line that reflects a current market equilibrium price.

We consider short-term stock price logarithms $[log(S(t_0)), log(S(t_1)), \ldots, log(S(t_N))]$ where $log(S(t_i))$ is assumed at times $t_i \in [0, 1, 2, \ldots, N]$. Taking short runs of window-data (window-size n, $0 < n < N$) from sample and of longer duration N, we get $f_1(\mu, \sigma^2), f_2(\mu, \sigma^2), \ldots, f_k(\mu, \sigma^2)$ k windows of n points when $nk = N$ from each historical window. These functions are converted. To A,B, and C and thence vote along σ^2 quadratic curves of the form given in (5.2.14).

In this section, we have shown that artificial intelligence can be adapted to detect stochastic processes in real market data to estimate market parameters. This technique was shown by (Onyango, S. N., & Ingleby, M. , 2006) to be robust and able to communicate complex ideas with tremendous simplicity, clarity speed, and power. It also supports effective or rapid decision making in the market. This is important because, in the market, the velocity of change is increasing amidst greater complexity and chaos, so processing a deep understanding of the market patterns can be critical to decision making whenever trading is to be undertaken.

5.3 PRICING MODELS

In this section, we define the costs as in (Merton, R. C. , 1987). For an introduction to the basic concepts for the pricing of derivative assets and real options under the uncertainty and incomplete information, we refer to (Bellalah, M., & Jacquillat, B. , 1995) and (Bellalah, M. , 1999). We use an extension of the analysis in the (Schwartz, E. , 1998, 1997) to account for the effects of incomplete information as it appears in the models of (Merton, R. C. , 1987) and (Bellalah, M. , 2001). We also use the aforementioned extension to describe the stochastic behavior of commodity prices in the presence of mean reversion and shadow costs of incomplete information.

The data used to test the models consist of weekly observations of futures prices for two commodity markets: Ethiopian Coffee Exchange (ECX) and Ivorian Coffee Exchange (ICE). In every case, ninety-nine futures contracts (i.e. N=99) were used in the estimation. For different commodities and different time periods, however, different specific futures contracts had to be used since they vary across commodities and through time for a particular commodity. The interest rate data consisted in yields on the Kenyan interest rates (Central Bank Rates). These data was used in the models requiring variable interest rates.

In this study, we assume that future coffee prices are stochastic in nature. We also assume that the interest rate s follow a stochastic process and therefore use CIR model to generate the interest rates. The risk-free part of the CIR model are ones generated by stripping the Kenyan coupon bonds and using the interpolation model developed in chapter 2 to find the risk free rates for the tenures not availed by the government. We use the Hough Transform to estimate the market drift and volatility for the futures pricing model.

The models used in this study for futures prices are closed-form solution. We use a three-factor model for the pricing of the futures; the three factors are the spot price of the commodity, the instantaneous convenience yield, and the instantaneous interest rate. When the interest rate follows a mean-reverting process as in (Cox, J. C., Ingersoll Jr, J. E., & Ross, S. A. , 1985), the joint stochastic process for the three factors under the equivalent martingale measure can be written as:

$$dS = (r - \delta)Sdt + \sigma_1 S dz_1^* \tag{5.3.1}$$

$$d\delta = \kappa(\hat{\alpha} - \delta)dt + \sigma_2 dz_2^* \tag{5.3.2}$$

$$dr_t = a(m^* r_t)dt + \sigma_3 dz_3^* \tag{5.3.3}$$

where α and m^* refer respectively to the speed of adjustment coefficient and the risk adjusted mean short rate of the interest rate process. In the context, futures prices must satisfy the following PDE.

$$\begin{aligned}
&\frac{1}{2}\sigma_1^2 S^2 F_S S + \frac{1}{2}\sigma_2^2 F_\delta \delta + \frac{1}{2}\sigma_3^2 F_r r + \\
&\qquad \sigma_1\sigma_2\rho_1 S F_S \delta + \sigma_2\sigma_3\rho_2 F_\delta r + \sigma_1\sigma_3\rho_3 S F_S r \\
&\qquad\qquad + (r-\delta)SF_S + [\kappa(\hat{\alpha}-\delta)]F_\delta + a(m^* - r)F_r - F_T = 0
\end{aligned} \tag{5.3.4}$$

Under the terminal boundary condition $F(S, \delta, r, 0) = S$;

Following the analysis in (Schwartz, E. , 1997), the solution is given by:

$$F(S, \delta, r, T) = Sexp\left[\frac{-\delta(1-e^{-\kappa T})}{\kappa} + \frac{r(1-e^{-\alpha T})}{\alpha} + C(T)\right] \tag{5.3.5}$$

This can be written in a log form as:

$$lnF(S, \delta, r, T) = lnS - \frac{\delta(1-e^{-\kappa T})}{\kappa} + \frac{r(1-e^{-\alpha T})}{\alpha} + C(T) \tag{5.3.6}$$

where

$$\begin{aligned}
C(T) = &\frac{(\kappa\hat{\alpha} + \sigma_1\sigma_2\rho_1)[(1-e^{-\kappa T}) - \kappa T]}{\kappa^2} - \frac{\sigma_2^2(4(1-e^{-\kappa T}) - (1-e^{-2\kappa T}) - 2\kappa T}{4\kappa^3} - \\
&\frac{(am^* + \sigma_1\sigma_3\rho_3)[(1-e^{-aT}) - aT]}{a^2} - \frac{\sigma_3^2(4(1-e^{-aT}) - (1-e^{-2aT}) - 2aT)}{4a^3} + \\
\sigma_2\sigma_3\rho_2 &\left[\frac{(1-e^{-\kappa T}) + (1-e^{-aT}) - (1-e^{-(\kappa+a)T})}{\kappa\alpha(\kappa+a)} + \frac{\kappa^2(1-e^{-aT}) + a^2(1-e^{-\kappa T}) - \kappa a^2 T - a\kappa^2 T}{\kappa^2 a^2(\kappa+a)}\right]
\end{aligned} \tag{5.3.7}$$

where α and m^* refer respectively to the speed of adjustment coefficient and the risk-adjusted mean short rate of the interest rate process, dz is an increment to a standard Brownian motion, measures the degree of mean reversion to the long run mean log price where $\alpha^* = \alpha - \lambda$, λ is the market price of risk and can be interpreted as market volatility; is the instantaneous convenience yield which can be seen as the cash flow of services to the holder of the commodity rather than the buyer of the futures contract, λ_S refers to an information cost for the asset S; r is the risk free rate, the sigmas the market risks associated with the given products and phi is the correlation matrix between the prices in the market.

According to both Schwartz and Bellalah, this model can also be used to price forwards and some options. The differentiating factor of these models would be the nature of the interest rate. Since in this particular model the interest rates are stochastic, then the futures prices will not equal to forward prices.

The calibration method used to estimate the parameters in the models is the L-BGFS-B model which reduces the number of iterations to be undertaken and also the attractive property of having boundary conditions. For estimation of drift and volatility, we use Hough Transform and generate the stochastic rates using CIR.

The use of stochastic rates instead of constant rates ensures that futures prices are not equal to forward prices. With the assumed risk-adjusted stochastic process for the instantaneous interest rate, the present value of a unit discount bond payable at time T when the interest rate is r is given by (see (Cox, J. C., Ingersoll Jr, J. E., & Ross, S. A. , 1985)):

$$B(r,T) = \frac{2(e^{\gamma}(T-r)-1)}{(\gamma+\kappa+\lambda)(e^{\gamma}(T-r)-1)+2\gamma} \tag{5.3.8}$$

The prices generated using this equation are compared with the prices from stripping the bonds (which in this case indicates observed market prices of discount bonds) to determine how appropriate the method is.

5.4 RESULTS AND DISCUSSION

5.4.1 Generated Spot Rates

Table 5.1: Kenyan Bond Tenures

Kenyan Bond Tenures	**B**	**Short Rate, r**	**Long Rate, R**
1	0.906005	0.110856	0.056916
2	1.646151	0.086423	0.051075
5	3.140629	0.092459	0.052518
7	3.728693	0.086499	0.051093
10	4.256073	0.054171	0.043365
15	4.64573	0.660614	0.041866
20	4.781058	0.544267	0.041397
30	4.844209	0.366483	0.040878

5.4.2 Interpolated Spot Rates

We use the interpolation method developed in the thesis to generate rates for the tenures not presented in Table 5.2 below.

Table 5.2: Short rates and long rates

Time(years)	Short Rate, r	Long Rate , R
0	0.19	0.07583581
0.25	0.165441135	0.069964941
0.5	0.1542443	0.067288305
0.75	0.138987475	0.063641117
1	0.110856292	0.056916275
1.25	0.101738221	0.054736574
1.5	0.102978684	0.05503311
1.75	0.107473232	0.056107545
2	0.0864235	0.051075545
2.25	0.087449043	0.051320704
2.5	0.079789168	0.049489588
2.75	0.0711277	0.047419039
3	0.074532756	0.048233028
3.25	0.081462153	0.049889521
3.5	0.06677602	0.046378758
3.75	0.068483269	0.04678688
4	0.073452557	0.047974803
4.25	0.077519586	0.048947038
4.5	0.094286071	0.052955115
4.75	0.091815666	0.052364558
5	0.092459676	0.05251851
5.25	0.093702882	0.052815702
5.5	0.090007541	0.05193232
5.75	0.094845644	0.053088883
6	0.096758688	0.053546202

Time(years)	Short Rate, r	Long Rate , R
6.25	0.086231132	0.051029559
6.5	0.093325978	0.052725603
6.75	0.097635338	0.053755768
7	0.086499504	0.051093714
7.25	0.08862438	0.051601672
7.5	0.086484652	0.051090163
7.75	0.069573522	0.047047509
8	0.06892234	0.046891842
8.25	0.074599898	0.048249078
8.5	0.065524853	0.046079662
8.75	0.067917434	0.04665161
9	0.053447548	0.043192547
9.25	0.055024574	0.04356954
9.5	0.056230414	0.0438578
9.75	0.056699141	0.04396985
10	0.054171291	0.04336556

5.4.3 Futures Pricing Calibrated Results

Table 5.3: Calibrated parameter values

Periods	1/15/05-5/16/12
Contracts	F1, F3, F5, F7, F9
Number of observations	347
μ	0.326 (0.0110)
k	1.156 (0.041)
α	0.248 (0.098)
σ_1	0.274 (0.012)
σ_2	0.280 (0.017)
σ_3	0.281 (0.016)
ρ_1	0.818 (0.020)
λ	0.256 (0.0114)
ρ_2	0.0621 (0.0124)

We use L-BFGS-B algorithm to calibrate the parameters above. With the parameters calculated in Table 5.3 above, we are then in a position to calculate the price of the Futures. The bracketed values are the error terms. The relatively high error margins are explained by the use of Brownian motion in pricing, which is not completely reflective of the movement of prices in the markets from which we got the data, but it is nevertheless adequate.

5.4.4 Analysis of the Coffee Futures Data

We look at different methods of analysis taking into account the non-linear nature of our data. The methods used include Correlation analysis with time lags and Frechet distance. The reason for the time lags is because we foresee a situation where the market would take longer to react to information, and thus account for the delay.

5.4.4.1 Correlation analysis with time lags

The following results of the correlation with time lags (up to lag 7) between ECX washed prices, ECX unwashed prices and ICE futures prices in comparison with the thesis futures pricing model's prices are shown in the table below. They were generated using R programming, and the codes and the output can be found in the appendices section (Appendix B).

Table 5.4: Correlation of the Coffee Futures Data with Time Lags (Up to Lag 7)

	LAGS	AUTO CORRELATION		
THESIS FUTURES PRICING MODEL		ECX WASHED	ECX UNWASHED	ICE FUTURES
	1	0.716	0.781	0.929
	2	0.579	0.737	0.860
	3	0.571	0.699	0.809
	4	0.555	0.674	0.780
	5	0.449	0.663	0.757
	6	0.370	0.614	0.735
	7	0.433	0.553	0.714

It is evident that Ivory Coffee Exchange futures prices are highly correlated with the thesis futures pricing model prices.

5.4.4.2 Frechet Distance

Frechet distance is a measure of the similarity between curves that takes into account the location and order of the points along the curves. Let S be a metric space. A curve A in S is a continuous map from the unit interval into S, i.e. $A : [0, 1] \to S$. A reparameterization α of $[0, 1]$ is a continuous, non-decreasing, surjection $\alpha : [0, 1] \to [0, 1]$. Let A and B be two given curves in S. Then, the Frechet distance between A and B is defined as the infimum over all reparameterizations α and β of $[0, 1]$ of the maximum over all $t \in [0, 1]$ of the distance in S between $A(\alpha(t))$ and $B(\beta(t))$. In mathematical notation, the Frechet distance $F(A, B)$ is defined by:

$$F(A,B) = \underset{(\alpha,\beta)}{inf} \quad \underset{(t \in [0,1])}{max} \quad \{d\left(A(\alpha(t)), B(\beta(t))\right)\} \tag{5.4.1}$$

where d is the distance function of S.

The results for Frechet distance were generated using R programming as shown in the table below.

Table 5.5: Frechet Distance for the Interest Rates Data

KENYA	FRECHET DISTANCE		
	BRAZIL	ETHIOPIA	IVORY COAST
	13.00839	4.113821	1.40443

The results above show that curves generated by ICE futures and the thesis futures pricing model have the shortest distance, which means that they are more similar to each other compared to the distance between the curves generated by the ECX data and thesis model.

