IOSR Journal of Mathematics (IOSR-JM)
e-ISSN: 2278-5728, p-ISSN: 2319-765X. Volume 12, Issue 2 Ver. I (Mar. - Apr. 2016), PP 79-89
www.iosrjournals.org

Pricing Coffee Futures in a Market with Incomplete Information: A Case of Nairobi Securities Exchange

Lucy Muthoni[1], Prof. Silas Onyango[2], Prof. Omolo Ongati[3]
[1] Assistant Lecturer, Institute of Mathematical Sciences (IMS), Strathmore University P. O. Box: 4877 – 00506, Nairobi, Kenya
[2]Dean, School of Business and Public Management, Public Management, KCA University, P.O Box: 56808-00200, Nairobi, Kenya.
[3]Dean, School of Mathematics and Actuarial Science, J.O.O.U.S.T. University, P.O Box: 210 - 40601 Bondo, Kenya.

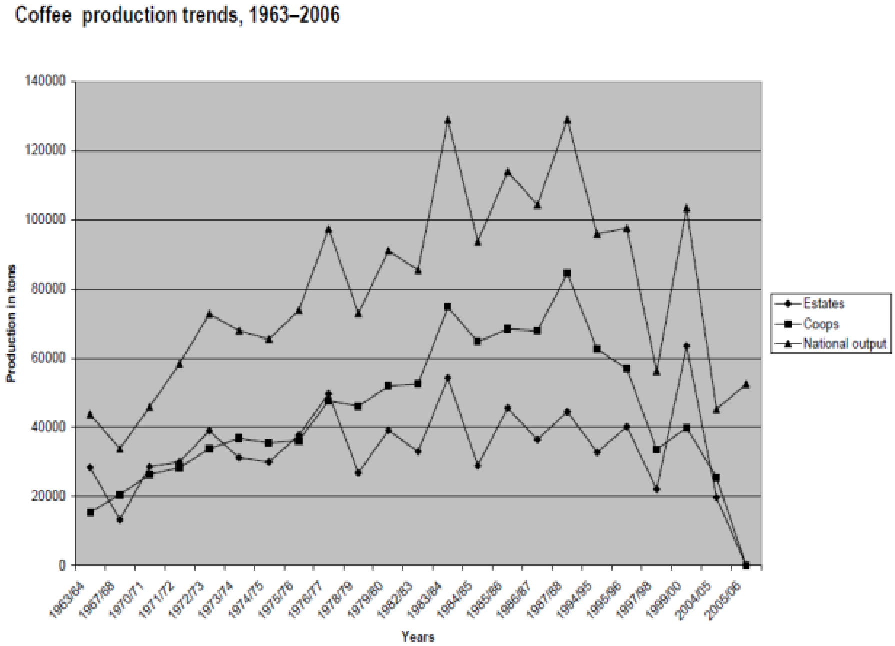
Abstract: *The objective of this paper is to apply Belallah's model in pricing of coffee futures. This caters to the need to price the coffee futures which will be traded at the Nairobi Securities Exchange (NSE) by the end of year 2017, according to a report given by the Capital Market Authority of Kenya in 2013. We apply Belallah's three-factor model. 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. 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. To test the efficiency and consistency of the model, we use the coefficient of determination, root mean squared error and root mean squared error. We also include a liquidity constraint in our pricing which reflects the illiquidity of the test market (Tunisian) and target market, NSE.*

I. Introduction

The coffee sector in Kenya is an important economic activity in terms of income generation, employment creation, foreign exchange earnings and tax revenue. 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 in the race for well-paying jobs, sinecures, and contracts in the various institutions that serve as gravy trains in the coffee sector cash cow.Kenya produces some of the best coffee in the world. Being the more flavorful Coffee Arabica rather than Coffee Canephora (Robusta), the "fully washed mild" belongs to the top quality group called "Colombian milds". This is attributed to the well-distributed rainfall; high altitude (1,500–2,000 meters above sea level) and therefore moderate temperatures (averaging 20° centigrade Celsius), with characteristically high equatorial ultraviolet sunlight diffusing through thick clouds; and deep red volcanic soils.

In Kenya, coffee is grown in the highland districts of Kenya: Kiambu, Muranga, Nyeri, Thika, Kirinyaga, Meru North, Meru Central, Meru South, Embu, Machakos, Kitui, Nakuru, West Pokot, Kajiado, Baringo, Kericho, Nandi, Laikipia, Transnzoia, Uasin-Gishu, Keiyo, Marakwet, Kajiado, Bungoma, Kakamega, Busia, Kisii, Siaya, Kisumu, South Nyanza, and Taita. The high production zone is a triangle formed by Mt. Kenya, the Aberdare Range and Machakos Town (see the coffee map of Kenya in appendix figure 1).Coffee producing areas contain about 45 per cent of Kenya's population, estimated at 36.4 million. Since some of these people are as much as 40 per cent income-dependent on coffee, their lives revolve around the fate of coffee. Kenya coffee sector is composed of two categories of farms: the plantation sub-sector comprising of about 3,300 farms of Which 300 are greater than 25 hectares; and the cooperative sub-sector of some 523 cooperative unions with about 700,000 smallholders cultivating about 120,000 hectares of coffee, equivalent to about 0.2 hectares apiece. It is estimated that a total of 170,000 hectare are under coffee and that 75 per cent of that total is organized around smallholder cooperatives.Kenya coffee production increased rapidly in ripples in the two decades after independence. As shown in figure 1 below; total production for both estates and cooperative sub-sectors rose from 43,778 tons in 1963–64 to 128,941 tons in1983–84. Since then, however, the coffee industry has been on a downward trend except for a brief spell in 1999–2000. As a result, coffee's contribution to incomes, employment creation and foreign exchange earnings has declined.

