LUND UNIVERSITY
School of Economics and Management

School of Economics and Management
Department of Economics

NEKN05
Economics -
Degree Project Master of Science in Business and Economics

Spring term of 2013

Can the forecast of the cotton price be improved using a model based upon economic variables?

Authors:

Camilla Olén
Tobias Norrman Andersson

Supervisor:

Karl Larsson

Abstract

The purpose of this thesis is to find a model, which is based on economic variables that can forecast the cotton price better than commonly used benchmark models. A vector error correction model is used because of the existence of non-stationary variables and one cointegration relation in the data. Two types of forecasting methods are used for out-of sample predictions. The dynamic forecasting in this thesis is used to predict the cotton price six days ahead and the static forecast only predicts one day ahead. Three different types of estimation windows are used to see which gives the best forecasting results. The residuals are then used to calculate the root mean squared error, RMSE, enabling the comparison with random walks and autoregressive processes. The static forecasts did result in significant better forecasts than the benchmark models while the dynamic forecasts did not produce significantly better nor worse results than the benchmark models. Including economic variables when predicting the cotton price only significantly improves static forecasts of one-day ahead predictions. A sign prediction test was conducted in order to test the static method for one-day speculative purposes, and significant results were found.

Keywords: Cotton, VECM, Forecasting, Out-Of-Sample,

Table of Contents

1. Introduction

Cotton as a commodity is fascinating all in its own right and its value as an export good for many countries inflates its importance even further. This thesis does not aim to revolutionize forecasting. Instead, the goal is to use current and accepted methods in order to test if a model can be established that surpasses commonly used benchmarks.

The discussion on forecasting, if it is possible at all and if so is the case, which model would yield the most accurate result, has been raging on in academic circles for decades. The most common case of forecasting studies concern inflation, GDP or currency forecasting due to their vast impact on the economy and the importance to stabilize the evolution of these variables. In this thesis the goal is to find a model that can predict future cotton prices with the help of econometric modelling, using a vector autoregression model and out-of-sample forecasting methodology. It must be stated that forecasting is a very difficult task, proven through trial and error of academics and professionals alike throughout the years. Nevertheless, the ambition of this thesis is to estimate a short run model that surpasses an autoregressive benchmark and a random walk benchmark. With the forecast horizon set to six days as well as one day, the forecasts produced will be deemed as short run. With the repeated testing over five rolling weeks per sample the goal is to achieve a model that can forecast the cotton price six days ahead while testing the vigour in both volatile and quiet markets. The static model is also evaluated using a sign prediction test in order to see if the predicted value is heading in the same direction as the actual cotton price.

In order to ensure the consistency of the estimated model, the tests conducted are repeated using different subsamples as well as a sample covering the full data range with unique in-sample as well as out-of-sample observations. The motivation for this approach is that it efficiently reduces the probability that the model is estimated during a relatively stable market and therefore will produce overly well-performing results. Some of the most severe periods of crisis and volatility during the late 20:th century and the 21:th century is included in the sample, and through the split into subsamples the model will be tested in both quiet markets as well as a volatile market.

In this thesis, both a static as well as a dynamic forecasting method will be employed. A static forecast will be executed in order to forecast a one-day ahead price level using actual observations up until the day in question, step-by-step forecasting six days into the future. A dynamic forecast will be used in order to conduct forecasts for one six days ahead using the predicted values instead of the observed values.

The estimated model is based on economic theory and fundamentals, and therefore the included variables are individually motivated in the *data* section. The variables are:

- Cotton
- Cotton Futures
- GDP total for OECD countries
- Maize
- Oil
- S&P 500
- Sugar
- Wool

A greater amount of variables was included from the start, but only the ones that are ultimately used are shown above, which did show significance.

1.1 Purpose and motivation

The purpose of this thesis is to investigate if a model based upon economic variables can predict the future cotton price level in a more accurate way than a random walk model and an autoregressive model

While forecasting is used within a wide array of different fields most of the economic research within the subject is based on forecasting inflation and exchange rates, while the area of commodity forecasting, and remarkably cotton, is not as explored. Cotton forecasting is interesting in particularly three ways. Firstly it is a commodity greatly used the world over in industries that greatly influence the economy of mainly third world countries where price predictability is key for a stable and long run development. Secondly it is useful in pricing of hedging instruments and understanding these, where these contracts carry great weight for producers, consumers, exporters and importers of the commodity. It is also attractive as a mean of detecting mispricing and thereby to conduct speculative trading. The speculative performance is evaluated using the sign prediction test.

1.2 Problem Definition / Outline

The aim of this thesis is to answer the question, can there exist a model that provides a better forecast than a random walk or an autoregressive model? The purpose of the thesis is to find such a model in order for both speculators and hedgers to be able to better predict future price movements both in quiet markets and times of great volatility.

The paper is organized as follows. The second section presents a brief *Background (2)* of cotton. In the third section a *Literature Review (3)* is given, providing a description of how previous literature has helped this thesis both in limitations and support. The next section handles the *Data (4)* used and the manipulations conducted on the data are explained. Empirical *Method (5)* explains the method regularly used in forecasting, how it is constructed and how it is used. The *Choice of Method (6)* is given as well as a motivation for said method. The *Result (7)* section discusses the forecasts and their evaluation. In the following sections, the *Conclusion (8)* is revealed and lastly suggestions on *Future Research (9)* are being discussed.

2. Background

Cotton History

Cotton is a widely used commodity and has been so for at least 7000 years with its origins in the Middle East. The first sign of commercial growing of cotton was in Pakistan, more precisely in the Indus valley where Alexander the Great brought back cotton to the northern hemisphere. This started a long and extensive period of cotton usage and trading of a commodity that proved to be extremely valuable. Applications of cotton are ranging from clothing to furniture and specialized usage such as fireproof chemically treated fibers. The mainstream success of the fibre however had to wait until the East Indian Company started to import cotton to England in the 17th century, where wool was the primary commodity for the apparel and textile industry at the time. Even with this development cotton did not reach the popularity of today until it was introduced to the North American market where it could be grown in large quantities, and in collaboration with the development of railroads it would aid in the development of the industrial revolution (Yafa, 2005).