5.5 CONCLUSION

In this Chapter estimated the parameter of Futures pricing model by applying findings by (Onyango, S. N., & Ingleby, M. , 2006) to calculate grey-scale votes in each cell within the accumulator array. This enables us to retrieve the drift and volatility of the price data. In conjunction with the interpolation model, we were able to generate the risk-free rates which we incorporated into the stochastically driven CIR model; the stochastic rates generated were then used in the pricing model, the parameters of which were estimated using L-BFGS-B, thereby making it possible for us to price coffee futures. (See Appendix A).

The results show that the model generated from this thesis is closely related to the observed market prices from Ivorian Futures exchange. The same conclusion is arrived at when using Frechet distance as a comparison tool, indicating that the model used can be relied upon to predict future prices of the futures traded in the exchange.

Chapter 6

CONCLUSIONS AND RECOMMENDATIONS

6.1 THESIS CONCLUSIONS

The main intention of this study was to develop a zero coupon yield curve (ZCYC) for Nairobi Securities Exchange (NSE) and an associated model for pricing coffee futures that are to be introduced in the Kenyan market. This modeling task required:

- The existence of zero coupon bonds in the market;
- Continuous flow of data from the market; and
- Historical prices of coffee futures,

The required information was neither adequate nor available in the market because only coupon bonds are traded: zero coupon bonds are not traded at all. To solve this, we stripped the coupon bonds to create hypothetical zero-bond or risk free rates. Secondly, the bonds available for trade had only specific, European-type, maturities: 1, 2, 3, 4, 5, 7, 10, 15, 20, 25 and 30 years, and there is no information on several tenures. Therefore, it was necessary for us to interpolate the data available so that we could estimate the rates for the missing maturities and construct a yield curve. Careful selection of interpolation technique had to be done, which is a significant part of this study.

In relation to interpolation, novel aspects covered in this thesis are two:

1. In the literature, there was a general weakness in the interpolation methods in lacking differentiability, especially at the knot points. In this study, we estimated forward rates from given spot rates, a process which required differentiation, using numerical differentiation technique. We call this the operator method.
2. We also developed a new interpolation method which removes the non-differentiability at the knot points by side stepping the monotonicity constraint set by Hyman, which was set to facilitate using of Fritsch Butland algorithm to estimate the parameters and at the same time avoid negative rates. In this study, we decided to relax this constraint given that negative rates are

actually being applied in different parts of the world. When compared to the current interpolation method, the function developed in this thesis performed the best when applied to current Kenyan bond data from NSE.

To compare the interpolation methods, we analyzed different interpolation methods by using test on stability, where we tested how perturbation at a point in a curve affected other points in the curve. We sampled three models: the model currently applied at the Johannesburg Securities exchange when tenures are missing from data, the monotone convex interpolation on $r(t)t$, the operator method and the thesis interpolation method. We tested the stability of both the interpolated spot rates and forward rates. Using monotone preserving interpolation on $r(t)t$ produced a mean of 26.52% and 28.86%, and a standard deviation of 2.86 and 11.93 for spot rates and forward rates respectively. The operator form produced a mean of 19.27% and 23.84%, and a standard deviation of 1.34 and 7.74 for spot rates and forward rates respectively. The thesis model produced a mean of 19.27% and 20.06%, and a standard deviation of 1.34 and 3.82 for spot rates and forward rates respectively. For the latter two methods, they produced the same results for spot rates because they were using the same interpolation function to generate the spot rates. However, when determining the forward rates, we used different methods. Comparing these results with the average of Kenya Central Banks lending rate and market rates (19.5-23%), we conclude that the model generated by this thesis produces rates that are reflective of the practical rates. The operator form performs better also compared to the monotone preserving interpolation on $r(t)t$ method.

Zero coupon yield curves can also be generated using parametric models. Many studies have been covered in establishing the best parametric models for risk free rates generation. In this study, we chose the most acclaimed three models: (Rezende, R. B. , 2011; Nelson, C. R., & Siegel, A. F. , 1987; Svensson, L. E. , 1994). We compare the models in terms of accuracy in pricing back the financial instruments, in this case the bonds, and smoothness. In terms of accuracy, we see that (Rezende, R. B. , 2011) performs the best with an accuracy of 97.38%, compared to (Svensson, L. E. , 1994) and (Nelson, C. R., & Siegel, A. F. , 1987) which have 96.93% and 96.54% respectfully. In terms of smoothness, Nelson-Siegel performs the best with smoothness value of $4.9822 \star 10^{-6}$, compared to (Svensson, L. E. , 1994) and (Rezende, R. B. , 2011) which have smoothness values of $6.3840 \star 10^{-6}$ and $10.7467 \star 10^{-6}$. Note that the smaller the smoothness value the smoother the curve. This result is consistent with other studies which have always shown that Nelson-Siegel performs better than the other models in its class. This is explained by the over-parametization of the newer models, which make them less smooth. In this category, we conclude that (Nelson, C. R., & Siegel, A. F. , 1987) is the best parametric model for generating the ZCYC for the NSE, despite the other two models having a higher accuracy.

In support of our accuracy findings, the Bank of International Settlements (2005), which reports on how central banks around the world calculate the ZCYC, has reported that central banks around the world , when determining monetary policy , do not typically require yield curve models that price back all the inputs exactly. Instead of 100% accuracy; 95% and above is deemed as adequate. We see that (Nelson, C. R., & Siegel, A. F. , 1987) Model has 96.54% accuracy level, which is above the required 95%.

We then compare the best parametric model with the best interpolation method. In this case, we compare the two in terms of both the accuracy of pricing back the input instrument and smoothness. We find that (Nelson, C. R., & Siegel, A. F. , 1987) performs better compared to the thesis interpolation model, with an accuracy of 96.54% compared to the thesis models 94.93%, and smoothness value of $4.9822 \star 10^{-6}$ compared to the thesis interpolation model which has $10.75 \star 10^{-6}$. After considering all the factors at hand, we come to a conclusion that the parametric model (Nelson-Siegel) is the slightly better than the model suggested by this thesis. However, we would like to point out that the decision on whether to use parametric models or spline-based methods to generate the yield curve depends on the intended use of the yield curve. Parametric methods are very popular amongst Central Banks. This is because Central banks typically do not require yield curves the prices back all inputs exactly, when determining monetary policy, and they find parametric methods easier to implement as compared to the spline models. On the other hand, Investment Bankers usually prefer spline-based models because parametric models might not completely reflect the information on the ground as compared to the spline models. For this reason, we use the interpolation method suggested in this thesis for the purpose of pricing the coffee futures, which has been shown to perform better than existing spline-based interpolation methods. Our method also offers the advantage of robustness in dealing with curves which suffer a sudden change of trend in response to events causing a change in market sentiment.

To develop a pricing model was a challenge because not only does the NSE not trade in coffee futures; coffee is actually not traded there as a commodity at all. The only available NSE data was the coffee beans prices from the auctions, we therefore had to look for alternative sources of data for the coffee commodity prices and futures. To found Nairobi auction prices on a traded commodity framework, we relate them to data from guide countries that grow coffee and at the same time trade in coffee futures. The countries we consider as guides of this nature are: Ethiopia, Cote dIvoire (Ivory Coast) and Brazil, and we relate their economies to that of Kenya by fairly crude interest-rate comparisons. The basis of comparison that we use is correlation analysis and Frechet distance.

In correlation analysis, the country which correlates most strongly with the Kenyan economy is taken as the best guide to starting an eventual commodity exchange price for Kenya. In problem formulation section 1.2 of the thesis, we consider instantaneous correlation of bank rates and also correlations with a time lag because the economic conditions driving auction bidding in Nairobi may well be those prevailing in the guide market at an earlier date. The results, over a range of plausible time-lags, were consistent in showing that data from Ivory Coast provides a better guide than those from the other two economies.

Seeing as we are dealing with comparison between the interest rate curves, we also used Frechet distance. This is a measure of similarity between two curves. It is defined as the minimum cord-length sufficient to join a point traveling forward along one curve and one traveling forward along the other curve, although the rate of travel for either point may not necessarily be uniform. The shorter the distance the closely related the curves are. From our analysis, Frechet distance comparison strongly supports Ivory Coasts interest rates as being more closely related to Kenyas interest rates than those of the other two countries, Brazil and Ethiopia.

Having concluded that Ivory Coasts coffee futures data was most suitable for pricing Kenyas coffee futures, we settled on using a simple variation of GBM, used by (Bellalah, M. , 1999; Bellalah, M., & Jacquillat, B. , 1995) because they price asset derivatives and real options under the uncertainty and incomplete information. This is a characteristic of emerging markets, NSE being one of them. In our pricing, we use an extension of the analysis in the (Schwartz, E. , 1998, 1997) to account for the effects of incomplete information as it appears in the models of (Bellalah, M. , 2001; Merton, R. C. , 1987). We also use the aforementioned extension to describe the stochastic behavior of commodity prices in the presence of mean reversion and shadow costs of incomplete information. The implications of the models are studied with respect to the valuation of financial and real assets. The parameters in the models of (Bellalah, M. , 2001; Merton, R. C. , 1987) using The L-BFGS-B algorithm by (Richard, L., & Burden, J. , 1988) is a standard method for solving large instances of $\underset{x \in R^n}{min} F(x)$ when F is a smooth function, typically twice differentiable with respect to the parameter of choice.

However, we estimated the drift and volatility from Ivorian futures prices data, using the Hough

Transform, an artificial intelligence pattern extraction method. We used the method suggested by (Onyango, S. N., & Ingleby, M. , 2006) to calculate grey-scale votes in each cell within the accumulator array. This is accomplished by using the χ^2 distribution to get the probability that the voting curve for window k in a price history passed through $\sigma^2 - \mu$ cell corresponding to σ_j^2 and μ_j. The grey voting technique is also used to simulate Ivorian coffee futures prices and evaluate the weights of votes of each cell. These voting weights are then accumulated in a 3-dimensional accumulator array. The cell curves pass through it. The corresponding values μ and σ^2 give the estimates of the mean rate of return and square of volatility of the underlying simulated asset prices respectively.

We find that with constant N and λ, the number of glitch points with noise (v=13) after which the values of the parameter estimate change is 8% of the total number of simulated asset prices. This shows that the transform is robust against glitches 8% data points (depending on size λ of the inaccuracy). This technique was shown by (Onyango, S. N., & Ingleby, M. , 2006) to be robust and able to communicate complex ideas with tremendous simplicity, clarity speed and power. It also supports effective or rapid decision making in the market suffering an unexpected change of sentiment. This is important because in the market, velocity of change is increasing amidst greater complexity and chaos, so processing a deep understanding of the market patterns can be critical to decision making whenever trading is to be undertaken. The futures market that we seek to serve is not free from glitch noise and sentiment shifts.

6.2 RECOMMENDATIONS FOR FURTHER STUDIES

This study has used linear trending univariate stochastic models for simplicity. A more comprehensive multivariate approach should be used where Kenyan data could be used alongside data from other markets. This research could be helpful for a market like Integrated East African market, where we see cross-trading among the East African securities markets. The results could be used in portfolio pricing and comparison, and also in establishing hedging positions using multiple derivative products.

The spline method suggested is still neither ideally accurate nor smooth as it should be. This means that there is need to improve the spline method suggested in this thesis, so as to improve on its accuracy which is at 94.93%. Internationally acceptable accuracy level is 95%.

In pattern recognition method applied in estimating the market parameters, the adaptation of the Hough Transform, though amply demonstrated, needs to be studied further. The method is able to deal with non-linearities such as logistic effects in place of the linear part of the stochastic models. It is also able to deal with vector asset price models that allow correlations between asset prices - particularly important where there are sector drags and contagion effects, possibly important also if comparing coffee futures at different exchanges around the world more thoroughly than in our search for a guide market such as Ivory Coast's.

Bibliography

1. Adams, K. J., & Van Deventer, D. R. (1994). Fitting yield curves and forward rate curves with maximum smoothness. The Journal of Fixed Income, 4(1), 5262. Retrieved from http://www.iijournals.com/doi/pdfplus/10.3905/jfi.1994.408102

2. Aguado, A. S., Montiel, E., & Nixon, M. S. (2000). Bias error analysis of the generalised Hough transform. Journal of Mathematical Imaging and Vision, 12(1), 2542. Retrieved from http://www.springerlink.com/index/U6118728J1705510.pdf

3. Akima, H. (1991). A method of univariate interpolation that has the accuracy of a third-degree polynomial. ACM Transactions on Mathematical Software (TOMS), 17(3), 341366. Retrieved from http://dl.acm.org/citation.cfm?id=116810

4. Akinci, O., Gurcihan, B., Gurkaynak, R., & Ozel, O. (2006). Devlet Ic Borclanma Senetleri Icin Getiri Egrisi Tahmini. Retrieved from https://ideas.repec.org/p/tcb/wpaper/0608.html

5. Alper, C. E., Akdemir, A., & Kazimov, K. (2004). Estimating the term structure of government securities in Turkey. Retrieved from https://papers.ssrn.com/soL3/papers.cfm?abstract_id=578921

6. Amihud, Y., Mendelson, H., & Lauterbach, B. (1997). Market microstructure and securities values: Evidence from the Tel Aviv Stock Exchange. Journal of Financial Economics, 45(3), 365390. Retrieved from http://www.sciencedirect.com/science/article/pii/S0304405X97000214

7. Anderson, L. (2007). Discount Curve Construction with Tension Spline. Notes on Review of Derivatives Research, 10, 227267. Retrieved from https://scholar.google.com/scholar?q=Anderson%2C+L.+%282007%29.+Discount+Curve+\Construction+with+Tension+Spline.+&btnG=&hl=en&as_sdt=1%2C5#

8. Anderson, N. (1996). Estimating and interpreting the yield curve (Vol. 6). John Wiley & Sons Incorporated.

9. Bachelier, L. (1900). Thorie de la spculation. Gauthier-Villars. Retrieved from http://physics.princeton.edu/~mcdonald/examples/statistics/bachelier_asens_17_21_00_english.pdf