Figure 1:



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 Trader's Association, No. 2/2006, p.9.

Table 1: tonnage and value of coffee marketed and average gross prices 2000-2005

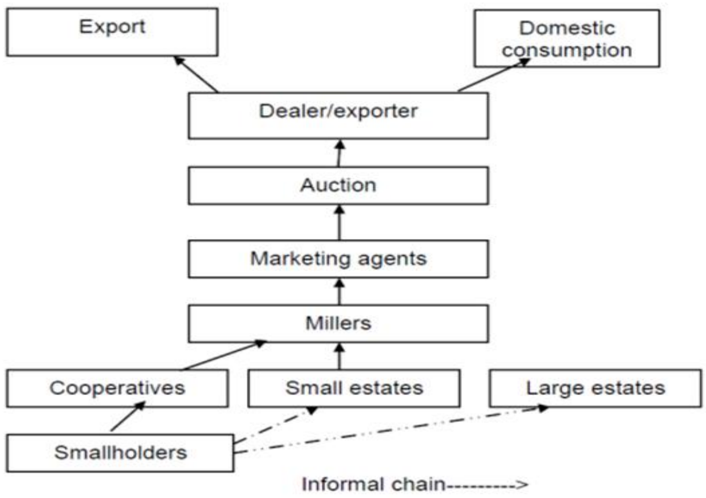
Year	Coffee sold (tons)	Value of sales (thousand KES)	Average gross farm prices (KES/kg)
2000	98000	1128200	11.51
2001	54600	642420	11.76
2002	45500	544110	11.96
2003	61200	595670	9.73
2004	49900	728450	14.60
2005	45000	680000	15.11

So why has coffee production not picked up as a result of the price increase? Could it be that the farmers feel discouraged by production costs and marketing constraints which show up at the farm-gate in the net price received and the domestic terms of trade? This would lead to neglect of the coffee bushes and substitution of the crop with other economic activities in some areas.

The value chain in coffee production involves the following steps:

i. Nursery operations to produce seedlings;
ii. farm-level operations (planting, weeding, fertilizing, pruning, spraying, picking/harvesting of red cherry);
iii. Transportation of cherries to the pulpier/coffee factory
iv. coffee factory primary processing: pulping, fermenting, washing and drying to produce parchment coffee, either at a cooperative facility or in a farm-based pulpery;
v. curing operations (removing parchment/peeling, cleaning and polishing the beans to produce green coffee beans), by a miller;
vi. milling plant operations: hulling, cleaning/polishing, sorting, grading, bagging, e.g. by Kenya Planters Cooperative Union (KPCU) and Thika Coffee Mills;
vii. Auctioning at the Nairobi Coffee Exchange (NCE) where dealers, roasters, marketers and exporters buy various grades of green coffee;
viii. Roasting, grinding, blending and packing/packaging by roasters and marketing agents, e.g. C. Dormans and Nairobi Java House. Can be done locally or in the importing country;
ix. Marketing and selling: locally, regionally, globally – packed or even in bulk – by dealers, roasters, marketers and exporters to supply coffee to consumers.

Figure 2. The coffee value chain



Source: *Final report on assessment of the value-adding opportunities in the Kenyan coffee industry*, European Commission, April 2004.

Once coffee is exported, it is traded in the international markets; it is in these markets that the local prices of coffee are determined. There are two markets in which coffee is traded: the physical market and the financial market, which sometimes is known as the futures market. The Physical market involves traders buying coffee from producers, and exporting this coffee to buyers. At each step in the process, physical coffee moves between market participants and payments are made for the physical coffee. The producers earn money by selling the coffee that they have grown. The trader (or intermediary / exporter) earns money by buying the coffee from the producers, processing the coffee and selling it to a buyer (importer). Often there will be a variety of intermediaries through whom the coffee passes, including primary and secondary cooperatives, processors, exporters, etcetera. However the physical market will always involve the trade in physical volumes of coffee with the ultimate purpose to deliver coffee from the producer who grows the coffee to the consumer who drinks the coffee. Each participant in the physical coffee supply chain aims to make money from the trade in the physical coffee. Futures Markets are very different from the physical market both in how they function and their purpose. Unlike the physical market, in the financial market contracts will only result in physical delivery of coffee on a limited basis. Rather these contracts are held for financial purposes and the contracts will be sold or terminated at, or prior to, the latest delivery date which will result in a financial settlement of the contract. Because such contracts are offset (buy and against sell and vice versa) they are often referred to a ‘paper contracts’. The vast majority of contracts traded on the exchange are traded as a means of providing buyers and sellers of coffee (in the physical market) with opportunities to manage their exposure to price risk. These markets are accessed by participants from all over the world which results in an extremely large number of transactions every day. An added benefit of this liquidity is that 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 as all those in the coffee business know is used by producers, traders, exporters, and roasters around the world as the reference price for coffee on any given day.