The role of cotton has been enormous and fuelled conflicts far and wide. The most notable is the American Civil War, where even then modern European economies of France and Britain was on the verge of supporting the southern states in order to secure their cotton supply. Their ideological opposition to slavery, among other variables, forced them out of the conflict, at least in the sense of an exclusion of a full-scale military support that could have changed the outcome of the war (Owsley and Owsley, 1959).

Cotton is used as a primary commodity in many developing countries, including Pakistan, Brazil and Uzbekistan, representing a significant part of their national accounts as a mean to finance the development of their economy. Due to this fact, the cotton industry is critically important for these economies and therefore a volatile price climate will drastically increase the level of economic uncertainty in these nations. In order to reduce the level of uncertainty futures contracts has grown increasingly important as a hedging instrument for commercial farmers. With organized commodity markets developing farmers were able to easily fix the future price of their crop and in that way secure future earnings.

The case of the U.S. is a particular one, where in 2011 the country was the third largest producer of cotton in the world, and by far the largest exporter. This can be explained by the theory of absolute and comparative advantages where the U.S. has great areas of fertile ground in the southern part of the nation, giving them an advantage in the land-intensive production that is agricultural farming. Due to expensive labor, the labor-intensive production such as the refinement of cotton to cloth and textile is more fitted for countries such as China and Bangladesh (the largest and second largest importer of cotton respectively). In order for U.S. cotton to be competitive, great subsidies was required. China, also subsidizing the fiber to a great extent, is the greatest producer, importer and consumer of cotton in the world, while stocking up on great cotton inventories to hedge themselves in case of future price swings.

Cotton is a part of a group of commodities known as soft commodities. Soft commodities differ from hard commodities in the sense that soft commodities are usually grown while hard commodities are mined. Soft commodity prices are determined by actual supply and demand in a greater degree than they are by non-fundamental speculation and behavioral overreactions. This is because commodities are deemed as necessities and are therefore not subject to rumors and speculations in the same degree as stocks or hard commodities. In a sense they are constantly in demand, following an increased demand trend due to an increasing population. However, it is important to note that speculators can choose to build up a private supply with the purpose of holding a part of the world supply, thus reducing the supply with the expectation that the commodity in question will claim a higher future sell price (Hamilton, 2009).

Subsidies

The cotton sector has been greatly subsidized throughout the years by governmental support. With more than a fifth of the worlds cotton production in 2001/2002 was being produced through government subsidized farming, mainly from the U.S., EU and China (Gillson et al. 2004). This has raised great debate throughout the years because cotton prices became depressed due to this artificial price created by the respective government. The most notable case is the American subsidized cotton and the USA – Brazil dispute of 2004 where Brazil won the dispute through a ruling by the World Trade Organization (WTO). In recent years the U.S. subsidies for cotton has decreased to a lower level, while China and the EU still heavily subsidizes the commodity (ICAC, 2012).

Weather dependency and a complex cultivating process in order to create the highest grade of cotton have made cotton a commodity most suited for a limited amount of nations. Cotton can be grown in a wide range of climates but the need for water is substantial in order to get a full harvest (Gerik et al. 1996). There are considerable research being conducted in the field of alternative methods, for example generically mutations for reduced water dependency, but as of now the progress is limited. As an effect of this dependency, the price of cotton varies greatly with weather conditions and therefore the development of cotton hedges has grown to be increasingly important. In the event of a draught all commodities in the affected area suffer greatly, but among the most affected are typically the cotton farmers.

Competitive Crop

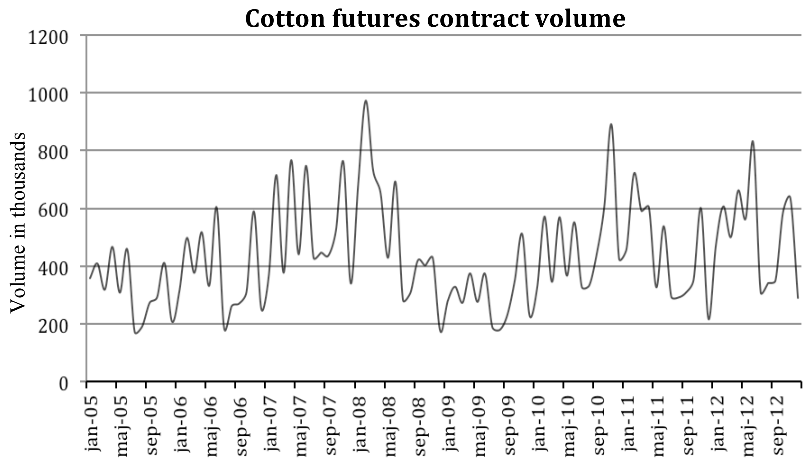
Cotton has only few competitors in the U.S. but on the other hand, it does also compete for acres with soybeans in the South and maize in the Midwest. Maize production has increased drastically in recent years due to an increased demand in ethanol, where maize is a major component used in the distillation process.

Cotton Trade

Cotton reaches back a long way, being traded in New York since 1870, on the New York Board of Trade.

The Intercontinental Exchange (ICE) acquired the New York Board of Trade in 2007, creating a new centre for the so-called soft commodities, which broadly represents commodities such as cotton, coffee, soybeans and orange juice. On the ICE, cotton is traded as future contracts as well as options on future contracts, under the contract name Cotton No. 2 futures and options under the ticker CT (ICE, 2012). Cotton futures are also traded on the Chicago Mercantile Exchange.