10. Ballard, D. H. (1981). Generalizing the Hough transform to detect arbitrary shapes. Pattern Recognition, 13(2), 111122. Retrieved from http://www.sci.utah.edu/~gerig/CS6640-F2010/FINALPROJECT/Ballard-GHT-1981.pdf

11. Bates, D. S. (1996). 20 Testing option pricing models. Handbook of Statistics, 14, 567611. Retrieved from http://www.sciencedirect.com/science/article/pii/S0169716196140220

12. Bellalah, M. (1999). Les biais des modles doptions rvisits. Revue Franaise de Gestion, (124), 9499. Retrieved from http://cat.inist.fr/?aModele=afficheN&cpsidt=10120393

13. Bellalah, M. (2001). A reexamination of corporate risks under shadow costs of incomplete information. International Journal of Finance & Economics, 6(1), 4158. Retrieved from http://onlinelibrary.wiley.com/doi/10.1002/ijfe.139/abstract

14. Bellalah, M., & Jacquillat, B. (1995). Option valuation with information costs: theory and tests. Financial Review, 30(3), 617635. Retrieved from http://onlinelibrary.wiley.com/doi/10.1111/j.1540-6288.1995.tb00849.x/abstract

15. Bessembinder, H., & Kaufman, H. M. (1997). A cross-exchange comparison of execution costs and information flow for NYSE-listed stocks. Journal of Financial Economics, 46(3), 293319. Retrieved from http://www.sciencedirect.com/science/article/pii/S0304405X97000329

16. Bingham, N. H., Goldie, C. M., & Omey, E. (2006). Regularly varying probability densities. Publications de lInstitut Mathmatique, 80(94), 4757.Retrieved from http://emis.ams.org/journals/PIMB/094/n094p047.pdf

17. Bingham, N. H., Goldie, C. M., & Teugels, J. L. (1989). Regular variation (Vol. 27). Cambridge university press.Retrieved from https://books.google.com/books?hl=en&lr=&id=VCU5D4jMfiEC&oi=fnd&pg=PR17&dq=nick+bingham+regular+variation+(1989)&ots=oGdZxHI7RS&sig=3isAfedq1ayeUhUtDOFcRW1O4yI

18. Black, F., & Scholes, M. (1973). The pricing of options and corporate liabilities. Journal of Political Economy, 81(3), 637654. Retrieved from http://www.journals.uchicago.edu/doi/abs/10.1086/260062

19. Blair, B. J., Poon, S.-H., & Taylor, S. J. (2010). Forecasting S& P 100 volatility: the incremental information content of implied volatilities and high-frequency index returns. In Handbook of Quantitative Finance and Risk Management (pp. 13331344). Springer. Retrieved from http://link.springer.com/chapter/10.1007/978-0-387-77117-5_88

20. Bliss, R. R. (1996). Testing term structure estimation methods. Working paper, Federal Reserve Bank of Atlanta. Retrieved from https://www.econstor.eu/handle/10419/100824

21. Bolder, D. J., & Strliski, D. (1999). Yield curve modelling at the bank of canada. Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1082845

22. Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. Journal of Econometrics, 31(3), 307327. Retrieved from http://www.sciencedirect.com/science/article/pii/0304407686900631

23. Brennan, M. J., & Schwartz, E. S. (1979). A continuous time approach to the pricing of bonds. Journal of Banking & Finance, 3(2), 133155. Retrieved from http://www.sciencedirect.com/science/article/pii/0378426679900116

24. Broyden, C. G. (1970). The convergence of a class of double-rank minimization algorithms 2. The new algorithm. IMA Journal of Applied Mathematics, 6(3), 222231. Retrieved from http://imamat.oxfordjournals.org/content/6/3/222.short

25. Byrd, R. H., Lu, P., Nocedal, J., & Zhu, C. (1995). A limited memory algorithm for bound constrained optimization. SIAM Journal on Scientific Computing, 16(5), 11901208. Retrieved from http://epubs.siam.org/doi/abs/10.1137/0916069

26. Chambers, D. R., Carleton, W. T., & Waldman, D. W. (1984). A new approach to estimation of the term structure of interest rates. Journal of Financial and Quantitative Analysis, 19(03), 233252. Retrieved from http://journals.cambridge.org/article_S0022109000020366

27. Christofi, A. C. (1998). Estimation of the Nelson-Siegel parsimonious modeling of yield curves using an exponential GARCH process. Managerial Finance, 24(9/10), 119. Retrieved from http://www.emeraldinsight.com/doi/pdf/10.1108/03074359810765813

28. Conn, A. R., Gould, N. I., & Toint, P. L. (1988). Global convergence of a class of trust region algorithms for optimization with simple bounds. SIAM Journal on Numerical Analysis, 25(2), 433460. Retrieved from http://epubs.siam.org/doi/abs/10.1137/0725029

29. Cox, J. C., Ingersoll Jr, J. E., & Ross, S. A. (1985). A theory of the term structure of interest rates. Econometrica: Journal of the Econometric Society, 385407. Retrieved from http://www.jstor.org/stable/1911242

30. De Boor, C., & Schwartz, B. (1977). Piecewise monotone interpolation. Journal of Approximation Theory, 21(4), 411416. Retrieved from http://www.sciencedirect.com/science/article/pii/0021904577900119

31. Diebold, F. X., & Li, C. (2006). Forecasting the term structure of government bond yields. Journal of Econometrics, 130(2), 337364. Retrieved from http://www.sciencedirect.com/science/article/pii/S0304407605000795

32. Dothan, L. U. (1978). On the term structure of interest rates. Journal of Financial Economics, 6(1), 5969. Retrieved from http://www.sciencedirect.com/science/article/pii/0304405X7890020X

33. Du Preez, P. F. (2011). An investigation into popular methods for constructing yield curves. University of Pretoria. Retrieved from http://www.repository.up.ac.za/handle/2263/25882

34. Duan, J.-C. (1995). The GARCH option pricing model. Mathematical Finance, 5(1), 1332. Retrieved from http://onlinelibrary.wiley.com/doi/10.1111/j.1467-9965.1995.tb00099.x/full

35. Duda, R. O., & Hart, P. E. (1972). Use of the Hough transformation to detect lines and curves in pictures. Communications of the ACM, 15(1), 1115. Retrieved from http://dl.acm.org/citation.cfm?id=361242

36. Duffie, D., & Kan, R. (1996). A yieldfactor model of interest rates. Mathematical Finance, 6(4), 379406. Retrieved from http://onlinelibrary.wiley.com/doi/10.1111/j.1467-9965.1996.tb00123.x/full

37. Engle, R. F. (1982).Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. Econometrica: Journal of the Econometric Society, 9871007. Retrieved from http://www.jstor.org/stable/1912773

38. Exchange-NSE, N. S. (1997). Annual Report 1997. Nairobi: Nairobi Stock Exchange. Retrieved from https://www.nse.co.ke/inverstor-relations/financial-reports-and-results/category/46-nse-annual-reports.html?download=8987%3A2015-annual-report-digital-version

39. Fisher, M. (2001). Forces that shape the yield curve. Economic Review-Federal Reserve Bank of Atlanta, 86(1), 1. Retrieved from http://search.proquest.com/openview/f162e1140b1f282d68ccf0c1a50a7d5f/1?pq-origsite=gscholar&cbl=34496

40. Fleming, J. (1998). The quality of market volatility forecasts implied by S& P 100 index option prices. Journal of Empirical Finance, 5(4), 317345. Retrieved from http://www.sciencedirect.com/science/article/pii/S0927539898000024

41. Fletcher, R. (1970). A new approach to variable metric algorithms. The Computer Journal, 13(3), 317322. Retrieved from http://comjnl.oxfordjournals.org/content/13/3/317.short

42. Fritsch, F. N., & Carlson, R. E. (1980). Monotone piecewise cubic interpolation. SIAM Journal on Numerical Analysis, 17(2), 238246. Retrieved from http://epubs.siam.org/doi/abs/10.1137/0717021

43. Garman, M. B., & Klass, M. J. (1980). On the estimation of security price volatilities from historical data. Journal of Business, 6778. Retrieved from http://www.jstor.org/stable/2352358

44. Gilli, M., Groe, S., & Schumann, E. (2010). Calibrating the nelson-siegel-svensson model. Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1676747

45. Goldfarb, D. (1970). A family of variable-metric methods derived by variational means. Mathematics of Computation, 24(109), 2326. Retrieved from http://www.ams.org/mcom/1970-24-109/S0025-5718-1970-0258249-6/

46. Gorton, G., Lewis, M., & Zuckerman, G. (2011). The Big Short: Inside the Doomsday Machine. JSTOR. Retrieved from https://books.google.com/books?hl=en&lr=&id=erDXCgAAQBAJ&oi=fnd&pg=PT11&dq=the+big+shot:+inside+the+doomsday+machine+lewis&ots=c-BsB2QOfX&sig=gRymcnH19H738WVyLHFjIDNikog

47. Hagan, P. S., & West, G. (2006). Interpolation methods for curve construction. Applied Mathematical Finance, 13(2), 89129. Retrieved from http://www.tandfonline.com/doi/abs/10.1080/13504860500396032

48. Hagan, P. S., & West, G. (2008). Methods for constructing a yield curve. Wilmott Magazine, May, 7081. Retrieved from https://www.researchgate.net/profile/Patrick_Hagan3/publication/228463045_Methods_for_constructing_a_yield_curve/links/54db8cda0cf23fe133ad4d01.pdf

49. Heath, D., Jarrow, R., & Morton, A. (1992). Bond pricing and the term structure of interest rates: A new methodology for contingent claims valuation. Econometrica: Journal of the Econometric Society, 77105. Retrieved from http://www.jstor.org/stable/2951677

50. Heston, S. L. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options. Review of Financial Studies, 6(2), 327343. Retrieved from http://rfs.oxfordjournals.org/content/6/2/327.short

51. Heston, S. L., & Nandi, S. (2000). A closed-form GARCH option valuation model. Review of Financial Studies, 13(3), 585625. Retrieved from http://rfs.oxfordjournals.org/content/13/3/585.short

52. Ho, T. S., & LEE, S.-B. (1986). Term structure movements and pricing interest rate contingent claims. The Journal of Finance, 41(5), 10111029. Retrieved from http://onlinelibrary.wiley.com/doi/10.1111/j.1540-6261.1986.tb02528.x/full

53. Hough, P. V. (1962). Method and means for recognizing complex patterns. Retrieved from http://www.osti.gov/scitech/biblio/4746348

54. Hull, J. C. (2006). Options, futures, and other derivatives. Pearson Education India. Retrieved from http://userhome.brooklyn.cuny.edu/bassell/myles_bassell/Myles_Bassell_704S.pdf

55. Hull, J., & White, A. (1987). The pricing of options on assets with stochastic volatilities. The Journal of Finance, 42(2), 281300. Retrieved from http://onlinelibrary.wiley.com/doi/10.1111/j.1540-6261.1987.tb02568.x/full

56. Hull, J., & White, A. (1990). Pricing interest-rate-derivative securities. Review of Financial Studies, 3(4), 573592. Retrieved from http://rfs.oxfordjournals.org/content/3/4/573.short

57. Hyman, J. M. (1983). Accurate monotonicity preserving cubic interpolation. SIAM Journal on Scientific and Statistical Computing, 4(4), 645654. Retrieved from http://epubs.siam.org/doi/abs/10.1137/0904045

58. Illingworth, J., & Kittler, J. (1988). A survey of the Hough transform. Computer Vision, Graphics, and Image Processing, 44(1), 87116. Retrieved from http://www.sciencedirect.com/science/article/pii/S0734189X88800331

59. Ingersoll, J. E. (1987). Theory of financial decision making (Vol. 3). Rowman & Littlefield. Retrieved from https://books.google.com/books?hl=en&lr=&id=_JTWMD_Cm-QC&oi=fnd&pg=PR9&dq=Ingersoll,+J.+E.+(1987).+Theory+of+financial+decision+making+&ots=EdbnmCLvr4&sig=FYO5xiia7TinkXa11_LvM4JyamI

60. Kenya, C. B. of. (1986). Central Bank of Kenya: its evolution, responsibilities and organization. The Bank.

61. Leavers, V. F. (1992).A Case Study: Circles and Ellipses. In Shape Detection in Computer Vision Using the Hough Transform (pp. 137172). Springer. Retrieved from http://link.springer.com/chapter/10.1007/978-1-4471-1940-1_7

62. Lewis, A. S., & Overton, M. L. (2013). Nonsmooth optimization via quasi-Newton methods. Mathematical Programming, 129.

63. Lyuu, Y.-D. (2001). Financial engineering and computation: principles, mathematics, algorithms. Cambridge University Press. Retrieved from https://books.google.com/books?hl=en&lr=&id=iispAwAAQBAJ&oi=fnd&pg=PP1&dq=Lyuu,+Y.-D.+(2001).+Financial+engineering+and+computation:+principles,+mathematics,+algorithms.\+Cambridge+University+Press.&ots=dS4hDNjVwL&sig=JuIQxcDByNVqwi0b575x6n49_TA

64. Madhavan, A. (1992). Trading mechanisms in securities markets. The Journal of Finance, 47(2), 607641. Retrieved from http://onlinelibrary.wiley.com/doi/10.1111/j.1540-6261.1992.tb04403.x/full

65. Malz, A. M. (2000). Do implied volatilities provide early warning of market stress? Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=211448

66. Maria, L. M., Leanez, C., & Moreno, M. (2009). Estimation of the Term Structure of Interest Rates: The Venezuelan Case. Journal of Future Markets, 23, 10751105.