Comparing these characteristic of the physical market to those in the financial markets shows the characteristics main differences between the two markets:

- **Location/ Place** – 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), and not physical coffee, is exchanged.
- **Activities / Purposes** – in the physical market, the primary activities is the buying and selling of green coffee between businesses that earn money from trading and moving coffee. The financial markets have a very different purpose and set of activities. In the financial markets, coffee contracts are traded with very little expectation of delivery of coffee. The primary purpose of these activities is for coffee sector participants to get information on current (and future) prices for coffee, and to enable them to manage their risk through the trading of these financial instruments.
- **Delivery Location** – in local physical coffee markets, delivery is usually affected by shipment from the port where the coffee is exported from. 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 FOB (Free on Board) export terms whereas the financial market contracts are priced 'in store' (also called ex dock); the coffee futures are priced basis 'delivered licensed warehouse' meaning that the coffee is presumed to have been shipped from origin and discharged into the licensed warehouse.
- **Unit of Measurement** – Local markets utilize their own units of measurement (for example Kilograms for East Africa / Quintales for Central America). But the New York Arabica futures markets works in pounds and the London Robusta futures market works in tones.

Each of these differences will affect the basis (the differential between local prices for physical coffee and international prices), as they each involve different costs. Often international coffee contract (traded on the commodity exchanges) will have higher prices (for the same type of coffee) than local contracts. This is primarily due to the differences in delivery conditions

Basis refers to the **difference between the international coffee price (the futures price) and the local physical market coffee price**. Many traders know this more commonly as the "differential" between the two markets. Hence some speak of "basis risk", others of 'differential risk' – the two mean the same.

This basis between these two markets is determined by a number of different factors. There are a variety of items that can make up this difference. These include coffee quality differences between one country and another, the costs of transportation, interest and insurance.

Basis can be either positive or negative. Positive basis is when the local market price for coffee expressed in FOB terms is greater than the international market price. When considering basis a simple way to think of it is:

Local coffee price FOB Price of Coffee – Future Price = Basis

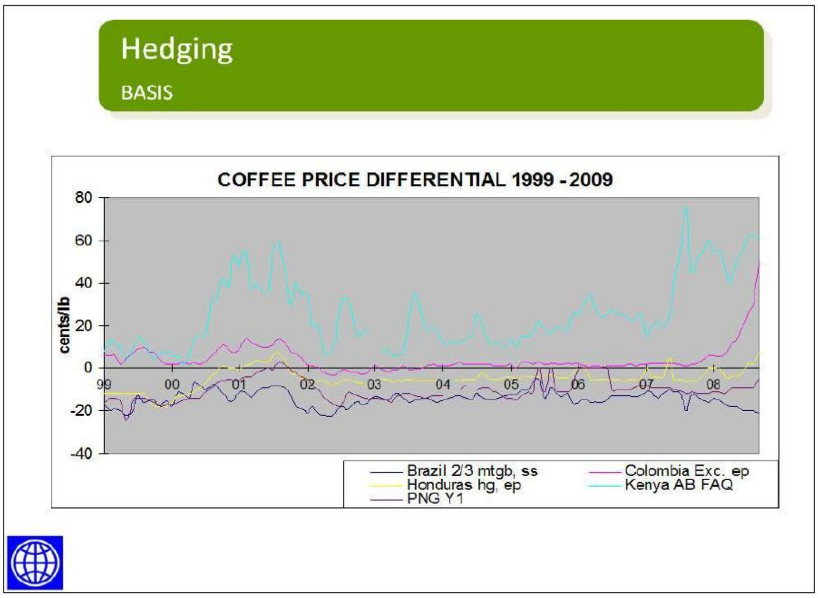
The following figure shows the coffee price differential for five coffee exporting countries.

II. Objective Functions

The objectives of this paper are to:

i. Give relevant information about coffee in Kenya and trading of coffee
ii. Find a suitable model for pricing coffee futures in Kenya's Nairobi Securities Exchange
iii. Test the model in terms of efficiency and consistency and give recommendations

Figure 3: Coffee Price Differential



III. Literature Survey

The stochastic behavior of commodity prices plays a critical role in the pricing of commodity derivatives and in capital budgeting decisions. The traditional approach for the valuation of the investment projects is the net present value approach, which hugely affected budgeting decisions. An alternative to the traditional approach is the certainty-equivalent approach which avoids the computation of a risk-adjusted discount factor using instead the relevant risk-free rate of interest. Earlier studies are based on constant interest rates and convenience yields in the pricing of financial and real commodity derivatives. This assumption implies that the distribution of future spot prices has a variance that increases without bound as the horizon increases. Brennan and Schwartz (1985) apply the option pricing theory for value investment projects in natural resources where the spot price of the commodity follows a geometric Brownian motion. The option pricing theory uses the information contained in the futures prices since the prices are used in the estimation of the convenience yield. The approach is based on the use of the risk free rate than a risk-adjusted discount rate and allows for managerial flexibility in the form of options.