Cotton Options were introduced in 1984, and the volume of contracts has been growing increasingly since the end of the 20:th century. The participants in cotton future contracts are mainly cotton farmers who are hedging for the price risk, and speculative traders. In the last years the market activity for cotton futures has been varying, with a somewhat stable trend over the past five years. While observing market activity, clear signs of seasonality can be detected which is natural for any agri-commodity.



Graph I – Cotton future contracts volumes from 2005 – 2012 traded at the ICE. Volumes are displayed in thousands. Source: ICE.

Disturbances and abnormal events

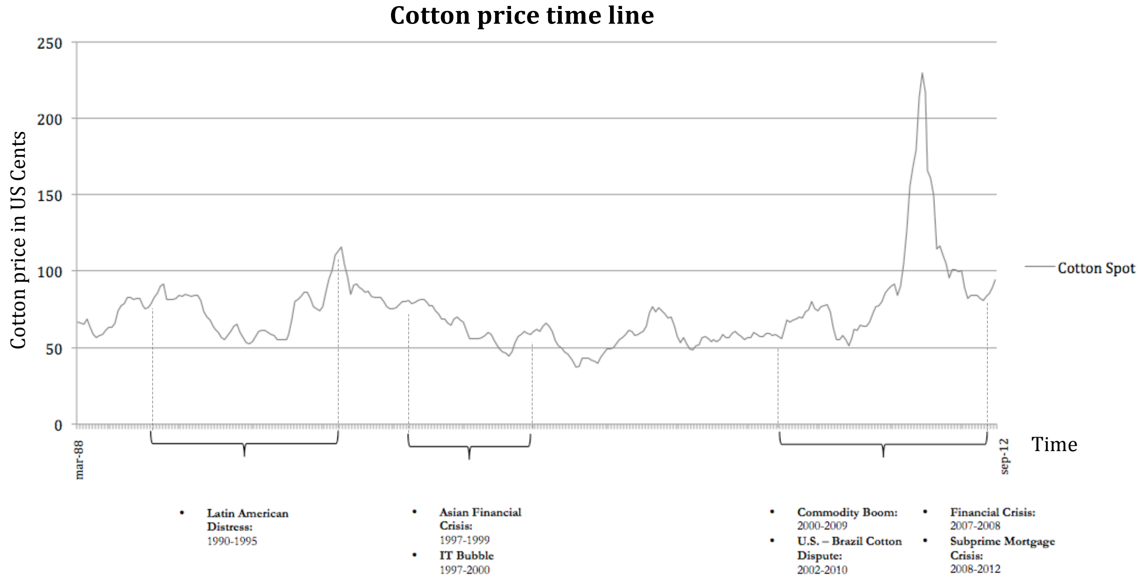
Latin American Distress: In the early 1990s there was a period of great uncertainty in Latin America, a geographical area with an extreme importance when it comes to commodity production. Brazil, one of the major cotton producers in the world, struggled with political unrest from the military regime that ruled the country up until 1990. The nation implemented harsh stabilization policies to reduce inflation without greater success while being hit by repeated corruption notices and doubt from the world (Dillinger, 1998). During the same decade Mexico was hit by an economic crisis, called the Mexican Peso Crisis, where the Peso was put under enormous pressure alongside with the burden of a deep debt crisis, a crisis that spread throughout the whole Southern Cone (Gil-Diaz, 1998). South and Central America had a series of difficulties during this era, difficulties that disrupted market stability and increased commodity price volatility, including the volatility in cotton.

Asian Crisis: As described by McKibbin and Martin (1999) commodity prices and a potential shock to these did not reflect the events that occurred in the East Asian Crisis. Instead, it was caused by reduced faith in the future profitability of the area and through that a reduced inflow of investments. They do describe how this differs greatly from previous crisis observed, such as the Latin American crisis of the 1990s, in the sense that commodity prices did not boom, rising production costs and lowering terms of trade as a cause of the crisis. Even if commodities where not the cause of the crisis, the price movements of commodity prices did reflect the volatility of the markets of the affected nations.

The Dot-Com Bubble: The IT bubble of the late 1990s did not affect commodities in a great degree per se, but greatly shook the financial stability of the global economy. The period was characterized by overinvestments in the IT sector that was based on no fundaments of any kind. Optimism spread to other parts of the stock market and soon inflated the whole financial system. When the bubble burst it left its footprint on the commodity markets as well, with signs of a depressed economic environment worldwide, if not in the same degree as it left on information technological stocks. Sharma (2012) describes how commodity stocks rose after the dot-com bubble burst to replace technology stock dominance, pushing commodity stock prices in place of technology stock prices.

The 2008 Commodity Boom: Soft commodities did not play as large a role in the 2007-2008 commodity boom as energy and metals did. Having said that, there was still a significant impact on the price of soft commodities from this boom. For examples the bankruptcy of three major U.S. Cotton Merchants brought down because of the spike of 2008 where the increase in future contract margin calls was not met (Carter and Janzen, 2009). The spike has been blamed on speculative traders as well as investors trying to diversify their portfolios and in the process driving up commodity prices. This theory is however widely contested (Östensson, 2012).

The U.S. – Brazil Cotton Dispute: The U.S. cotton subsidies was a market distortion destroying competitive markets and free trade according to Brazil, among others. In 2002 the World Trade Organization was initiated in a cotton subsidy dispute requested by Brazil against the U.S. At the time USA was the second largest producer of cotton, as well as the largest exporter. This was partly due to the subsidized price that U.S. farmers could claim. The dispute was not finalized until several years later due to the fact that both parts repeatedly applied to recourse the issue. The resulting ruling was in Brazils favour, with the two parties signing a Memorandum of Understanding in 2010. Both the U.S. and EU were announced to have used loopholes in order to stay competitive in a way that was not sustainable. It was unfair to developing country farmers that possessed an actual advantage when it comes to growing agricultural commodities such as cotton, but could not compete with these depressed cotton prices (Baffes, 2011). Part of the drastic price increase in 2009 until 2011 could be explained by this deregulation of the market where prices were allowed to soar to their natural level, even if this level was severely overshot before reverting to a slightly higher mean than before the dispute.