67. McCulloch, J. H. (1971). Measuring the term structure of interest rates. The Journal of Business, 44(1), 1931. Retrieved from http://www.jstor.org/stable/2351832

68. Melino, A., & Turnbull, S. M. (1990). Pricing foreign currency options with stochastic volatility. Journal of Econometrics, 45(12), 239265. Retrieved from http://www.sciencedirect.com/science/article/pii/0304407690901008

69. Merton, R. C. (1973). Theory of rational option pricing. The Bell Journal of Economics and Management Science, 141183. Retrieved from http://www.jstor.org/stable/3003143

70. Merton, R. C. (1987). A simple model of capital market equilibrium with incomplete information. The Journal of Finance, 42(3), 483510. Retrieved from http://onlinelibrary.wiley.com/doi/10.1111/j.1540-6261.1987.tb04565.x/full

71. Modena, M. (2008). An empirical analysis of the curvature factor of the term structure of interest rates. Retrieved from https://ideas.repec.org/p/gla/glaewp/2008_35.html

72. Mor, J. J., & Toraldo, G. (1989). Algorithms for bound constrained quadratic programming problems. Numerische Mathematik, 55(4), 377400. Retrieved from http://link.springer.com/article/10.1007/BF01396045

73. Munasib, A., & Haurin, D. (2004). Time to First Homeownership: Racial Differences, and the Impact of 1986 Tax Reform Act. Retrieved from http://spears.okstate.edu/home/munasib/papers/FirstTimeHomeownership.pdf

74. Muthoni, L., Onyango, S., & Ongati, N. O. (2015a). Construction of nominal yield curve for Nairobi securities exchange: an improvement on monotone preserving r (t) t interpolation method. Retrieved from http://62.24.102.115:8080/handle/123456789/237

75. Muthoni, L., Onyango, S., & Ongati, N. O. (2015b). Extraction of Zero Coupon Yield Curve for Nairobi Securities Exchange: Finding the Best Parametric Model for East African Securities Markets. SIAM Journal on Financial Mathematics. Retrieved from http://www.ss-pub.org/wp-content/uploads/2015/09/JMSS15062201.pdf

76. Nelson, C. R., & Siegel, A. F. (1987). Parsimonious modeling of yield curves. Journal of Business, 473489. Retrieved from http://www.jstor.org/stable/2352957

77. Ngugi, R., & Agoti, J. (2007). Microstructure Elements of The Bonds Market in Kenya. Kenya Institute for Public Policy Research and Analysis. Retrieved from https://www.afdb.org/fileadmin/uploads/afdb/Documents/Knowledge/30753253-FR-123-NGUGI-BONDS-PAPER-AEC.PDF

78. Nocedal, J. (1980). Updating quasi-Newton matrices with limited storage. Mathematics of Computation, 35(151), 773782. Retrieved from http://www.ams.org/mcom/1980-35-151/S0025-5718-1980-0572855-7/

79. Onyango, S. N., & Ingleby, M. (2006). On the pattern recognition of It$\hat{o}$ processes in market price data. Environmental Economics and Investment Assessment, 98, 243. Retrieved from https://books.google.com/books?hl=en&lr=&id=ZRnVJmBXX-YC&oi=fnd&pg=PA243&dq=Onyango,+S.+N.,+%26+Ingleby,+M.+(2006).+On+the+pattern+recognition+of+It%C3\%B4+processes+in+market+price+data.+Environmental+Economics+and+Investment+\Assessment,+98,+243.&ots=G-AlWgOu9F&sig=k9mXUAzJd3Bf5SO85RCDDxs5qUO

80. PIM(2004). Pacific Investment Management Company LLC: Private Company Information - Bloomberg. Retrieved January 23, 2017, from http://www.bloomberg.com/research/stocks/private/snapshot.asp?privcapId=907219

81. Pearson, K. (1905). The problem of the random walk. Nature, 72(1865), 294.

82. Place, J. (2000). Basic bond analysis. Bank of England. Retrieved from http://www.bankofengland.co.uk/education/Documents/ccbs/handbooks/pdf/ccbshb20.pdf

83. Rezende, R. B. (2011). Giving flexibility to the Nelson-Siegel class of term structure models. Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1290784

84. Richard, L., & Burden, J. (1988). Douglas Faires, Numerical Analysis. PWS Publishing Co., Boston, MA.

85. Rosadi, D. (2011). Modelling of Yield Curve and Its Computation by Software Packages Remdr-Plugin. Econometric. Proceeding FMIPA UnDip, 514523.

86. Schwartz, E. (1998). Valuing long-term commodity assets. Journal of Energy Finance & Development, 3(2), 8599. Retrieved from http://www.sciencedirect.com/science/article/pii/S1085744399800702

87. Schwartz, E. (1997). The stochastic behavior of commodity prices: Implications for valuation and hedging. The Journal of Finance, 52(3), 923973. Retrieved from http://onlinelibrary.wiley.com/doi/10.1111/j.1540-6261.1997.tb02721.x/full

88. Ser, P.-K., & Siu, W.-C. (1995). Memory compression for straight line recognition using the Hough transform. Pattern Recognition Letters, 16(2), 133145. Retrieved from http://www.sciencedirect.com/science/article/pii/016786559400078H

89. Shanno, D. F. (1970). Conditioning of quasi-Newton methods for function minimization. Mathematics of Computation, 24(111), 647656. Retrieved from http://www.ams.org/mcom/1970-24-111/S0025-5718-1970-0274029-X/home.html

90. Subramanian, K. V. (2001). Term structure estimation in illiquid markets. The Journal of Fixed Income, 11(1), 7786. Retrieved from http://www.iijournals.com/doi/pdfplus/10.3905/jfi.2001.319292

91. Svensson, L. E. (1994). Estimating and interpreting forward interest rates: Sweden 1992-1994. National Bureau of Economic Research. Retrieved from http://www.nber.org/papers/w4871

92. Svensson, L. E. (1995).Estimating forward interest rates with the extended Nelson & Siegel method. Sveriges Riksbank Quarterly Review, 3(1), 1326.

93. Tam, C.-S., & Yu, I.-W. (2008). Modelling sovereign bond yield curves of the US, Japan and Germany. International Journal of Finance & Economics, 13(1), 8291. Retrieved from http://onlinelibrary.wiley.com/doi/10.1002/ijfe.353/full

94. Toft, P. A., & Srensen, J. A. (1996). The Radon transform-theory and implementation. Technical University of DenmarkDanmarks Tekniske Universitet, Department of Informatics and Mathematical ModelingInstitut for Informatik og Matematisk Modellering. Retrieved from http://orbit.dtu.dk/fedora/objects/orbit:85899/datastreams/file_5529668/content

95. Vasicek, O. (1977). An equilibrium characterization of the term structure. Journal of Financial Economics, 5(2), 177188. Retrieved from http://www.sciencedirect.com/science/article/pii/0304405X77900162

96. Weatherburn, C. E. (1949). A first course mathematical statistics (Vol. 158). CUP Archive. Retrieved from https://books.google.com/books?hl=en&lr=&id=xFA4AAAAIAAJ&oi=fnd&pg=PR13&dq=Weatherburn,+C.+E.+(1949).+A+first+course+mathematical+statistics+\(Vol.+158).+CUP+Archive.&ots=SW87w4bHuB&sig=cZ8WaRWApD9VcBuAunI6a8Hv0Yg

97. Wilmott, P. (1998). Derivatives. Retrieved from http://www.citeulike.org/group/2384/article/1204422

98. Wright, S., & Nocedal, J. (1999). Numerical optimization. Springer Science, 35, 6768. Retrieved from http://home.agh.edu.pl/~pba/pdfdoc/Numerical_Optimization.pdf

99. Xu, X., & Taylor, S. J. (1995). Conditional volatility and the informational efficiency of the PHLX currency options market. Journal of Banking & Finance, 19(5), 803821. Retrieved from http://www.sciencedirect.com/science/article/pii/037842669500086V

100. Yip, R. K., Tam, P. K., & Leung, D. N. (1995). Modification of Hough transform for object recognition using a 2-dimensional array. Pattern Recognition, 28(11), 17331744. Retrieved from http://www.sciencedirect.com/science/article/pii/S1077314296900417

101. Yuen, S. Y., & Ma, C. H. (1997). An investigation of the nature of parameterization for the Hough transform. Pattern Recognition, 30(6), 10091040. Retrieved from http://www.sciencedirect.com/science/article/pii/S0031320397852673

102. Zhu, C., Byrd, R. H., Lu, P., & Nocedal, J. (1994). LBFGS-B: Fortran subroutines for large-scale bound constrained optimization. Report NAM-11, EECS Department, Northwestern University. Retrieved from https://pdfs.semanticscholar.org/5a22/53e72c71012c3a970689c1627492172440b7.pdf

103. Zhu, C., Byrd, R. H., Lu, P., & Nocedal, J. (1997). Algorithm 778: L-BFGS-B: Fortran subrou-

tines for large-scale bound-constrained optimization. ACM Transactions on Mathematical Software (TOMS), 23(4), 550560. Retrieved from http://dl.acm.org/citation.cfm?id=279236

APPENDICES

APPENDIX A: THE L-BFGS-B ALGORITHM

A.1 Introduction

The problem addressed is to find a local minimizer of the non-smooth minimization problem.

$$\min_{x \in \mathbb{R}^n} f(x) \tag{A1}$$

$$s.t. l_i \leq x_i \leq u_i$$

$$i = 1, \ldots, n$$

where $f : \mathbb{R}^n \rightarrow \mathbb{R}$ is continuous but not differentiable anywhere and n is large. l_i and u_i are respectively an upper limit and lower limit parameters. $f(x)$ is NLS (Non Linear Schrdinger) function of residual functions of Nelson-Siegel model class and x is a parameter of the Nelson-Siegel model class.

The L-BFGS-B algorithm by (Richard, L., & Burden, J. , 1988) is a standard method for solving large instances of $\min_{x \in \mathbb{R}^n} f(x)$ when f is a smooth function, typically twice differentiable.

The name BFGS stands for Broyden, Fletcher, and Goldfarb and Shanno, the originators of the BFGS quasi-Newton algorithm for unconstrained optimization discovered and published independently by them in 1970 (Broyden, C. G. , 1970; Fletcher, R. , 1970; Goldfarb, D. , 1970; Shanno, D. F. , 1970). This method requires storing and updating a matrix which approximates the inverse of the Hessian $\nabla^2 f(x)$ and hence requires $O(n^2)$ operations per iteration. According to (Nocedal, J. , 1980), the L-BFGS variant where the L stands for "Limited-Memory" and also for "Large" problems, is based on BFGS but requires only $O(n)$ operations per iteration, and less memory. Instead of storing the $n \times n$ Hessian approximations, L-BFGS stores only m vectors of dimension n, where m is a number much smaller than n. Finally, the last letter B in L-BFGS stands for bounds, meaning the lower and upper bounds l_i and u_i. The L-BFGS-B algorithm is implemented in a FORTRAN software package, according to (Zhu, C., Byrd, R. H., Lu, P., & Nocedal, J. , 1994). We discuss how to modify the algorithm for non-smooth functions.

A.2 BFGS

BFGS is standard tool for optimization of smooth functions. It is a line search method. The search direction is of type $= -B_k \nabla f(x_k)$ where B_k approximation to the inverse Hessian of f [1]. This k^{th} step approximation is calculated via the BFGS formula.

$$B_{k+1} = \left(I - \frac{s_k {y_k}^T}{{y_k}^T s_k}\right) B_k + \left(I - \frac{y_k {s_k}^T}{{y_k}^T s_k}\right) + \frac{s_k {s_k}^T}{{y_k}^T s_k} \tag{A2}$$

where $y_k = \nabla f(x_{k+1}) - \nabla f(x_k)$ and $s_k = x_{k+1} - x_k$. BFGS exhibits super-linear convergence on generic problems but it requires $O(n^2)$ operations per iteration, according to (Wright, S., & Nocedal, J. , 1999).

In the case of non-smooth functions, BFGS typically succeeds in finding a local minimizer. However, this requires some attention to the line search conditions. This conditions are known as the Armijo and weak Wolfe line search conditions and they are a set of inequalities used for computation of an appropriate step length that reduces the objective function "sufficiently".

A.3 L-BFGS

L-BFGS stands for Limited-memory BFGS. This algorithm approximates BFGS using only a limited amount of computer memory to update an approximation to the inverse of the Hessian of f. Instead of storing a dense $n \times n$ matrix, L-BFGS keeps a record of the last m is a small number that is chosen in advance. For this reason the first m iterations of BFGS and L-BFGS produce exactly the same search directions if the initial approximation of B_0 is set to the identity matrix.

Because of this construction, the L-BFGS algorithm is less computationally intensive and requires only $O(mn)$ operations per iteration. So it is much better suited for problems where the number of dimensions n is large.

[1] When it is exactly the inverse Hessian this method is known as Newton method. Newton‘s method has quadratic convergence but requires the explicit calculation of the Hessian at every step

A.4 L-BFGS-B

Finally L-BFGS-B is an extension of L-BFGS. The B stands for the inclusion of Boundaries. L-BFGS-B requires two extra steps on top of L-BFGS. First, there is a step called gradient projection that reduces the dimensionality of the problem. Depending on the problem, the gradient projection could potentially save a lot of iterations by eliminating those variables that are on their bounds at the optimum reducing the initial dimensionality of the problem and the number of iterations and running time. After this gradient projection comes to second step of subspace minimization. During the subspace minimization phase, an approximate quadratic model of (A1) is solved iteratively in a similar way that the original L-BFGS algorithm is solved. The only difference is that the step length is restricted as much as necessary in order to remain within the l,u-box defined by equation (A1).