Many models consider relations between prices of futures contracts and corresponding spot prices, e.g. Anderson (1983), Hischileifer (1989) and (1990). We also see a textbook by Duffie (1989) trying to explain the relationship between the prices of futures contracts and corresponding prices, but applying the concept on pricing of sugar. Schwartz (1997) compared three models of stochastic behavior of commodity prices: a one-factor model and three-factor model. Schwartz (1998) developed a one-factor model that preserves the main characteristics of two-factor models. In this paper, we define the costs as in Black Sholes Merton (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 (1995), and (1999b). We use an extension of the analysis in the Schwartz (1997) and Schwartz (1998) to account for the effects of incomplete information as it appears in the models of Merton (1987) and Bellalah (2001). This paper also uses 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.

IV. Empirical Methodology

Model Selection

4.1.1 The Interest Rate Model

For the interest rate model, we use the Cox Ingersoll Ross (CIR) model which was introduced in 1985 by John C. Cox, Jonathan E. Ingersoll and Stephen A. Ross, as an extension of the Vasicek Model. The CIR bond pricing model assumes that the risk natural process for the short interest rate is given by:

$$dr_t = a\,(b - r_t\,)dt \;+ \sigma\sqrt{r_t}dW_t \qquad (4.1)$$

$a\,(b\, -\, r\,)$ is the same as in the Vasicek Model. The standard deviation factor $\sigma\sqrt{r_t}$, avoids the possibility of negative or null rates if the condition $2a \succ b\sigma^2$ is met.

Taking the integral of equation (4.1) above, we have:

$$r_t - r_u = a\,\int_u^t\,(\,b\, -\, r\,_s)ds + \sigma\int_u^t\sqrt{r_s}\,dW_s, \qquad (4.2)$$

Defining: $f\,(x_t\,)\, =\, x_t{}^2$ and $x_t\, =\, r_t$, and finally applying Itô's lemma, equation (4.2) becomes:

$$r_t{}^2 = r_u^2 + \int_u^t 2r_s\left[a(b - r_s)ds + \sigma\sqrt{r_s}dW_s\right] + \frac{1}{2}\int_u^t 2\sigma^2 r_s ds,$$

$$= r_u^2 + \int_u^t 2r_s\left[abds - ar_s ds + \sigma\sqrt{r_s}dW_s\right] + \int_u^t \sigma^2 r_s ds,$$

$$= r_u^2 + \int_u^t 2r_s abds - 2ar_s^2 ds + 2\sigma\sqrt[3]{r_s}dW_s + \int_u^t \sigma^2 r_s ds,$$

$$= r_u^2 + \int_u^t (2r_s ab - 2ar_s^2 + \sigma^2 r_s)ds + \int_u^t 2\sigma\sqrt[3]{r_s}dW_s,$$

$$r_t = r_u^2(2ab + \sigma^2)\int_u^t r_s ds - 2a\int_u^t r_s^2 + 2\sigma\int_u^t \sqrt[3]{r_s}\,dW_s \qquad (4.3)$$

If $\mu\, =\, 0$ then we have:

$$r_t = r_0 + a\int_0^t (b - r_s)ds + \sigma\int_0^t \sqrt{r_s}dW_s \qquad (4.4)$$

Thus, the unconditional mean is:

$$E^Q\;(r_t\,) = r_0 + a\left(bt - \int_0^t E^Q\,(r_s)ds\right) \qquad (4.5)$$

Solving the equation $\Phi(t)\, =\, r_0\, +\, a\, \left(bt\, - \int_0^t \Phi(s)ds\right)$ which can be transformed into the ODE (ordinary differential equation) $\Phi(t)' +\, a\Phi\,(\,t\,)\, =\, ab$.

So, the unconditional mean becomes:

$E^Q\,(r_t\,)\, =\, b\, +\, (\,r_0\, -\, b\,)e^{-at}$ which can also be written as:

$$\{E^Q r_t|F_u\} = r_u e^{-a(t-u)} + b\left(1 - e^{-a(t-u)}\right) \qquad (4.6)$$

Similarly, we can write equation (4.6) as:

$$E^Q(r_t^2) = r_0^2(\,2ab\, +\, \sigma^2\,)\int_0^t\, E^Q(r_s^2)ds - 2a\int_0^t\, E^Q(r_s^2)ds, \qquad (4.7)$$

Substituting the value of $E^Q(r_t)$ into (4.3) and applying second derivative, then we can get can get variance:

$$Var(r_t) = \frac{\sigma^2}{a}(1 - e^{-bt})\left[r_0 e^{-bt} + \frac{b}{2}(1 - e^{-bt})\right], \qquad (4.8)$$

or

$$Var\{r_t|F_u\} = \frac{r_u\left(\sigma^2\left(e^{-a(t-u)} - e^{-2a(t-u)}\right)\right)}{a} + \frac{b\sigma^2\left(1 - e^{-a(t-u)}\right)^2}{2a} \qquad (4.9)$$

The instantaneous short rate dynamics corresponds to a continuous time first-order autoregressive process where the randomly moving interest rate is elastically pulled toward a central location or long term

value, b, which leads to mean reversion property. If $\sigma^2 > 2ab$ then r_t can be zero. If $\sigma^2 \le 2ab$ then the upward drift is sufficiently large to make the original inaccessible.
So from (4.8) equation we get that there is no explicit form for the solution to the CIR model. It is known that the model has unique positive solution. If the interest rate reaches zero then it can subsequently become positive. More generally, when the rate is low (close to zero), then the standard deviation also becomes close to zero.