Graph II – The evolution of the cotton price since 1988 with notations of the market shocks and crisis.

3. Literature review

In this section previous literature related to this thesis is being presented. The papers in this section is however only the papers that represents a wider relation such as common linkages among commodities as a whole. More specific papers such as the relationship between the stock market and specific commodities are given space in the *data section* of this thesis under the respective variables subsection.

The early work of Pindyck and Rotemberg (1988) tests the co-movements of commodities that are otherwise unrelated, by testing their response to changes in macroeconomic variables such as inflation and exchange rates. This long run relationship points out that a factor may contain some predictability even if the short run relationship is inadequate. They do not find a significant result and conclude that fundamental variables cannot explain co-movements of commodities. Instead point out unobservable variables such as psychological variables resulting in herd movements. This is very early work in a market that is much different than todays, both in terms of efficiency and liquidity, but their work is still worth to be mentioned.

The linkage between agricultural and energy prices have previously been examined by Hertel and Beckman (2011). In their working paper they examine the relationship between ethanol, oil and agricultural prices such as corn and sugar, not taking any notes of, for example, cotton. Their work shows that ethanol prices are highly related to that of sugar, but not as related to the price of oil, and in all cases no long run relationships are shown but instead high degrees of mean-reversion. The strongest link in their research is that between corn and gasoline, a relationship that helps motivate the inclusion of energy prices.

Ai et al. (2006) examined the co-movements of commodity prices among five different commodities but they chose to exclude cotton due to insufficient data. They were examining the commodity prices using quarterly data from 1957 until 2002. This thesis uses high-frequent data and a less historical approach to the relationship and is data shortage is therefore not an issue. Their findings are that commodities tend to correlate due to common variables in demand and supply. They examine the correlation between the different commodities, not the causal variables of single commodity price movements.

As mentioned by Tse (2012) the debate on whether speculative, non-commercial trade disrupts commodity prices and increases their volatility is highly divided. Tse discusses how different research provides different results on the subject. The effect is that it cannot conclude that speculative trading does not significantly affect commodity prices in a general case. This surges the motivation for an inclusion of a speculation proxy, ranging from an agricultural ETF (exchange-traded funds) or non-commercial future contracts. In this thesis, future contracts with six months to maturity will be included as a proxy for speculative trading.

Greater commodity booms are rare, but still existent. In the event of such a boom, parameter estimates will be influenced by this abnormal event, lowering the prediction power of said estimates due to the inflated parameter estimates. In the sample covered there is only one peak severe enough to be classified as a commodity boom, being the boom occurring in 2008. Before this peak the previous boom occurred in 1974 and between the time of the boom and the data range used in this thesis the market will have had time to revert to its mean and thereby cancelling out any distortions resulting from this abnormal behaviour. However, the boom of 2008 is in the middle of the data series and will be reflected in the estimates, a fact that is under consideration when the discussion regarding structural breaks is being conducted. The shocks and their impact is being described in great detail by Carter et al. (2011), as well as other, less severe crisis from a commodity price perspective.

4. Data

In order to conduct a proper and precise econometrical study a vast amount of accurate data is required. The data covered is historical data gathered using the Datastream database, created by Thomson-Reuters. All data is converted into logarithmic levels and are then manipulated further by taking the first difference of the prices to counter the issue of non-stationarity, which will be discussed later on. The variables are all quoted in high frequency, daily observations when available, with the single exception of Gross Domestic Product, which is released on a quarterly basis. To correct for this the variable has been interpolated. All of these manipulations are discussed in greater detail later in this section under the subsection *data manipulation*. GDP and S&P500 is treated as exogenous throughout the thesis.

The following table shows a summary of all the used variables:

Variable Name	**Description**	**Data Source**
Cotton	Daily data over low middling 1-1/16" cotton, a standard basis of cotton traded, quoted in USC/pound, provided by the US department of Agriculture.	Datastream
Cotton futures	Quoted at the Coffee, Sugar and Cocoa Exchange, settlement price, sold in USC/pound. Time till maturity is six months.	Datastream
GDP	OECD Total GDP on a quarterly basis, quoted in USD, interpolated to daily basis.	Datastream
Maize	Quoted at the CBOT, Corn No. 2, the most commonly traded type of corn, sold in USC/Bushel.	Datastream
Oil	Crude West Texas Intermediate, spot price, sold in USD/Barrel.	Datastream
S&P 500	CME S&P500 index, settlement price, in USD.	Datastream
Sugar	Raw, non-refined sugar, Quoted at the International Sugar Arrangement, sold in USC/pound.	Datastream
Wool	Continuous average settlement price, quality of the wool is 21 micron, sold in A$/Hank.	Datastream

Table I - The eight included variables and a short description, where micron is a definition of the diameter of the fibre measuring the quality of the fibre, a hank is a measurement of the length of a fibre, in a looped bundle, 560 yards in the case of wool.

Originally the model started out with more variables, such as different currency rate pairs as well as interest rates such as LIBOR. However, they were deemed unfitting due to low prediction power or providing a prediction already included in another variable, such as the GDP variable providing the necessary power that an interest rate variable would bring. This is motivated by the fact that interest rate fluctuations in some part reflected GDP growth and business cycles in the way searched for in this context. The method to start off with several variables and then dropping insignificant variables is called General to Specific and helps the model to be as precise as possible. VAR models tend to perform better with fever variables as well, supporting this method.