A.5 Gradient Projection

The L-BFGS-B algorithm was designed for the case when n is large and f is smooth. Its first step is the gradient projection similar to the one outlined in (Conn, A. R., Gould, N. I., & Toint, P. L. , 1988) and (Mor, J. J., & Toraldo, G. , 1989), which is used to determine an active set corresponding to those variables that are on either their lower or upper bounds. The active set is defined at point x^* is:

$$A(x^*) = \{i \in \{1, \dots, n\} | x_i^* = l_i V x_i^* = u_i\} \tag{A3}$$

Working with this active set is more efficient in large scale problems. A pure line search algorithm would have to choose to step length short enough to remain within the box defined by l_i and u_i. So if at the optimum, a large number $\boldsymbol{B}$ of variables are either on the lower or upper bound, as many as B of iterations might be needed. Gradient projection tries to reduce this number of iterations. In the best case, only one iteration is needed instead of $\boldsymbol{B}$.

Gradient projections works on the linear part of the approximation model:

$$m_k(x) = f(x_k) + \nabla f(x_k)^T (x - x_k) + \frac{(x - x_k)^T H_k (x - x_k)}{2} \tag{A4}$$

whereH_k is a L-BFGS-B approximation to the Hessian $\nabla^2 f$ stored in the implicit way defined by L-BFGS.

In this first stage of the algorithm a piece-wise linear path starts at the current point x_k in the direction-$\nabla f(x_k)$. Whenever this direction encounters one of the constraints the path runs corners in order to remain feasible. The path is nothing but feasible piece-wise projection of the negative gradient direction on the constraint box determined by the values l and u. At the end of this stage, the value of x that minimizes $m_k(x)$ restricted to this piece-wise gradient path is known as the "Cauchy point"x^c.

From this description of the estimation and optimization, following steps can be summarized:

- Find the residual function (r) of each model.
- Find NLS estimation, i.e. $f(x_i) = \frac{1}{2}\sum_{i=1}^{p}[x_i]^2$, of each model.
- Find the $p \times p$ matrix value for$B_1 = I$, p is the number of parameters estimated in each model.
- Find the initial value of parameter vector with rank $p \times 1$, p is the number of parameters estimated in each model.
- Find gradient from step 2 with every parameter in models. e.g. $\nabla f(x_i)_i$.
- Substitute the initial value of the parameter (step 3) to gradient of step 5 with result. e.g.$\nabla f(x_1)$.
- Find the value of p_1.

Find the value of$f(x_1)$ so it will obtain of d_1 and s_1.

APPENDIX B: DATA, CODES AND CURVES

B.1 Interest Rates Data and Codes

B.1.1 Brazil

The following is the analysis of the interest rates of Brazil using correlation with time lags and Frechet distance methods:

1. **Correlation**

R codes:

Brazil $<-$c(45, 42, 42, 39.5, 34, 32, 32, 29.5, 27, 23.5, 23.5, 22, 21, 21, 19.5, 19.5, 19.5, 19, 19, 19, 19, 19, 19, 19, 18.5, 18.5, 18.5, 18.5, 17.5, 17.5, 17, 16.5, 16.5, 16.5, 16.5, 16.5, 16.5, 15.75, 15.75, 15.25, 15.25, 15.25, 15.75, 15.75, 16.25, 16.25, 16.75, 16.75, 18.25, 18.25, 19, 19, 19, 19, 19, 19, 19, 19, 18.75, 18.75, 18.5, 18.5, 18.5, 18.5, 18.5, 18, 18)

Kenya $<-$c(21.76, 21.63, 23.1, 24.08, 22.09, 21.53, 21.61, 21.44, 21.42, 21.02, 20.35, 19.44, 18.45, 19.69, 26.2, 27.15, 26.78, 26.36, 26.28, 26.33, 26.74, 26.98, 26.38, 25.48, 24.67, 23.74, 22.47, 20.59, 17.66, 12.56, 10.7, 8.95, 8.84, 9.03, 9.63, 11.44, 14.47, 14.84, 15.78, 17.63, 18.14, 19.97, 20.3, 14.84, 11.28, 12.44, 11.22, 10.47, 9.9 , 9.25, 10.36, 10.65, 11.17, 12.9, 14.76, 15.3, 14.97, 12.9, 10.52, 12.07, 12.87, 12.84, 12.39, 11.63, 11.5, 11.01, 10.85)

```
Brazil_ Kenya<-data.frame(Brazil,Kenya)
acf<-acf(Brazil_ Kenya,lag.max=7,plot=F)
acf
plot(acf)
```

Output:

Autocorrelations of series Brazil_ Kenya, by lag

, , Brazil

Brazil		Kenya	
1.000	(0)	0.389	(0)
0.883	(1)	0.397	(−1)
0.777	(2)	0.404	(−2)
0.658	(3)	0.400	(−3)
0.545	(4)	0.383	(−4)
0.464	(5)	0.383	(−5)
0.388	(6)	0.394	(−6)
0.303	(7)	0.408	(−7)

, , Kenya

Brazil		Kenya	
0.389	(0)	1.000	(0)
0.329	(1)	0.946	(1)
0.277	(2)	0.855	(2)
0.226	(3)	0.757	(3)
0.178	(4)	0.652	(4)
0.141	(5)	0.555	(5)
0.108	(6)	0.461	(6)
0.076	(7)	0.369	(7)

2. **Frechet Distance**

R codes:

$$Brazil < -c(45, 42, 42, 39.5, 34, 32, 32, 29.5, 27, 23.5, 23.5, 22, 21, 21, 19.5, 19.5, 19.5, 19, 19, 19, 19, 19, 19, 19, 18.5, 18.5, 18.5, 18.5, 17.5, 17.5, 17, 16.5, 16.5, 16.5, 16.5, 16.5, 16.5, 15.75, 15.75, 15.25, 15.25, 15.25, 15.75, 15.75, 16.25, 16.25, 16.75, 16.75, 18.25, 18.25, 19, 19, 19, 19, 19, 19, 19, 19, 18.75, 18.75, 18.5, 18.5, 18.5, 18.5, 18.5, 18, 18)$$

$$Kenya < -c(21.76, 21.63, 23.1, 24.08, 22.09, 21.53, 21.61, 21.44, 21.42, 21.02, 20.35, 19.44, 18.45, 19.69, 26.2, 27.15, 26.78, 26.36, 26.28, 26.33, 26.74, 26.98, 26.38, 25.48, 24.67, 23.74, 22.47, 20.59, 17.66, 12.56, 10.7, 8.95, 8.84, 9.03, 9.63, 11.44, 14.47, 14.84, 15.78, 17.63, 18.14, 19.97, 20.3, 14.84, 11.28, 12.44, 11.22, 10.47, 9.9, 9.25, 10.36, 10.65, 11.17, 12.9, 14.76, 15.3, 14.97, 12.9, 10.52, 12.07, 12.87, 12.84, 12.39, 11.63, 11.5, 11.01, 10.85)$$

```
brazil<-sample(Brazil,10,replace=TRUE)
brazil
kenya<-sample(Kenya,10,replace=TRUE)
kenya
distFrechet(Kenya,Brazil,kenya,brazil)
```

Output:

Result = 13.00839

[1]13.00839

B.1.2 Ethiopia

The following is the analysis of the interest rates of Ethiopia using correlation with time lags and Frechet distance methods:

1. **Correlation**

R codes:

$$Ethiopia <- c(10,10,10,10,7,7,7,7,7,7,7,7,7,7,7,7,7,7,7,6,6,6,6,6,6,6,6,6,6,6,6,6,6\\ ,6,6,6,6,6,6,6,6,6,6,6,6,6,6,6,6,6,6,6,6,6,6,6,6,6,6,6,6,6,6,6,6,6)$$

$$Kenya <- c(21.76,21.63,23.1,24.08,22.09,21.53,21.61,21.44,21.42,21.02,20.35,19.44,18.45,19.69,26.2,\\ 27.15,26.78,26.36,26.28,26.33,26.74,26.98,26.38,25.48,24.67,23.74,22.47,20.59,17.66,12.56,10.7,8.95,\\ 8.84,9.03,9.63,11.44,14.47,14.84,15.78,17.63,18.14,19.97,20.3,14.84,11.28,12.44,11.22,10.47,\\ 9.9,9.25,10.36,10.65,11.17,12.9,14.76,15.3,14.97,12.9,10.52,12.07,12.87,12.84,12.39,11.63,11.5,\\ 11.01,10.85)$$

Ethiopia_ Kenya<-data.frame(Ethiopia,Kenya)

acf<-acf(Ethiopia_ Kenya,lag.max=7,plot=F)

acf

plot(acf)

Output:

Autocorrelations of series Ethiopia_ Kenya, by lag

, , Ethiopia

Ethiopia		Kenya	
1.000	(0)	0.433	(0)
0.824	(1)	0.453	(−1)
0.649	(2)	0.470	(−2)
0.473	(3)	0.476	(−3)
0.297	(4)	0.469	(−4)
0.285	(5)	0.479	(−5)
0.274	(6)	0.488	(−6)
0.262	(7)	0.489	(−7)

, , Kenya

Ethiopia		Kenya	
0.433	(0)	1.000	(0)
0.349	(1)	0.946	(1)
0.273	(2)	0.855	(2)
0.208	(3)	0.757	(3)
0.141	(4)	0.652	(4)
0.112	(5)	0.555	(5)
0.100	(6)	0.461	(6)
0.092	(7)	0.369	(7)

2. **Frechet Distance**

   R codes:

   Ethiopia < −c(10, 10, 10, 10, 7, 7, 7, 7, 7, 7, 7, 7, 7, 7, 7, 7, 7, 7, 7, 6, 6, 6, 6, 6, 6, 6, 6, 6, 6, 6, 6, 6,
   6, 6, 6, 6, 6, 6, 6, 6, 6, 6, 6, 6, 6, 6, 6, 6, 6, 6, 6, 6, 6, 6, 6, 6, 6, 6, 6, 6, 6, 6, 6, 6, 6, 6, 6)

   Kenya < −c(21.76, 21.63, 23.1, 24.08, 22.09, 21.53, 21.61, 21.44, 21.42, 21.02, 20.35,
   19.44, 18.45, 19.69, 26.2, 27.15, 26.78, 26.36, 26.28, 26.33, 26.74, 26.98, 26.38, 25.48, 24.67, 23.74, 22.47, 20.59,
   17.66, 12.56, 10.7, 8.95, 8.84, 9.03, 9.63, 11.44, 14.47, 14.84, 15.78, 17.63, 18.14, 19.97, 20.3, 14.84,
   11.28, 12.44, 11.22, 10.47, 9.9, 9.25, 10.36, 10.65, 11.17, 12.9, 14.76, 15.3, 14.97, 12.9, 10.52, 12.07,
   12.87, 12.84, 12.39, 11.63, 11.5, 11.01, 10.85)

   ethiopia<-sample(Ethiopia,10,replace=TRUE)
   ethiopia
   kenya<-sample(Kenya,10,replace=TRUE)
   kenya
   distFrechet(Kenya,Ethiopia,kenya,ethiopia)

   Output:

   Result = 4.113821
   [1]4.113821

B.1.3 Ivory Coast

The following is the analysis of the interest rates of Ethiopia using correlation with time lags and Frechet distance methods:

1. **Correlation**

   R codes:

   Ivory < −c(4.25, 4.25, 4.25, 4.25, 4.25, 4.25, 4.25, 4.25, 4.25, 4.25, 4.25, 4.25, 4.25, 4.25,
   4.25, 4.25, 4.25, 4.25, 4, 4, 4, 4, 4, 4, 4, 4, 4, 3.75, 3.75, 3.75, 3.75, 3.75, 3.75, 3.5, 3.5, 3.5, 3.5, 3.5, 3.5,
   3.5, 3.5, 3.5, 3.5, 3.5, 3.5, 3.5, 3.5, 3.5, 3.5, 3.5, 3.5, 3.5, 3.5, 3.5, 3.5, 3.5, 3.5, 3.5, 3.5, 3.5, 3.5, 3.5,
   3.5, 3.5, 3.5, 3.5, 3.5)

$Kenya <- c(21.76, 21.63, 23.1, 24.08, 22.09, 21.53, 21.61, 21.44, 21.42, 21.02, 20.35,$
$19.44, 18.45, 19.69, 26.2, 27.15, 26.78, 26.36, 26.28, 26.33, 26.74, 26.98, 26.38, 25.48, 24.67, 23.74, 22.47,$
$20.59, 17.66, 12.56, 10.7, 8.95, 8.84, 9.03, 9.63, 11.44, 14.47, 14.84, 15.78, 17.63, 18.14, 19.97, 20.3,$
$14.84, 11.28, 12.44, 11.22, 10.47, 9.9, 9.25, 10.36, 10.65, 11.17, 12.9, 14.76, 15.3, 14.97, 12.9, 10.52,$
$12.07, 12.87, 12.84, 12.39, 11.63, 11.5, 11.01, 10.85)$

```
Ivory_ Kenya<-data.frame(Ivory,Kenya)
acf<-acf(Ivory_ Kenya,lag.max=7,plot=F)
acf
plot(acf)
```

Output:

Autocorrelations of series Ivory_ Kenya, by lag

, , Ivory

Ivory		Kenya	
1.000	(0)	0.791	(0)
0.966	(1)	0.783	(−1)
0.932	(2)	0.771	(−2)
0.898	(3)	0.749	(−3)
0.863	(4)	0.725	(−4)
0.829	(5)	0.705	(−5)
0.795	(6)	0.687	(−6)
0.752	(7)	0.671	(−7)