4.1.2 The Pricing Model

This section presents the model of commodity prices and the formulas for futures contracts. These models are closed form solutions for futures prices.

4.1.2.1 Model 1

In this model, Schwartz (1997) assumed that the commodity spot price follows the stochastic process:

$$dS = \kappa(\mu - \ln S)Sdt + \sigma S dz \qquad (4.10)$$

Where dz is an increment to a standard Brownian motion and κ refers to the speed of adjustment; when $X = \ln S$, and applying lto's lemma to characterize the log price by an Ornstein-Uhlenbeck stochastic process, we have:

$$dX = \kappa(\alpha - X)dt + \sigma dz \qquad (4.11)$$

With

$$\alpha = \mu - \frac{\sigma^2}{2\kappa} \qquad (4.12)$$

Where κ measures the degree of mean reversion to the long run mean log price α. Under standard assumptions, Schwartz (1997) gives the following dynamics of the Ornstein-Uhlenbeck stochastic process under the equivalent martingale measure

$$dX = \kappa(\alpha^* - X)dt + \sigma dz^* \qquad (4.13)$$

Where $\alpha^* = \alpha - \lambda$ where λ is the market price of risk; from equation (4.13), the conditional distribution of X at time T under the equaivalent martingale measure is normal. The mean and variance of X is:

$$E_0[X(T)] = e^{-\kappa T}X(0) + (1 - e^{\kappa T})\alpha^*$$

$$Var[X(T)] = \frac{\sigma^2}{2\kappa}(1 - e^{-2\kappa T}) \qquad (4.14)$$

When the interest rate is constant, the futures or the forward price of commodity corresponds to the expected price of the commodity for the maturity T. Using the properties of the long-normal distribution, the futures or the forward price given by:

$$F(S,T) = E[S(T)] = exp\left(E_0[X(T)] + \frac{1}{2}Var_0[X(T)]\right) \qquad (4.15)$$

And

$$F(S,T) = \exp\left(e^{-\kappa T}\ln S + (1 - e^{-\kappa T})\alpha^* + \frac{\sigma^2}{4\kappa}(1 - e^{-2\kappa T})\right) \qquad (4.16)$$

This equation can be written in a log form as

$$\ln F(S,T) = e^{-\kappa T}\ln S + (1 - e^{-\kappa T})\,\alpha^* + \frac{\sigma^2}{4\kappa}(1 - e^{-2\kappa T}) \qquad (4.17)$$

Equation (4.16) is solution to the partial differential equation

$$.\frac{1}{2}\sigma^2 S^2 F_{SS} + (\kappa(\mu - \lambda) - \ln S + 1 + (1 - e^{\kappa}T)\alpha * + \frac{\sigma^2}{4k}(1 - e^{-2\kappa T}) \qquad (4.18)$$

Under the terminal boundary condition $F(S,0) = S$

4.1.2.2 Model 2

In this two factor model, the first factor corresponds to the spot price of the commodity with the following dynamics.

$$dS = (\mu - \delta)Sdt + \sigma_1 Sdz_1 \qquad (4.19)$$

Where δ 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.
The second factor corresponds to the convenience yield with the following dynamics

$$d\delta = \kappa(\alpha - \delta)dt + \sigma_2 dz_2 \qquad (4.20)$$

Where

$$dz_1 dz_2 = \rho dt \qquad (4.21)$$

Hence equation (4.19) allows for stochastic convenience yield, which follows and Ornstein-Uhlenbeck stochastic process. When δ is the deterministic function of $S, \delta(S) = \kappa \ln S$, this model reduces to model 1. When $X = ln\, S$, applying lto's Lemma allows to characterize the log price as:

$$dX = \left(\mu - \delta - \frac{1}{2}\sigma_1^2\right)dt + \sigma_1 dz_1 \quad (4.22)$$

The commodity is viewed as an asset paying a stochastic dividend yield δ and the risk adjusted drift of the commodity is $(r + \lambda_S - \delta)$ where λ_S refers to an information cost for the asset S. In fact, it can be shown as in Bellalah (2001) that under the equivalent martingale measure, the stochastic processes for the two factors can be written as

$$dS = (r + \lambda_S - \delta) - \lambda]dt + \sigma_1 dz_1^* \quad (4.23)$$

$$d\delta = [\kappa(\alpha - \delta) - \lambda]dt + \sigma_2 dz_2^* \quad (4.24)$$

$$dz_1^* dz_1^* = \rho dt \quad (4.25)$$

Where λ refers in this model to the market price of convenience yield risk. Using the same approach as in Bellalah (2001), Futures prices satisfy the following PDE

$$\frac{1}{2}\sigma_1^2 S^2 F_{SS} + \frac{1}{2}\sigma_2^2 F_{\delta\delta} + \sigma_1\sigma_2\rho S F_{S\delta} + (r + \lambda_S - \delta)SF_S + [\kappa(\alpha - \delta) - \lambda]F_\delta - F_T = 0 \quad (4.26)$$