The variables used to predict cotton price fluctuations in this paper are cotton futures, maize, oil, sugar, wool**,** S&P500 and the a total GDP for OECD parameter. The data used for the variables are prices using daily data over a period of 17 years between 1995/06-2012/12, obtained from the Datastream (2013) database. This results in 4586 observations per time series.

As the variables are cited during the text and in tables the following notations will be used:

Notation	**Description**	**Example**
LNx	the logarithm of x	Lncotton
Dx	the first difference of x	Dcotton
x(-p)	the p lag of x	cotton(-p)

Table II - The notations used for the variables and a short description.

The choice of using high frequency prices on a daily basis, is based upon the fact that the assets are highly liquid and volatile and measuring daily returns are therefore of the greatest interest. The forecast horizon is limited to six banking days, and the model estimates coefficients using fifteen lags of the endogenous variables. Market liquidity is important since liquidity tells how much information the price contains, and how efficient the market is (Holmström and Tirole, 1993). With daily basis, the loss of information is minimal, but the data is also exposed to a great deal of white noise and disturbance. If by chance the estimated week happens to occur during abnormal circumstances, the estimated model coefficients will be sub-optimal and the forecasts will be suffering. However this phenomenon is mended by testing the model in several subsamples over five rolling weeks per sample.

4.1 Cotton

Cotton is traded in a wide array of different standards, ranging from the quality of the cotton to the staple length of the fibre. Among the commonly traded types of cotton there exist 27 different grades of official colours alone, each with 23 different staple lengths. The choice of actual data for the Cotton variable is chosen by observing the type of cotton with the highest liquidity, which represents the most commonly form of traded cotton. The cotton variable chosen is of the low middling type, 1-1/16” in staple length, and is provided by the US Department of Agriculture.

4.2 Cotton Futures

Cotton futures are sold at the Intercontinental Exchange, the CME and the CSCE where they are quoted in USD (Commodity yearbook, 2010). Cotton futures activity is notably high at the Zhengzhou Commodity Exchange in China, partly due to the fact that China is the world’s largest producer of cotton. For the purpose of continuity, the data selected is the future with six months maturity, traded at the Coffee, Sugar and Cocoa Exchange, which merged with the New York Cotton Exchange in 1998. The accuracy of the data is more questionable when considering the Zhengzhou Commodity Exchange due to its young age and the political conditions of its home market.

Commodities are often used as a way to diversify away risk from asset portfolios. Often stable commodities such as gold are used in order to hedge against major fluctuations in the market but less known assets are being traded frequently, such as cotton. Not all commodity trading is used for risk management and mainly two groups use cotton future contracts. The first group includes commercial traders and farmers where the contract is used as a hedge, while the second group includes speculative traders looking for profits (Amann et al. 2012). The inclusion of a cotton future contract is motivated according to economic theory where speculative trading is said to possibly drive up asset prices even though there is a high degree of market liquidity.

Janzen et al. (2012) studies the co-movements between cotton prices and structural explanations such as speculative trading and demand for inventories using a vector autoregressive model. They find that the major drivers for cotton prices are variables specific to supply and demand, and therefore speculative trading is not a major concern when it comes to mispricing of futures contracts. The main focus of their paper is more recent movements of the price and more specifically the effect of shocks to create pricing spikes. Considering this a speculative proxy will still be included due to the contested conclusion regarding the impact of speculative trading.

4.3 Gross Domestic Product

In a study conducted by Camacho and Pérez-Quirós (2013) the relationship between GDP and commodities is examined in greater detail. The paper fails to detect a long run relationship between GDP and commodity prices, where it is argued that short-term increases in commodity prices can be caused by short-term demand shocks. Also the matter of how commodity price shocks are procyclical is discussed. This further motivates how GDP values could explain current commodity prices and should be included in a forecasting analysis.

4.4 Maize and Sugar

Maize is used in the distillation process of ethanol, where in 2010 roughly 35% of all maize grown in the US was intended for ethanol production (Commodity yearbook, 2010). The inclusion of maize and sugar is based upon the competitiveness among agricultural products. Maize and sugar are both land-intensive in their production but the sensitivity of the price should differ from that of cotton in the sense of the final product use. Maize and sugar, in some extent, are not refined in a greater degree but are grown for final consumption as basic crops, while cotton is a part of a longer production chain. However, they still compete over acres as primary commodities and are highly traded and important variables for many nations national accounts.

4.5 Oil

Energy costs are relevant in most forecasting studies, where in the case of cotton it also affects the costs of the whole production chain of cotton. Oil is used as a proxy for overall energy costs that will affect the marginal production cost of each bale of cotton in the sense of plantation operations, logistics and changing customer costs. By customer costs the aim is to describe the costs that producers of final goods, such as clothes and fabrics, are faced with while operating their factories as well as their logistic process. Both oil and ethanol prices were investigated by Hertel and Beckman (2011) where the link between gasoline and corn is proven to be very interesting. Gellins and Parmenter (2004) describe how energy costs accounts for close to 80% of fertilizer costs, increasing the importance of the inclusion of an energy proxy. As well as representing an overall energy cost proxy oil is also the main component of PET, which in turn is used in the manufacturing process of polyester.

Mutuc et al. (2010) plotted the relationship between cotton prices in the US and the crude oil price, trying to determine the effect on cotton prices of supply and demand shocks in the crude oil market. The paper described how the cotton industry in the US has decreased dramatically due to increased fuel production in the form of ethanol. The rising oil prices and their effect on biofuel demand, production costs and logistics does also contribute to the decrease in demand.