, , Kenya

Ivory		Kenya	
0.791	(0)	1.000	(0)
0.765	(1)	0.946	(1)
0.736	(2)	0.855	(2)
0.702	(3)	0.757	(3)
0.665	(4)	0.652	(4)
0.630	(5)	0.555	(5)
0.592	(6)	0.461	(6)
0.550	(7)	0.369(7)	

2. **Frechet Distance**

   R codes:

$$
\begin{gathered}
\text{Ivory} <-\text{c}(4.25, 4.25, 4.25, 4.25, 4.25, 4.25, 4.25, 4.25, 4.25, 4.25, 4.25, 4.25, 4.25, 4.25,\\
4.25, 4.25, 4.25, 4.25, 4, 4, 4, 4, 4, 4, 4, 4, 3.75, 3.75, 3.75, 3.75, 3.75, 3.75, 3.5, 3.5, 3.5, 3.5, 3.5, 3.5,\\
3.5, 3.5, 3.5, 3.5, 3.5, 3.5, 3.5, 3.5, 3.5, 3.5, 3.5, 3.5, 3.5, 3.5, 3.5, 3.5, 3.5, 3.5, 3.5, 3.5, 3.5, 3.5, 3.5,\\
3.5, 3.5, 3.5, 3.5, 3.5)
\end{gathered}
$$

$$
\begin{gathered}
\text{Kenya} <-\text{c}(21.76, 21.63, 23.1, 24.08, 22.09, 21.53, 21.61, 21.44, 21.42, 21.02, 20.35,\\
19.44, 18.45, 19.69, 26.2, 27.15, 26.78, 26.36, 26.28, 26.33, 26.74, 26.98, 26.38, 25.48, 24.67, 23.74, 22.47,\\
20.59, 17.66, 12.56, 10.7, 8.95, 8.84, 9.03, 9.63, 11.44, 14.47, 14.84, 15.78, 17.63, 18.14, 19.97, 20.3,\\
14.84, 11.28, 12.44, 11.22, 10.47, 9.9, 9.25, 10.36, 10.65, 11.17, 12.9, 14.76, 15.3, 14.97, 12.9, 10.52,\\
12.07, 12.87, 12.84, 12.39, 11.63, 11.5, 11.01, 10.85)
\end{gathered}
$$

```
ivory<-sample(Ivory,10,replace=TRUE)
ivory
kenya<-sample(Kenya,10,replace=TRUE)
kenya
distFrechet(Kenya,Ivory,kenya,ivory)
```

   Output:

   Result = 1.40443

   [1]1.40443

B.2 Codes Used in Pricing the Coffee Futures

Using the futures pricing model 3 as deduced in chapter 5, the estimated parameters (generated using CIR model and the Hough transform), and the closing prices gotten from the Ethiopian Coffee Exchange (ECX) and the Ivorian Coffee Exchange (ICE), we get the thesis futures pricing model prices for ECX washed, ECX unwashed and ICE futures respectively as shown in the tables below. The similarities between the curves (ECX Washed, ECX unwashed and ICE futures, with the thesis futures pricing model) is observed through Frechet distance and correlation which are generated using R programming.

B.2.1 ECX Washed

The following shows the analysis of the ECX washed prices with the thesis futures pricing model prices. The analysis methods used include:

1. Correlation with time lags, and
2. Frechet distance.

For both methods, R programming has been used to generate the results.

1. **Correlation**

   R codes:

   Thesis_Futures_Pricing_Model < −c(136.99, 136.80, 134.79, 134.21, 133.48, 129.598, 131.84, 133.87, 130.842, 130.96, 131.08, 131.202, 126.01, 130.44, 129.802, 131.32, 131.44, 131.564, 131.52, 133.25, 132.771, 134.67, 134.75, 134.871, 134.99, 132.49, 128.359, 127.46, 131.15, 134.758, 134.88, 135.00, 134.312, 133.42, 136.51, 141.249, 144.03, 144.16, 144.293, 141.31, 136.64, 139.31, 139.52, 137.63, 137.755, 137.88, 139.82, 139.985, 140.90, 145.11, 145.204, 145.34, 145.47, 148.335, 145.73, 145.57, 147.488, 151.16, 151.30, 151.431, 151.53, 139.42, 137.245, 142.27, 148.34, 148.474, 148.61, 150.84, 150.345, 151.91, 154.32, 156.018, 156.16, 156.30, 157.282, 162.71, 164.76, 160.883, 161.03, 161.17, 161.313, 160.90, 166.28, 165.452, 163.59, 161.61, 161.75, 161.89, 155.75, 156.698, 154.31, 155.90, 155.14, 155.28, 155.42, 155.554, 162.89, 160.27, 158.083)

   ECX_Washed < −c(137.33, 144.81, 155.37, 155.37, 144.73, 154.43, 149.14, 145.12, 153.53, 153.53, 153.53, 157.78, 153.76, 152.49, 146.03, 162.53, 162.53, 162.53, 159.58, 159.58, 169.61, 149.51, 162.98, 162.98, 162.98, 165.34, 160.36, 175.69, 169.72, 158.95, 158.95, 158.95, 162.34, 153.79, 165.13, 167.65, 163.66, 163.66, 163.66, 162.59, 161.5, 160.88, 164.43, 150.63, 150.63, 150.63, 171.19, 169.11, 161.52, 156.95, 160.72, 160.72, 160.72, 160.4, 164.39, 172.25, 156.46, 171.22, 171.22, 171.22, 159.26, 162.52, 172.29, 182.06, 154.2, 154.2, 154.2, 154.2, 150.31, 142.67, 152.14, 148.78, 148.78, 148.78, 144.27, 146.48, 147.91, 145.32, 145.32, 145.32, 145.32, 149.71, 154.93, 143.75, 137.65, 154.35, 154.35, 154.35, 147.79, 147.79, 125.94, 109.14, 128.19, 128.19, 128.19, 128.19, 151.27, 138.99, 132.18)

   ECX_ Washed_ Thesis_ Futures_ Pricing_ Model<-data.frame(ECX_ Washed,Thesis_ Futures_ Pricing_ Model)
   acf<-acf(ECX_ Washed_ Thesis_ Futures_ Pricing_ Model,lag.max=7,plot=F)
   acf
   plot(acf)

   Output:

   Autocorrelations of series ECX_ Washed_ Thesis_ Futures_ Pricing_ Model, by lag
   , , ECX_ Washed

ECX_Washed		Thesis_Futures_Pricing_Model	
1.000	(0)	−0.453	(0)
0.716	(1)	−0.415	(−1)
0.579	(2)	−0.399	(−2)
0.571	(3)	−0.410	(−3)
0.555	(4)	−0.376	(−4)
0.449	(5)	−0.344	(−5)
0.370	(6)	−0.317	(−6)
0.433	(7)	−0.265	(−7)

, , Thesis_ Futures_ Pricing_ Model

ECX_Washed		Thesis_Futures_Pricing_Model	
−0.453	(0)	1.000	(0)
−0.483	(1)	0.962	(1)
−0.476	(2)	0.918	(2)
−0.456	(3)	0.880	(3)
−0.479	(4)	0.855	(4)
−0.503	(5)	0.831	(5)
−0.532	(6)	0.810	(6)
−0.576	(7)	0.790	(7)

2. **Frechet Distance**

   R code:

Thesis_Futures_Pricing_Model $<-$c(136.99, 136.80, 134.79, 134.21, 133.48, 129.598, 131.84, 133.87, 130.842, 130.96, 131.08, 131.202, 126.01, 130.44, 129.802, 131.32, 131.44, 131.564, 131.52, 133.25, 132.771, 134.67, 134.75, 134.871, 134.99, 132.49, 128.359, 127.46, 131.15, 134.758, 134.88, 135.00, 134.312, 133.42, 136.51, 141.249, 144.03, 144.16, 144.293, 141.31, 136.64, 139.31, 139.52, 137.63, 137.755, 137.88, 139.82, 139.985, 140.90, 145.11, 145.204, 145.34, 145.47, 148.335, 145.73, 145.57, 147.488, 151.16, 151.30, 151.431, 151.53, 139.42, 137.245, 142.27, 148.34, 148.474, 148.61, 150.84, 150.345, 151.91, 154.32, 156.018, 156.16, 156.30, 157.282, 162.71, 164.76, 160.883, 161.03, 161.17, 161.313, 160.90, 166.28, 165.452, 163.59, 161.61, 161.75, 161.89, 155.75, 156.698, 154.31, 155.90, 155.14, 155.28, 155.42, 155.554, 162.89, 160.27, 158.083)

ECX_Washed < −c(137.33, 144.81, 155.37, 155.37, 144.73, 154.43, 149.14, 145.12, 153.53, 153.53, 153.53, 157.78, 153.76, 152.49, 146.03, 162.53, 162.53, 162.53, 159.58, 159.58, 169.61, 149.51, 162.98, 162.98, 162.98, 165.34, 160.36, 175.69, 169.72, 158.95, 158.95, 158.95, 162.34, 153.79, 165.13, 167.65, 163.66, 163.66, 163.66, 162.59, 161.5, 160.88, 164.43, 150.63, 150.63, 150.63, 171.19, 169.11, 161.52, 156.95, 160.72, 160.72, 160.72, 160.4, 164.39, 172.25, 156.46, 171.22, 171.22, 171.22, 159.26, 162.52, 172.29, 182.06, 154.2, 154.2, 154.2, 154.2, 150.31, 142.67, 152.14, 148.78, 148.78, 148.78, 144.27, 146.48, 147.91, 145.32, 145.32, 145.32, 145.32, 149.71, 154.93, 143.75, 137.65, 154.35, 154.35, 154.35, 147.79, 147.79, 125.94, 109.14, 128.19, 128.19, 128.19, 128.19, 151.27, 138.99, 132.18)

thesis_ futures_ pricing_ model<-sample(Thesis_ Futures_ Pricing_ Model,10,replace=TRUE)
thesis_ futures_ pricing_ model
ecx_ washed<-sample(ECX_ Washed,10,replace=TRUE)
ecx_ washed
distFrechet(Thesis_ Futures_ Pricing_ Model,ECX_ Washed,thesis_ futures_ pricing_ model,ecx_ washed)

Output:

Result = 35.98278

[1]35.98278

Note: Repeating the procedure many times will always give a different value which will always be below 50.

B.2.2 ECX Unwashed

The following shows the analysis of the ECX unwashed prices with the thesis futures pricing model prices.

1. **Correlation**

   R code:

   Thesis_Futures_Pricing_Model < −c(136.99, 136.80, 134.79, 134.21, 133.48, 129.598, 131.84, 133.87, 130.842, 130.96, 131.08, 131.202, 126.01, 130.44, 129.802, 131.32, 131.44, 131.564, 131.52, 133.25, 132.771, 134.67, 134.75, 134.871, 134.99, 132.49, 128.359, 127.46, 131.15, 134.758, 134.88, 135.00, 134.312, 133.42, 136.51, 141.249, 144.03, 144.16, 144.293, 141.31, 136.64, 139.31, 139.52, 137.63, 137.755, 137.88, 139.82, 139.985, 140.90, 145.11, 145.204, 145.34, 145.47, 148.335, 145.73, 145.57, 147.488, 151.16, 151.30, 151.431, 151.53, 139.42, 137.245, 142.27, 148.34, 148.474, 148.61, 150.84, 150.345, 151.91, 154.32, 156.018, 156.16, 156.30, 157.282, 162.71, 164.76, 160.883, 161.03, 161.17, 161.313, 160.90, 166.28, 165.452, 163.59, 161.61, 161.75, 161.89, 155.75, 156.698, 154.31, 155.90, 155.14, 155.28, 155.42, 155.554, 162.89, 160.27, 158.083)

ECX_Unwashed < −c(128.74, 126.66, 135.16, 135.16, 135.31, 140.43, 130.06, 131.44, 135.28, 135.28, 135.28, 128.23, 132.87, 135.57, 142.68, 128.46, 128.46, 128.46, 129.05, 129.05, 131.74, 128.92, 129.16, 129.16, 129.16, 132.24, 130.86, 129.81, 130.65, 128.75, 128.75, 128.75, 133.13, 134.44, 134.21, 140.78, 135.2, 135.2, 135.2, 138.63, 140.62, 144.1, 143.78, 145.94, 145.94, 145.94, 149.81, 149.84, 148.54, 151.5, 150.9, 150.9, 150.9, 144.77, 145.38, 145.15, 150.1, 146.6, 146.6, 146.6, 145.64, 147.27, 147.52, 145.52, 145.18, 145.18, 145.18, 151, 148.26, 130.31, 147.09, 146.66, 146.66, 146.66, 145.06, 147.31, 147.69, 151.42, 151.42, 151.42, 151.42, 156.22, 154.51, 155.5, 155.65, 148.9, 148.9, 148.9, 145.49, 138.22, 137.38, 140.02, 136.72, 136.72, 136.72, 136.72, 165.36, 141.76, 146)

ECX_ Unwashed_ Thesis_ Futures_ Pricing_ Model<-data.frame(ECX_ Unwashed,Thesis_ Futures_ Pricing_ Model)
acf<-acf(ECX_ Unwashed_ Thesis_ Futures_ Pricing_ Model,lag.max=7,plot=F)
acf
plot(acf)

Output:

Autocorrelations of series ECX_ Unwashed_ Thesis_ Futures_ Pricing_ Model, by lag

, , ECX_ Unwashed

ECX_Unwashed		Thesis_Futures_Pricing_Model	
1.000	(0)	0.708	(0)
0.781	(1)	0.705	(−1)
0.737	(2)	0.702	(−2)
0.699	(3)	0.670	(−3)
0.674	(4)	0.670	(−4)
0.663	(5)	0.668	(−5)
0.614	(6)	0.654	(−6)
0.553	(7)	0.651	(−7)