Under the terminal boundary condition $F(S, \delta, 0) = S$; As in Schwartz (1997), the solution is given by:

$$F(S, \delta, T) = S\, exp\left[-\delta \frac{1-e^{-\kappa T}}{\kappa} + A(T)\right] \quad (4.27)$$

This can be written in the log form as:

$$ln\, F(S, \delta, T) = ln\, S - \delta \frac{1-e^{-\kappa T}}{\kappa} + A(T) \quad (4.28)$$

Where:

$$A(T) = \left(r + \lambda_S - \hat{\alpha} + \frac{1}{2}\frac{\sigma_2^2}{\kappa^2} - \frac{\sigma_1\sigma_2\rho}{\kappa}\right)T + \frac{1}{4}\sigma_2^2 \frac{1-e^{-2\kappa T}}{\kappa^3} + (\hat{\alpha}\kappa + \sigma_1\sigma_2\rho - \frac{\sigma_2^2}{\kappa})\frac{1-e^{-2\kappa T}}{\kappa^2}$$

$$\hat{\alpha} = \alpha - \frac{\lambda}{\kappa} \quad (4.29)$$

The main difference between this solution and that in Schwartz (1997) concerns the discount rate in $A(T)$ which appears to be the interest rate plus the information cost on the asset S rather than that interest rate only.

4.1.2.3 Model 3

In this three-factor model, 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 Vasicek (1997), using equations (4.23) and (4.24), the joint stochastic process for the three factors under the equivalent martingale measure can be written as

$$dS = (r + \lambda_S - \delta)Sdt + \sigma_1 S dz_1^* \quad (4.30)$$

$$d\delta = \kappa(\hat{\alpha} - \delta)dt + \sigma_2 dz_2^* \quad (4.31)$$

$$dr = a(m^* - r)dt + \sigma_3 dz_3^* \quad (4.32)$$

Where $dz_1^* dz_2^* = \rho_1 dt, dz_2^* dz_3^* = \rho_2 dt, dz_1^* dz_3^* = \rho_3 dt$ (4.33)

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.

$$\frac{1}{2}\sigma_1^2 S^2 F_{SS} + \frac{1}{2}\sigma_2^2 F_{\delta\delta} + \frac{1}{2}\sigma_3^2 F_{rr} + \sigma_1\sigma_2\rho_1 S F_{S\delta} + \sigma_2\sigma_3\rho_2 F_{\delta r} + \sigma_1\sigma_3\rho_3 S F_{Sr} + (r + \lambda_S -$$
δSFS+κα−δFδ+a(m∗−r)Fr−FT=0 (4.34)

Under the terminal boundary condition $F(S, \delta, r, 0) = S$; Following the analysis in Schwartz (1997), the solution is given by:

$$F(S, \delta, r, T) = S\, \exp\left[\frac{-\delta(1-e^{-\kappa T})}{\kappa} + \frac{(r+\lambda_S)(1-e^{-\alpha T})}{\alpha} + C(T)\right] \quad (4.35)$$

This can be written in a log form as

$$ln(S, \delta, r, T) = ln\, S - \frac{\delta(1-e^{-kT})}{\kappa} + \frac{(r+\lambda_S)(1-e^{-kT})}{\alpha}C(T)] \quad (4.36)$$

Where

$$C(T) = \frac{(\kappa\hat{\alpha}+\sigma_1\sigma_2\rho_1)(1-e^{-\kappa T})-\kappa T]}{\kappa^2} - \frac{\sigma_2^2(4(1-e^{-\kappa T})-(1-e^{-2\kappa T})-2\kappa T}{4\kappa^3} - \frac{(am^*+\sigma_1\sigma_3\rho_3)[(1-e^{-aT})-aT]}{a^2} - \frac{\sigma_3^2(4(1-e^{-aT})-(1-e^{-2aT})-2aT}{4a^3} + \sigma_2\sigma_3\rho_2\frac{(1-e^{-\kappa T})+(1-e^{-aT})-(1-e^{-(\kappa+a)T})}{\kappa\alpha(\kappa+a)} + \frac{\kappa^2(1-e^{-aT})+a^2(1-e^{-kT})-\kappa a^2 T-a\kappa^2 T)}{\kappa^2 a^2(\kappa+a)} \quad (4.37)$$

V. Liquidity-Weighting and Optimization Functions

Most African securities markets are illiquid. Subramanian (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:

$$W_i = \frac{\left[\left(1-e^{-\frac{v_i}{v_{max}}}\right)+\left(1-e^{-\frac{n_i}{n_{max}}}\right)\right]}{\sum_i W_i} = \left(1-e^{-\frac{v_i}{v_{max}}}\right)+\left(1-e^{-\frac{n_i}{n_{max}}}\right) \qquad (4.38)$$

$$W_i = \frac{\left[tanh\left(-\frac{v_i}{v_{max}}\right)+tanh\left(-\frac{n_i}{n_{max}}\right)\right]}{\sum_i W_i} = tanh\left(-\frac{v_i}{v_{max}}\right)+tanh\left(-\frac{n_i}{n_{max}}\right) \qquad (4.39)$$