4.6 Standard & Poor's 500 index

S&P500 is included to show how the financial climate affects the price of cotton in the sense that a more volatile financial situation would tend to increase futures contract activities as a mean of hedging your operational costs. Tse (2012) states how instruments are greatly affected by their local market climate even if their home market is stable, motivating the inclusion of an American stock index as cotton is traded in New York. The use of local market climate proxies such as the S&P 500 in the case of NYSE has been previously motivated when forecasting by, among others, Froot and Dabora (1999) and Chan et al. (2003). This is further motivated by the fact that securities traded at a market are more influenced by the local trading climate than their home market, and the cotton price used is the one quoted in USD in New York.

As stated by Jansson (2013) commodities tend to lag behind the stock market, a situation resembling that of a leading indicator, providing support for the inclusion of a broad stock index in the forecasting process. The phenomenon of leading and lagging indicators describe variables that tend to hold some prediction power over the future evolution of other variables, usually serving as a warning (or indicator) for short-term business-cycle movements. The Standard and Poor's 500 stock index is a very common leading indicator for the economy as a whole which can hold predicting power over other assets that tend to lag behind the stock market as well due to affect that greater stock movements tend to have on the expectations of common investors. A lagging indicator displays signs of a delayed co-movement with leading indicators. Jack Caffrey of JP Morgan expressed his view on commodities tracing behind economic growth effectively acting as a lagging indicator in an interview for CNBC (Caffrey, 2013).

4.7 Wool

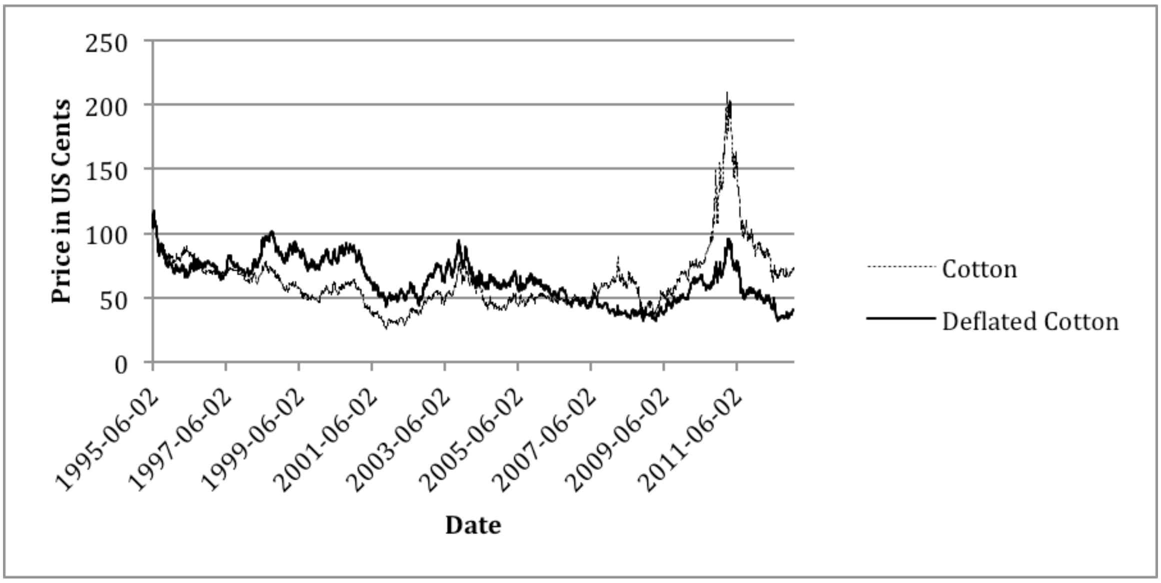
The use of wool in clothing is a direct substitute to cotton in many production processes and one that is not as dependent on weather conditions in such a degree as cotton. This makes it more reliable as a hedge as well, where insecure weather conditions will decrease the supply for cotton or possibly increase the demand for cotton futures to hedge against said conditions.

4.8 Data Manipulation

The data is transformed into logarithmic prices due to the variables being quoted in different units, ranging from USD/Barrel to A$/Hank. A logarithmic transformation will display the relative change in the factor and therefore will present comparable data.

To reduce the trend behaviour and in that reducing the probability of non-stationary data in the variables they are deflated using the S&P GSCI Agricultural spot index. The decision to use the agricultural index instead of the broader commodity index is because the broader index is focused in a large degree on energy commodities, overshooting the inflation in the agricultural commodities. The gross domestic product and the S&P500 index are not deflated as they are neither price levels nor endogenous. In graph III, the difference in the evolution of the original cotton price and the deflated price can be observed.

The evolution of the cotton price and the deflated cotton price



Graph III – The deflated and actual cotton price evolution from 1995 until 2012, displaying the trend reduction that a deflation brings to the variable. The cotton variable has been deflated using the S&P GSCI Agricultural spot index to remove overall inflation in the price.

The data of GDP in OECD countries that are used, as with all GDP measures, exists in its most high-frequent frequency on a quarterly basis. Because the data of the other variables contains daily observations, log-linear interpolation is used to generate high-frequency data from the quarterly observations.

S&P500 and GDP are considered to be exogenous in the estimated model but still to be included as they are considered to be important factor in order to forecast future cotton prices. In doing so they are not included in lags in a VAR analysis but only in their current level, reducing the risk of a singular matrix with interpolated variables such as the GDP variable used in this thesis.

4.9 Limitations

Ethanol would be interesting to include in the model because of its link to agricultural production, whereas maize and sugar are considered as proxies in this development in some degree. Polyester is also a variable that would be relevant to include due to its part as a substitute good to cotton, but lack of data limits the inclusion of the variable. This is further discussed in the *Future Research* part of this thesis. Polyester would also be of interest due to the attractiveness that a substitute commodity that is synthetic posses. Polyester contains many of the attributes sought for in cotton and can be used in a wide array of different textile industries such as clothing and furnisher. It also reduces the dependency that comes with organic fibres, including weather dependencies and the need for a great amount of hectares of fertile grounds. This variable is dropped by the same reason as Ethanol, the data available to date is not sufficient for a proper analysis to be conducted.