, , Thesis_ Futures_ Pricing_ Model

ECX_Unwashed		Thesis_Futures_Pricing_Model	
0.708	(0)	1.000	(0)
0.679	(1)	0.962	(1)
0.651	(2)	0.918	(2)
0.638	(3)	0.880	(3)
0.613	(4)	0.855	(4)
0.587	(5)	0.831	(5)
0.565	(6)	0.810	(6)
0.550	(7)	0.790	(7)

2. **Frechet Distance**

   R code:

```
Thesis_Futures_Pricing_Model < −c(136.99, 136.80, 134.79, 134.21, 133.48,
129.598, 131.84, 133.87, 130.842, 130.96, 131.08, 131.202, 126.01, 130.44, 129.802, 131.32, 131.44, 131.564,
131.52, 133.25, 132.771, 134.67, 134.75, 134.871, 134.99, 132.49, 128.359, 127.46, 131.15, 134.758,
134.88, 135.00, 134.312, 133.42, 136.51, 141.249, 144.03, 144.16, 144.293, 141.31, 136.64, 139.31, 139.52,
137.63, 137.755, 137.88, 139.82, 139.985, 140.90, 145.11, 145.204, 145.34, 145.47, 148.335, 145.73,
145.57, 147.488, 151.16, 151.30, 151.431, 151.53, 139.42, 137.245, 142.27, 148.34, 148.474, 148.61, 150.84,
150.345, 151.91, 154.32, 156.018, 156.16, 156.30, 157.282, 162.71, 164.76, 160.883, 161.03, 161.17,
161.313, 160.90, 166.28, 165.452, 163.59, 161.61, 161.75, 161.89, 155.75, 156.698, 154.31, 155.90, 155.14,
155.28, 155.42, 155.554, 162.89, 160.27, 158.083)
```

```
ECX_Unwashed < −c(128.74, 126.66, 135.16, 135.16, 135.31, 140.43, 130.06, 131.44, 135.28, 135.28,
135.28, 128.23, 132.87, 135.57, 142.68, 128.46, 128.46, 128.46, 129.05, 129.05, 131.74, 128.92, 129.16, 129.16, 129.16,
132.24, 130.86, 129.81, 130.65, 128.75, 128.75, 128.75, 133.13, 134.44, 134.21, 140.78, 135.2, 135.2, 135.2, 138.63,
140.62, 144.1, 143.78, 145.94, 145.94, 145.94, 149.81, 149.84, 148.54, 151.5, 150.9, 150.9, 150.9, 144.77, 145.38, 145.15,
150.1, 146.6, 146.6, 146.6, 145.64, 147.27, 147.52, 145.52, 145.18, 145.18, 145.18, 151, 148.26, 130.31, 147.09, 146.66,
146.66, 146.66, 145.06, 147.31, 147.69, 151.42, 151.42, 151.42, 151.42, 156.22, 154.51, 155.5, 155.65, 148.9, 148.9, 148.9,
145.49, 138.22, 137.38, 140.02, 136.72, 136.72, 136.72, 136.72, 165.36, 141.76, 146)
```

```
thesis_ futures_ pricing_ model<-sample(Thesis_ Futures_ Pricing_ Model,10,replace=TRUE)
thesis_ futures_ pricing_ model
ecx_ unwashed<-sample(ECX_ Washed,10,replace=TRUE)
ecx_ unwashed
distFrechet(Thesis_ Futures_ Pricing_ Model,ECX_ Unwashed,thesis_ futures_ pricing_ model,ecx_ unwashed)
```

   Output:

   Result = 43.51305

   [1]43.51305

B.2.3 ICE Futures

The following shows the analysis of the ICE futures prices with the thesis futures pricing model prices.

1. **Correlation**

R codes:

Thesis_Futures_Pricing_Model < −c(136.99, 136.80, 134.79, 134.21, 133.48,
129.598, 131.84, 133.87, 130.842, 130.96, 131.08, 131.202, 126.01, 130.44, 129.802, 131.32, 131.44, 131.564,
131.52, 133.25, 132.771, 134.67, 134.75, 134.871, 134.99, 132.49, 128.359, 127.46, 131.15, 134.758,
134.88, 135.00, 134.312, 133.42, 136.51, 141.249, 144.03, 144.16, 144.293, 141.31, 136.64, 139.31, 139.52,
137.63, 137.755, 137.88, 139.82, 139.985, 140.90, 145.11, 145.204, 145.34, 145.47, 148.335, 145.73,
145.57, 147.488, 151.16, 151.30, 151.431, 151.53, 139.42, 137.245, 142.27, 148.34, 148.474, 148.61, 150.84,
150.345, 151.91, 154.32, 156.018, 156.16, 156.30, 157.282, 162.71, 164.76, 160.883, 161.03, 161.17,
161.313, 160.90, 166.28, 165.452, 163.59, 161.61, 161.75, 161.89, 155.75, 156.698, 154.31, 155.90, 155.14,
155.28, 155.42, 155.554, 162.89, 160.27, 158.083)

ICE_Futures < −c(160.50, 168.75, 168.9, 168.90, 168.10, 163.15, 165.85, 168.25, 164.3, 164.3, 164.3, 164.3,
157.65, 163.05, 162.1, 163.85, 163.85, 163.85, 163.65, 165.65, 164.9, 167.1, 167.05, 167.05, 167.05, 163.8, 158.55,
157.3, 161.7, 166, 166, 166, 165, 163.75, 167.4, 173.05, 176.3, 176.3, 176.3, 172.5, 166.65, 169.75, 169.85, 167.4,
167.4, 167.4, 169.6, 169.65, 170.6, 175.55, 175.5, 175.5, 175.5, 178.8, 175.5, 175.15, 177.3, 181.55, 181.55, 181.55,
181.5, 166.85, 164.1, 169.95, 177.05, 177.05, 177.05, 179.55, 178.8, 180.5, 183.2, 185.05, 185.05, 185.05, 186.05,
192.3, 194.55, 189.8, 189.8, 189.8, 189.8, 189.15, 195.3, 194.15, 191.8, 189.3, 189.3, 189.3, 181.95, 182.9, 179.95,
181.65, 180.6, 180.6, 180.6, 180.6, 188.95, 185.75, 183.05)

ICE_ Futures_ Thesis_ Futures_ Pricing_ Model<-data.frame(ICE_ Futures,Thesis_ Futures_ Pricing_ Model)
acf<-acf(ICE_ Futures_ Thesis_ Futures_ Pricing_ Model,lag.max=7,plot=F)
acf
plot(acf)

Output:

Autocorrelations of series ICE_ Futures_ Thesis_ Futures_ Pricing_ Model, by lag
, , ICE_ Futures

ICE_Futures		Thesis_Futures_Pricing_Model	
1.000	(0)	0.984	(0)
0.929	(1)	0.939	(−1)
0.860	(2)	0.887	(−2)
0.809	(3)	0.843	(−3)
0.780	(4)	0.819	(−4)
0.757	(5)	0.798	(−5)
0.735	(6)	0.778	(−6)
0.714	(7)	0.760	(−7)

, , Thesis_ Futures_ Pricing_ Model

ICE_Futures		Thesis_Futures_Pricing_Model	
0.984	(0)	1.000	(0)
0.931	(1)	0.962	(1)
0.878	(2)	0.918	(2)
0.836	(3)	0.880	(3)
0.809	(4)	0.855	(4)
0.784	(5)	0.831	(5)
0.763	(6)	0.810	(6)
0.743	(7)	0.790	(7)

2. **Frechet Distance**

R code:

$$Thesis\_Futures\_Pricing\_Model < -c(136.99, 136.80, 134.79, 134.21, 133.48, 129.598, 131.84, 133.87, 130.842, 130.96, 131.08, 131.202, 126.01, 130.44, 129.802, 131.32, 131.44, 131.564, 131.52, 133.25, 132.771, 134.67, 134.75, 134.871, 134.99, 132.49, 128.359, 127.46, 131.15, 134.758, 134.88, 135.00, 134.312, 133.42, 136.51, 141.249, 144.03, 144.16, 144.293, 141.31, 136.64, 139.31, 139.52, 137.63, 137.755, 137.88, 139.82, 139.985, 140.90, 145.11, 145.204, 145.34, 145.47, 148.335, 145.73, 145.57, 147.488, 151.16, 151.30, 151.431, 151.53, 139.42, 137.245, 142.27, 148.34, 148.474, 148.61, 150.84, 150.345, 151.91, 154.32, 156.018, 156.16, 156.30, 157.282, 162.71, 164.76, 160.883, 161.03, 161.17, 161.313, 160.90, 166.28, 165.452, 163.59, 161.61, 161.75, 161.89, 155.75, 156.698, 154.31, 155.90, 155.14, 155.28, 155.42, 155.554, 162.89, 160.27, 158.083)$$

$$ICE\_Futures < -c(160.50, 168.75, 168.9, 168.90, 168.10, 163.15, 165.85, 168.25, 164.3, 164.3, 164.3, 164.3, 157.65, 163.05, 162.1, 163.85, 163.85, 163.85, 163.65, 165.65, 164.9, 167.1, 167.05, 167.05, 167.05, 163.8, 158.55, 157.3, 161.7, 166, 166, 166, 165, 163.75, 167.4, 173.05, 176.3, 176.3, 176.3, 172.5, 166.65, 169.75, 169.85, 167.4, 167.4, 167.4, 169.6, 169.65, 170.6, 175.55, 175.5, 175.5, 175.5, 178.8, 175.5, 175.15, 177.3, 181.55, 181.55, 181.55, 181.5, 166.85, 164.1, 169.95, 177.05, 177.05, 177.05, 179.55, 178.8, 180.5, 183.2, 185.05, 185.05, 185.05, 186.05, 192.3, 194.55, 189.8, 189.8, 189.8, 189.8, 189.15, 195.3, 194.15, 191.8, 189.3, 189.3, 189.3, 181.95, 182.9, 179.95, 181.65, 180.6, 180.6, 180.6, 180.6, 188.95, 185.75, 183.05)$$

thesis_ futures_ pricing_ model<-sample(Thesis_ Futures_ Pricing_ Model,10,replace=TRUE)

thesis_ futures_ pricing_ model

ice_ futures¡-sample(ICE_ Futures,10,replace=TRUE)

ice_ futures

distFrechet(Thesis_ Futures_ Pricing_ Model,ICE_ Futures,thesis_ futures_ pricing_ model,ice_ futures)

Output:

Result = 29.37018

[1]29.37018

B.3 Thesis Interpolation Model Curves

Figure 6.1: Thesis Model-Spot Curve

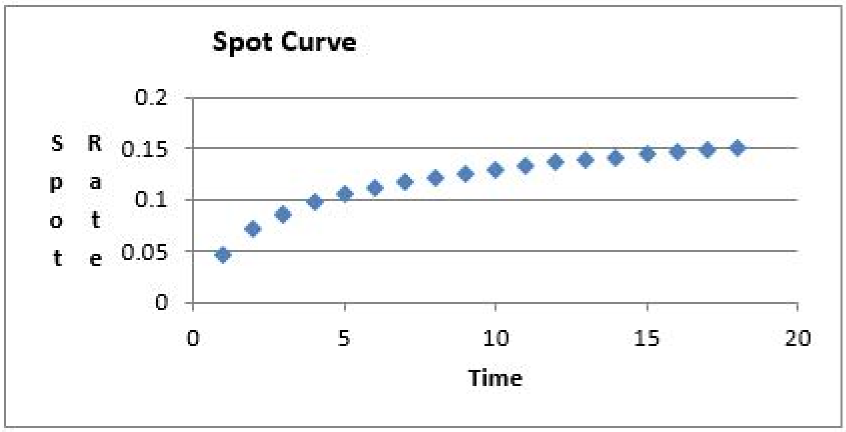
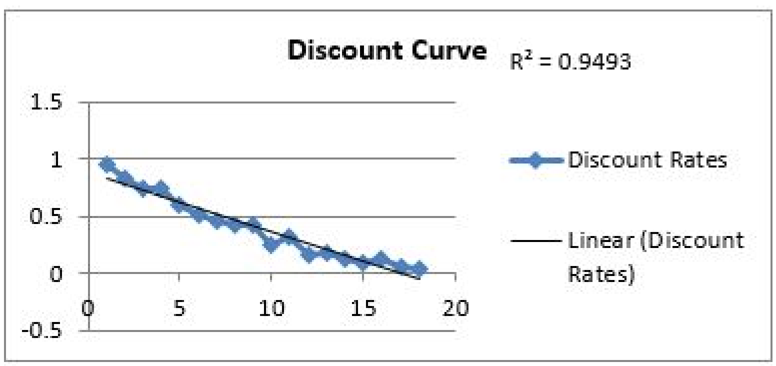