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 (4.38) and (4.39) 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:

$$Min \sum_{i=1}^{n} w_i (P_i - B_i)^2 = Min \sum_{i=1}^{n} w_i \varepsilon_i^2 = 0 \qquad (4.40)$$

Where P_i is the observed futures price and B_i the model-generated future's price

VI. 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:

The Coefficient of Determination (R^2)

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(P_i - \widehat{B_i})^2/(n-k)}{\sum_{i=1}^{n}(P_i - \bar{P_i})^2/(n-1)} \qquad (4.41)$$

Where $\bar{P}$ is the mean average price of all observed futures prices, $\widehat{B_i}$ is the model price of a futures i, n the number of futures 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 observed prices and associate a low R^2 with a poor fit.

Root Mean Squared Error (RMSE)

Denoted as the RMSE, a low value for this measure is assumed to indicate that the model is flexible, on average, and is able to fit the observed prices.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(P_i - \widehat{B_i}\right)^2} \qquad (4.42)$$

Root Mean Squared Percentage Error (RMSPE)

Denoted as the RMSPE, a low value for this measure is also assumed to indicate that the model is flexible, on average, and is able to fit the observed futures prices.

$$RMSPE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\frac{P_i - \widehat{B_i}}{P_i}\right)^2} * 100\% \qquad (4.43)$$

VII. Empirical Results

Data

The data used in this study is from Tunisian market the source being Bloomberg. This is because NSE has not yet started trading in the futures.Data and information to enable analyses of the problems were gathered mostly from secondary sources, mainly the publications of the Coffee Board of Kenya (CBK), the Central Bureau of Statistics and the government printer. Some data were collected from brief field visits and the author's knowledge of the sector.

VIII. Calibration results

To calibrate parameter values in the above models, we use both the L-BGFS-B and Gaussian approximation, which is made possible by noting that transition densities solve parabolic PDEs (the Kolmongorov equation). The latter is very useful when it comes to calibrating and approximating the terms in the CIR model, while the former, with combination of numerical methods make it possible to calibrate the other parameters in the equations.It is also possible to use a double grid search routine to estimate the state variables S and δ, which minimize the squared deviation between model and market prices. The terms F1, F2… correspond to futures contracts with different maturities. The calibrated values for the parameters are as follows (with the standard errors indicated in the brackets):

Table 2: Calibrated Parameters values

Period	1/15/99-5/16/09
Contracts	F1, F3, F5, F7, F9
Number of observations	347
μ	0.326 (0.0110)
κ	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)

Table 3: Test Statistics Results

	RMSPE	RMSE	R^2
mean	0.0131	1.4914	0.9693
Std.dev	0.0061	0.7299	0.0368

IX. Results and Discussion

From the table 3 above, we see that the coefficient of determination is greater than 95% which is an indication of a good fit of observed prices to the model. The model applied in this study was model 3, which is a three factor model.We also see that the root mean squared percentage error is lower than 2%, another indication of model's attractiveness and fitness to the observed data.For all the three test statistics, we find that the standard deviation is less than 1.0, which is an indication that the average observed prices do not deviate much from the model's generated prices.In conclusion, we feel that Bellalah's three factor model is good enough to be applied in markets with incomplete information, such as Nairobi Securities Exchange.

References

[1]. A. R. Conn, N. G. (1988). Global covergence of a class of trust region algorithms for optimization with simple bounds, . SIAM J. Numer. Anal., 25(2): , 433-460.
[2]. Anderson, R. W. (1983). Hedger Diversity in Futures Markets. The Economic Journal , 370-389.
[3]. Bellalah M. (2001). A reexamination of Corporate Risks Under Incomplete Inforamtion . International Journal of Finance and Economics .
[4]. Bellalah, M. a. (1995). Option Valuation with Infromation Costs:Theory and Tests. Financial Review , 617-635.
[5]. Bellalah, M. (1999b). Les biais des modeles d'options revites. Revenue Franaise de Gestion, Juin , 94-100.
[6]. Broyden., C. (1970). The convergence of a class of double-rank minimization algorithms. II.The New algorithm. . J. Inst. Math. Appl.,6 , 222-231.
[7]. Ciyou zhu, R. B.-B.-B. (2011). www.ece.northwestern.edu. Retrieved from http://www.ece.northwestern.edu/nocedal/lbfgsb.html,.
[8]. Deventer, A. a. (1994). Fitting Yield Curves and Forward Rate Curves with Maximum Smoothness. Journal of Fixed Income , 52-62.
[9]. Duffie, D. (1989). Futures Markets. New Jersey: Prentice hall Hall Inc.
[10]. Fletcher., R. (1970). A new approach to viable metric algorithms. The computer Journals, 13(3) , 317-322.
[11]. Goldfarb, D. (1970). A family of variable metricmethods derived by variational means. math comp, 24: , 23-26.
[12]. Hirschleifer, D. (1990). Hedging Pressure and Futures in a General Equilibrium Model. Econometrica, 58, 2, , 411-428.
[13]. Hischileifer, D. (1989). Determinants of Hedgig and Risk premia in Commodity Futures Markets . Journal of Financial and Quantitative Analysis, 24 , 313-331.
[14]. Merton, R. (1987). A simple Model of Capital Risks Under Incomplete Information. Journal of Finance 42 , 483-510.
[15]. Nelson, C. a. (1987). Parsimonius Mofelling of Yield curves. The Journal of Business , 473-489.
[16]. Nocedal., J. (1980). Updating quasi-Newton matrices with limited storage. Math Comp., 35(151): , 773-782.
[17]. Overton., A. S. (2013). Nonsmooth optimization via quasi-Newton methods. . Math. Program., 141(1-2, Ser. A): , 135-163 .
[18]. Richard, H. B. (1995). A limited memory algorithm for bound constrained optimizaton . SIAM J. Sci. Comput, 16(5) , 1190-1208.
[19]. Schwartz E. (1997). Valuation Long-Term Commodity Prices: Implications for Valuation and Hedging. Journal of Finance, Vol LII, No 3 , 923-972.
[20]. Schwartz. E, .. (1998). Valuing Long-Term Commodty Assets. Journal of Energy Finance and development , 85-99.
[21]. Settlements, B. f. (2005). Zero-Coupon Yield Curves: Technical Documentation. BIS Papers .
[22]. Shanno, D. (1970). Conditioning of quasi-Newton methods for fucntion minimization. . Math Comp., 24 , 647-656.
[23]. Subramanian, K. (2001, June). Term Structure Estimation in Illiquid Markets. The Journal of Fixed Income, 11 , 77-86.