An interest rate variable was also considered but its close relationship with the GDP variable renders it unnecessary when a total OECD GDP variable was available. The interest rate variable would also be strictly linear when interpolated from lower frequency data making it even less attractive as an explanatory variable in this case, due to the low variance of interest rates.

The discussion on however currency exchange rate pairs would to be included was also conducted, with the falling judgment leaving it excluded. The reasons are several; including the difficulty to choose a representative pair of currencies or if it should be a producer currency, importer currency or exporter currency pair for example. The largest currency pairs also tend to covariate with the overall evolution of the developed world, an evolution capture by the GDP variable in a sufficient degree.

5. Empirical method

Choosing the correct model that fits the data needs careful consideration to make valid forecasts. Two types of models are a vector autoregressive model and a multiple regression. In previous studies, the majority of these types of investigations use a VAR model when forecasting. A VAR model simplifies the handling of lagged variables in a much greater degree than a multiple regression model. Due to the amount of variables and lags of said variables, a VAR model creates a more observable and easily manipulated model than a multiple regression would. The vector form of the model makes it more manageable and easy to overview. In addition to this most econometric software purposed with forecasting is based on the assumption that a VAR model is used. There are different types of VAR model that can be used. In order to decide on a proper VAR model a series of tests needs to be conducted.

5.1 Unit root / stationarity

Before making tests on time series data, the variables are often tested for stationarity in order to make correct inferences. When testing for long run relationships, the stationary term (ε_t) indicates that the variance, the covariance and the mean are constant. If this is not the case it can result in a spurious relationship, which means that two variables that are unrelated can show a clear dependence due to a trend over time. This trend can be due to variables that are completely independent of the two variables but it creates what looks like a strong relationship between them.

If the variables are found to be non-stationary, there are several ways to cope with this problem. A common way to correct for the non-stationarity is by differentiating the variables. When a variable is not differentiated but in its original form it is said to be in level. Differentiating a variable means that the value from the previous observation is subtracted from the value of the current observation as can be seen below:

$$\Delta y_t = y_t - y_{t-1} \qquad (1)$$

A series that needs to be differentiated once to induce stationarity results in variables in first differences. It can also be called integrated of order one, labelled I(1), or that the series contains a unit root. After differentiation, the series is stationary which is labelled I(0).

When using the variables for forecasting and thereby investigating the relationship in a longer perspective, non-stationarity may not be a problem. Differentiating the variables might cause more of a problem than leaving the variables non-stationary due to loss of information in the longer time perspective (Brooks, 2008).

When testing for a unit root, a Dickey-Fuller test is appropriate to use. The null hypothesis of a unit root is tested against the alternative hypothesis of no unit root in the time-series (Brooks, 2008). EViews uses critical values that are estimated by MacKinnon (1991, 1996), which, unlike the Dickey-Fuller test statistics, applies a larger number of simulations. The test can only be used if the series follows an autoregressive process of order one. An Augmented Dickey-Fuller test, ADF, can therefore be used when the dependent variable is autocorrelated of a higher order.

$$\Delta y_t = \alpha_0 + \gamma y_{t-1} + \sum_{i=2}^{p} \beta_i \Delta y_{t-i+1} + \varepsilon_t \qquad (2)$$

If the gamma coefficient, γ in equation (3), is equal to zero, the equation contains a unit root because there are only first differences left (Enders, 2010). The number of lags included in the equation needs to be defined when using the ADF. There are two ways to decide on the number of lags to use. The first one is only suitable if the data is sampled with low frequency, for example if monthly data is used, 12 lags would be chosen, which is the same as the number of months in a calendar year. The second one uses a measure called information criterion and there are a number of different criteria to choose from, such as the Likelihood Ratio-, the Schwarz- and Akaike Information Criteria. When comparing the result per selection criteria, the amount of lags showing the lowest F-value of a criterion should be chosen to represent the number of lags. The number of lags that is chosen in this chapter is only valid for the ADF test and not in the following chapter where the numbers of lags are selected for the VAR model.

The use of optimal number of lags is important because remaining autocorrelation might be the result of few lags and the standard errors of the coefficients will be larger if the numbers of lags are too high (Brooks, 2008).

Because the Dickey-Fuller test has low power if the series in question is stationary but is close to being non-stationary it is good practice to test the hypothesis using another test as well. To confirm the result, the Kwiatkowski-Phillips-Schmidt-Shin test, KPSS, is appropriate because it investigates if the series is stationary under the null hypothesis (Brooks, 2008). The test is specified as:

$$y_t = d_t^{'} + \beta x_t \tag{3}$$

where t is a number between one and T. The component $d_t^{'}$ is deterministic and the last term, x_t is stochastic. The test statistic uses residuals that can be obtained by regressing y_t on $d_t^{'}$ (Kurozumi et al, 2010).

5.2 Cointegration test

Johansen's cointegration test is often used when two variables are found to be non-stationary and a long run relationship wants to be found. If two variables are cointegrated, there exists a linear combination between them that are stationary (Brooks, 2008). In this sense it is motivated to test variables further for long run commitments even if the short run shows none.

In order to test for cointegration in a vector autoregression (VAR) is used on which the Johansen's multivariate method is based on. Only if the variables are non-stationary, this test is applicable. Before conducting the test an alternative of five different trend choices has to be specified. The option to test for "intercept (no trend) in CE and test VAR" should be chosen if the trends are believed to be stochastic. The original version of the Johansen test does not cover stochastic trends but only linear deterministic trends allowing for a constant term in the regression system. As argued by Campbell and Perron (1993) stochastic cointegration is often the part of interest in an empirical application. The model has been updated since then, allowing for tests for stochastic trends. There is an option available in the case that the most suitable alternative is unknown for the user, which provides a summary of all choices available. This alternative allows one to compare the results of cointegration relations from the five different options and through that determine the likelihood of a deterministic or stochastic trend or the existence of a intercept.