Figure 6.2: Thesis Model- Discount Curve



B.4 Nelson-Siegel Curves

Figure 6.3: Nelson-Siegel- Spot Curve

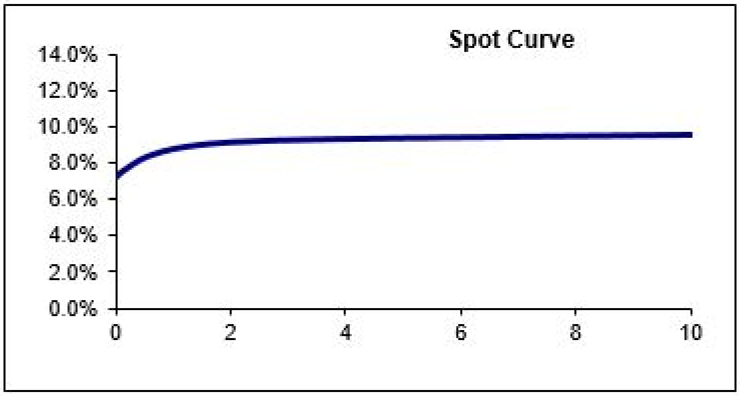
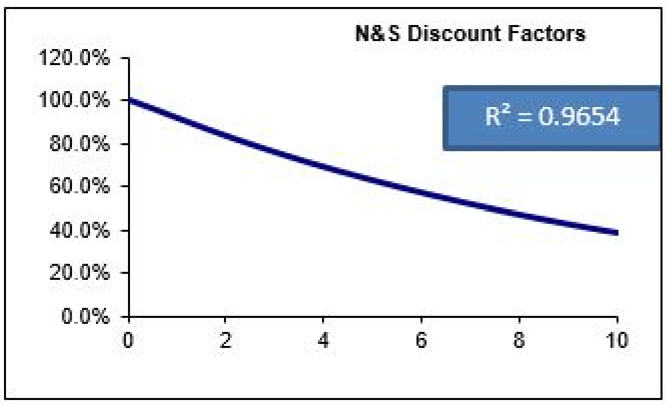


Figure 6.4: Nelson-Siegel- Discount Curve



APPENDIX C : PUBLICATIONS FROM THIS THESIS

1. (Muthoni, L., Onyango, S., & Ongati, N. O. , 2015a) Construction of Nominal Yield Curve for Nairobi Securities Exchange: An Improvement on Monotone Preserving r(t)t Interpolation Method. Journal of Mathematical Finance,5, 377-392. doi: 10.4236/jmf.2015.54032.
   (Presented at Southern Africa Mathematical Sciences Association 33rd Conference, 26th November 2014, Victoria Falls, Zimbabwe)

2. (Muthoni, L., Onyango, S., & Ongati, N. O. , 2015b) Extraction of Zero Coupon Yield Curve for Nairobi Securities Exchange: Finding the Best Parametric Model for East African Securities Markets, Journal of Mathematics and Statistical Science, 4, 2015, 51-74:
   http://www.ss-pub.org/wp-content/uploads/2015/09/JMSS15062201.pdf

(Presented at the 2nd Kenyatta University International Mathematics Conference, June 17th 2015, Nairobi, Kenya)

3. Muthoni, L. (2015) In Search of the Best Zero Coupon Yield Curve for Nairobi Securities Exchange: Interpolation Methods vs. Parametric Models. Journal of Mathematical Finance, 5, 360-376. doi: 10.4236/jmf.2015.54031.

   (Presented at Southern Africa Mathematical Sciences Association 33rd Conference, 25th November 2015, Windhoek Namibia)

4. Muthoni, L. , Onyango, S. and Ongati, O. (2015) Pricing Coffee Futures in a Market with Incomplete Information: A Case of Nairobi Securities Exchange, 12, 79-89. http://www.iosrjournals.org/iosr-jm/papers/Vol12-issue2/Version-1/K012217989.pdf (Presented at the 3rd Strathmore International Mathematics Conference on 6th August 2015, Nairobi, Kenya).

Glossary and Terms

Accumulator array: A histogram showing what estimate of pattern-recognition parameters are supported by different data of a sampled pattern.

American-type option: An option which can be exercised at any time up to the time of maturity of the option, in contrast to a European-type option.

Area: In image processing, this is the number of pixels in the region.

Arbitrage: The search of unlimited profits without accompanying risks (risk free profit) by playing off differences between inefficient markets in the same asset.

Arbitrageur: One who engages in arbitrage Asset: Some item in value, a property or right owned by a company or person.

At-the-money: A characteristic of an option when its strike price is at or near the current market price of the underlying asset.

Basis: the difference between the international coffee price (the futures price) and the local physical market coffee price.

Bid price: The price that a potential buyer is willing to pay for a security.

Binary image: A digital image in which each pixel is assigned (bits or bytes) an intensity value using binary numbers (bits or bytes).

Black-Scholes-Merton formula: An analytical option pricing formula developed to estimate the market value of option contracts, based on geometric Brownian motion.

Bond: A security paying regularly interest (cash bond) or lump sum at maturity.

Broker: A person who buys and sells assets on behalf of other people.

Brownian motion: The basic process in financial mathematics that describes the uncertainty of the market in terms of the cumulative effect of many small adjustments during trade. In the Black-Scholes-Merton model, it is the simplest form of random models for price movement.

Call option: A contract that entitles, but does not oblige the holder to buy a security at/by a future instant of time, the maturity date of the option.

Capital: Amount of money that is invested or used to start a business.

Cash bond: A continuously compounded bond that is appreciated at an instantaneous interest rate set by financial regulators.

Change of measure: A mathematical technique that is used to transform the output of a stochastic process to another variable.

Claim: A payment to be made in the future according to a contract such as an option or a future.

Contingent claim: The value of a financial derivative usually measured by the difference between an agreed strike price and the actual value of the underlying asset at maturity.

Delivery rate: The date when a derivative contract ends with an amount of cash paid in exchange of underlying asset.

Demand: This is a schedule of how much of an asset people will purchase during a specific period of time, depending on price and other factors.

Demand function: This is a mathematical expression of relationship between the quantity demanded and other factors, notably price.

Derivative: A financial instrument whose value is derived from an underlying asset.

Digital image: An image that has been converted into an array of pixels, each of which has an associated intensity value. It may be colored (separate intensities red, green and blue spectral regions), or monochrome (intensity on a single gray scale).

Discounting: Scaling a future price or reward down to make it comparable to the present prices.

Dividend: Portion of profit paid out in cash to the shareholders of a company.

Drift: A steady tendency to show systematic movement in the same direction (up or down).

Efficient market: This is a market in which prices of assets immediately and fully reflect all the relevant market information.

Envelope: Lines surrounding the trend line of a time-varying index or indicator to indicate the possible paths that may be followed.

Equilibrium market: A market in which there is a balance between market demand and supply of assets.

Equilibrium price: A price at which the demand and supply of a traded asset are equal.

European-type option: An option that may be exercised only at time of maturity, in contrast to American-type option.

Excess demand: A disequilibrium condition in a competitive market in which the quantity of an asset demanded is greater than the quantity supplied, hence there is excess demand.

Excess supply: A disequilibrium condition in a competitive market in which the quantity of an asset supplied is greater than the quantity demanded, hence there is excess supply.

Exercise date: A last future date (set in an option contract) by which an option may be exercised.

Exercise price: In an option contract, fixed price at which an asset may be bought or sold.

Exotic option: Financial derivatives based on a combination of assets or standard derivatives. They are over-the-counter contracts designed by individual companies hedging particular risks.

Expiry time: The date when an option contract matures, also known as exercise date.

Feature: In pattern recognition, this is an attribute of a pattern that may contribute to pattern classification-for example, slope of an edge between regions, textures of region, area within a curved edge, etc.

Forward rate: Forward price of instantaneous borrowing, as used in assessing present value of future returns.

Future: A contract with the obligation to sell or buy agreed amounts of assets at fixed price per unit by/at an agreed date.

Geometric Brownian motion: A stochastic process with an output whose logarithms execute simple Brownian motion with linear drift. It is the basic price process in the Black-Scholes Merton model.

Grey level: An intensity value, which is associated with a pixel in a digital image. Usually it represents the darkest pixel intensity by a binary (00000000) and the brightest pixel by one byte of (11111111).

Grey-scale: The range of grey levels that occur in an image, typically 256 if intensities are coded using an 8-bit byte.

Heteroscedasticity: Time-varying, or time-dependent, variance in a time-series of random variables.

Hedge: A trading strategy to protect against risk of loss in a market.

Hedger: A person who buys securities to reduce the investment risk in a portfolio.

Homoscedasticity: Time-independent variance, the opposite of heteroscedasticity

Instantaneous rate: Rate at which interest is paid.

Hough transform: A transformation of pixel coordinates in an image to the parameters of curves on which pixels may be located.

Image: In image processing, a two-dimensional representation of a scene as an array of pixel intensities.

It$\hat{o}$ calculus: A stochastic calculus that gives the rules for integrating functions of Brownian motion. In particular, the It$\hat{o}$ integral is an integral with respect to a variable undergoing Brownian motion and the It$\hat{o}$ lemma is a rule of this calculus.

It$\hat{o}$ lemma: A mathematical technique, which allows stochastic processes to be transformed by nonlinear functions to new processes with changed parameters.

Market equilibrium: A situation in which markets clear because demand for an asset is matched to willing supply.

Marker price of risk: A parameter measuring how much the average investor in an asset is risk seeking or risk averse.

Market value: The spot price obtained through offer and demand from sellers and buyers on the market.

Maturity date: Expiration of a contract.

Noise: Irrelative data that hamper recognition of pattern and the interpretation of data interest.

Numeraire: A basic asset relative to which the value of another asset is determined, for example the cash bond in the Black-Scholes Merton model.

Object: A real thing that can be seen partially or fully in an image.

Occlusion: The partial or complete hiding (loss) of in or more object(s) by another object or by the loss of data.

Optimization: Putting together a portfolio in such a way that return is maximized for a given risk level or risk is minimized for a given expected return level.

Option: A contract that entitles, but does not oblige one to buy/sell something by/at a future date. Options to buy are called calls, option to sell puts.

Pattern: A meaningful regularity that may be used to classify objects in images or trends and spreads in price histories, for example.

Pattern recognition: The analysis, description, identification and classification of objects by automatic and semiautomatic means.

Peak: A point of local maximum in a graph or histogram.

Pixel: The smallest element of a digital image that can be assigned an intensity level.

Portfolio: A collection of securities.

Price adjustment: The movement of supply and demand towards equilibrium at a market price.

Put-call parity: Relationship between the price C at time t of a European call option and the price P at the same time of a European put option, with both having the same strike price E and maturity time T. The expression is $P = Ee^{-rT} + C - S_0$, r is the risk-free interest rate.

Put option: A contract that entitles one to sell something by/at a future maturity date (see also call option).

Security: A contract with financial promise.

Self-financing: A trading strategy that changes the portfolio value only through the price changes of the asset bought sold or held.

Stochastic process: A collection of random variables indexed by time.

Strike price: A price fixed by contract at which an asset may be bought or sold, according to a futures or options agreement.

Supply: This is a schedule of how much of a good or service people supply during a specified period of time, depending on price and other factors.

Supply function: This is a mathematical expression of relationship between the quantity supplied and other factors, notably price.

Tatonnement: A French word, meaning groping, is an iterative process in which buyers and sellers enter into negotiations, discover a price and use this price as a base for re-negations until a price acceptable to all is reached, i.e. no further negotiations.

Tomography: The creation and study of plane image and 3-dimensional renderings by slicing partitioned data.

Transaction cost: A charge for buying or selling an asset.

Underlying: A basic market asset (stock, bond, currency, etc.,) on which other derivative securities can be based upon.

Volatility: The risk, or uncertainty, measure that is associated with short-term fluctuations in a financial time series. It is estimates as a mean square deviation from a trend pattern.

Votes: Entries into a Hough histogram that accumulates evidence for presence of patterns in data.

Window: Part of an image of asset price history.

Yield: The average interest rate returns from an asset, for example interest, a bond, dividend from equity shares etc.

List of Abbreviations

List of Symbols

C	Price of a European Call Option
d_1, d_2	Parameters used in option pricing formula
$D(P)$	Value of demand when price has value P
$dZ(t)$	Differential change in a Wiener diffusion process at time t
E	Exercise price of a derivative contract
$ED(P)$	Excess demand function
$Exp(.)$	Exponential of a variable
$f(.)$	A function of any variable x
f_T	Value of a derivative at time T
f_0	Value of a derivative at time zero, the start of a derivative contract
L	Back shift operator of a discrete time series
$log_e(.)$	Natural logarithm of any variable, log_e
$N(x)$	Cumulative probability that a variable with a standard normal distribution is less than x
$N(0,1)$	Standard normal distribution
$P(t,T)$	Used price of an asset
P^*	Walrasian equilibrium price
P	Walrasian price vector in a multi-asset market
q	Dividend yield rate
$Q_D(P)$	Demand function
$Q_S(P$	Supply function
r	The risk-free interest rate
$S(t)$	Price at time t of an underlying asset on which a derivative is written
$S(T)$	Price of an asset at maturity time T
$S(0)$	Price of an asset at time zero, the start of a derivative contract
t	A future point in time between start at time 0 and maturity at time T
T	Time at maturity of a derivative
$tr(x)$	Trace of a matrix $\mathbf{x}$
$u(i)$	Return provided on the asset between the observation, $S(i-1)$ and an observation $S(i)$

V_{ij}	Vote strength in grey scale histogram
$\mathbf{x}$	Column vector of variables with component $x_1, x_2, , x_i, , x_N$
x'	Transpose of a column vector $\mathbf{x}$, to a row vector
$\mathbf{X}$	A square matrix
X'	Transpose of a matrix X
$Z(t)$	Standard Wiener process having a distribution $N(0, \sqrt{t})$
Δ	Delta coefficient of derivative or a portfolio of derivatives of base assets
$\Delta_x f$	Small change in function f as argument changes from x to $x + \Delta x$
ϵ	Random sample from a standardized normal distribution i.e. $\epsilon \sim N(0,1)$
μ	Expected rate of return on an asset
α, β	Arbitrary constants
$\hat{\mu}$	Estimate of expected rate of return on an asset
θ	Polar coordinate angle between perpendicular to a line segment and x-axis
π	Value of a portfolio of derivatives of base assets
σ	Volatility of an underlying asset
$\hat{\sigma}$	Estimate of volatility of an underlying asset
ρ	Coefficient of statistical correlation
∇	Vector gradient operator. If the operator acts on a function f, then ∇f is its gradient
R^2	Coefficient of determination
χ^2	Chi-square distribution