[24]. Svensson, L. (1992). Estimating and Interpreting Foward Interest Rates.
[25]. Toraldo., J. J. (1989). Algorithms for bound constrained quadratic programming problem . Numer. Math., 55(4): , 377- 400.
Wright, J. N. (1999). Numerical optimization. Springer series in Operations Research Springer-Vrlag, New York,

X. APPENDIX I: The L-BFGS-B Algorithm

A.1.1. Introduction

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

$$\min_{x\epsilon\mathbb{R}^n} f(x) \quad \text{(A1)}$$
$$s.t.\ l_i \leq x_i \leq u_i$$
$$i = 1, \dots, n.$$

Where$f\colon \mathbb{R}^n \to \mathbb{R}$ 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 Schrödinger) 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 (1995) is a standard method for solving large instances of $\min_{x\epsilon\mathbb{R}^n} f(x)$ 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 (1970), Fletcher (1970), Goldfarb (1970) and Shanno (1970)]. This method requires storing and updating a matrix which approximates the inverse of the Hessian $\nabla^2 f(x)$ and hence requires $\mathcal{O}(n^2)$ operations per iteration. According to Nocedal (1980), the L-BFGS variant where the L stands for "Limited-Memory" and also for "Large" problems, is based on BFGS but requires only $\mathcal{O}(n)$ operations per iteration, and less memory. Instead of storing the $n \times n$ Hessian approximations, L-BFGS stores only m vectors of dimesion $n,$ 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 et al (2011). We discuss how to modify the algorithm for non-smooth functions.

A.1.2. BFGS

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

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

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 $\mathcal{O}(n^2)$ operations per iteration, according to Wright (1999). In the case of non-smooth functions, BFGS typically succeeds in finding a local minimizer, as indicated by Overton (2013). 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.1.3. 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 $\mathcal{O}(mn)$ operations per iteration. So it is much better suited for problems where the number of dimensions n is large.

A.1.4. 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 *lu*-box defined by equation (A1).

A.1.5. 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 (1988) and Toraldo (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:

$$\mathcal{A}(x^*) = \{i \,\epsilon \{1 \ldots . n\} \backslash x_i^* = l_i \vee x_i^* = u_i\} \quad \text{(A3)}$$

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 $\mathcal{B}$ of variables are either on the lower or upper bound, as many as $\mathcal{B}$ 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 $\boldsymbol{\mathcal{B}}$.

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

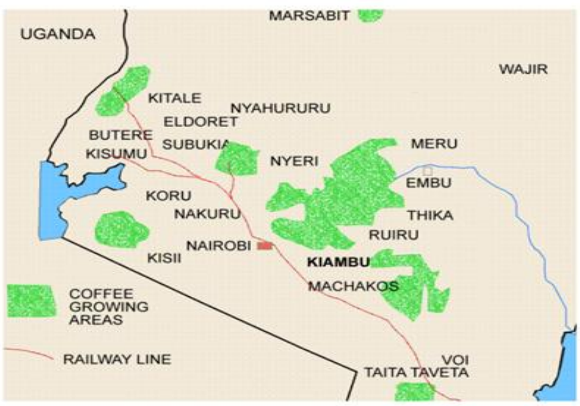
$$m_k(x) = f(x_k) + \nabla f(x_k)^T(x - x_k) + \frac{(x-x_k)^T H_k (x-x_k)}{2} \quad \text{(A4)}$$

Where H_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. $f(x_i) = \frac{1}{2}\Sigma_{i=1}^{p}[x_i]^2$, 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.

XI. Appendix II: Coffee Map Of Kenya

Annex figure 1. Coffee map of Kenya



Source: Coffee Board of Kenya.