The lags that need to be specified are in first differences and the optimal lag length is achieved through an information criterion. The results display two different statistics, the trace test statistic and the maximum eigenvalue test. The trace statistic tests the null hypothesis that there are r cointegration relations that is being displayed in a list from zero up to the number of endogenous variables minus one, while the alternative hypothesis contain the same number of cointegration relations as the number of endogenous variables included. The maximum eigenvalue statistics uses the same null hypothesis while the alternative hypothesis is *r* plus at least one cointegration relations. It has been shown that the trace test is more robust against both skewness and excess kurtosis than the eigenvalue test. If the two tests would provide different suggestions then it is advised to proceed with the trace tests suggestion (Cheung and Lai, 1993). If the critical value, obtained from the results, is smaller than the test statistics; the null hypothesis is rejected (Brooks, 2008). If the test indicates the same number of cointegration relations as the number of endogenous variables, a VAR model can be used without making any corrections for the non-stationarity. A vector error correction, VEC, model can be used if one cointegration relation is found. If no cointegration relations can be detected, the problem of non-stationarity has to be corrected manually by first-differences (Van Aarle et al, 2000).

5.3 VAR/VECM

A vector autoregression is used to forecast future price levels because of the construction of the model. The dependent variable does not only depend on white noise and lags of the variable itself, but is also explained by other variables. Generally, VAR models tend to generate better predictions in more cases when forecasting than other structural models. A restricted VAR model, known as a vector error correction model, is used in the case when the series is non-stationary and there exists cointegrating relationships (Brooks, 2008). Below, a general VECM is shown for u_t that is a vector of *n* variables:

$$\Delta u_t = ab' u_{t-1} + \sum_{j=1}^{p-1} \psi_j \Delta u_{t-j} + \varepsilon_t \tag{4}$$

where *t* is a number between one and *T* and *ab*' is a matrix of dimension (*n*h*), the prime is the sign for transposition and h denotes the cointegrating rank. The columns of *b'* contains *n* cointegration vectors and the columns of *a* contains *n* adjustment vectors. The residual, ε_t is normally, identically and independently distributed. Ψ_j denotes a *n* * *n* matrix and consists of the coefficient estimates (Seong et al, 2011).

The error correction term u_t is also known as the cointegration term and is the variation around the long run equilibrium and is partly adjusted by short run adaptions. The term is in the long run equilibrium equal to zero only if u_t does not vary around the long run equilibrium. In the latter case, the error correction term is non-zero and the return to the equilibrium relation will occur due to adjustments of the variables in the model. The speed towards the equilibrium is measured by the coefficient *a* of an endogenous variable, in the simple case above; the number one or two.

When forecasting time series, a vector autoregression also enables oneself to analyse the effect of random noise. The VAR model treat all variables as dependent variables of the endogenous variables that are used and the results are shown in a vector (Enders, 2010).

Information criteria such as the Swartz- and Akaike Information Criterion help to decide the number of lags that should be used in the model. The smallest values of the criterion are chosen to represent the selection of the preferred model (Rassi, 2012).

5.4 Structural Breaks

After studying the graph of the evolution of a price, one can reason and assume that the data contains structural changes by observing the swings and changes in the mean of the price. If this is the case, the test statistic obtained from the Dickey-Fuller test is biased (Enders, 2010). In order to determine if there are any structural breaks in the data sample, a Quandt-Andrews test is applied. There are a few different tests that can be used when testing for structural breaks. The Chow test is common but it requires specific dates to be specified by the user, which can be difficult if one does not know exactly when the breaks occur. Quandt-Andrews test can be used instead when the dates for breaks are unknown (Brooks, 2008). The test examines the time-series for at least one breakpoint within the sample that is specified by performing multiple chow tests automatically, providing the user with information on the break with the highest F-value. The numbers of test statistics are summarized to obtain one test statistic with the purpose of testing if there are no breakpoints within the sample, specified as the null hypothesis (Adom, 2013).

5.5 Granger Causality test

The question whether the changes in one variable y_1 are caused by changes in another variable y_2 can be answered by the Granger Causality test. If the answer is yes, all lags of y_2 should be significant when y_1 is the dependent variable. If the other way around is not true, then y_2 Granger causes y_1. There is evidence for causality when the value of y_1 today is correlated with earlier values of y_2. This is a suitable test when the model includes many explanatory variables but mostly when there are many lags of each of the variables and therefore makes it hard to conclude which variables have significant effect on the other ones (Brooks, 2008).

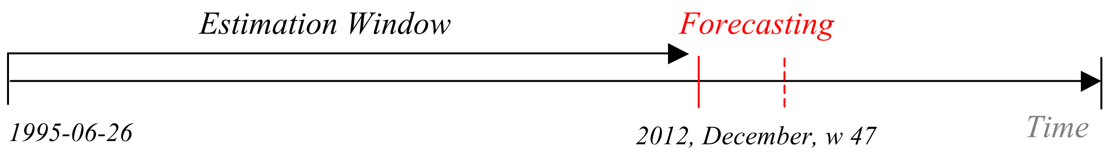
5.6 Forecasting

Static and Dynamic forecasts are two different types of forecasting methods. They can be used when making the out-of-sample forecasts, and therefore to predict the future prices which are treated as unknown. Static forecasting uses the observed parameters over time while the parameters used in the dynamic forecasting method vary and are updated with estimates but not actual observed values. In EViews, static forecasts only forecasts the next time period, one day, when the data contains daily observations, and therefore only uses the observed values in the estimation window. Dynamic forecasting, on the other hand, uses the same method when calculating the first observation as a static forecast but after that, the estimation window includes the previously forecasted values instead of observed values (Brooks, 2008).

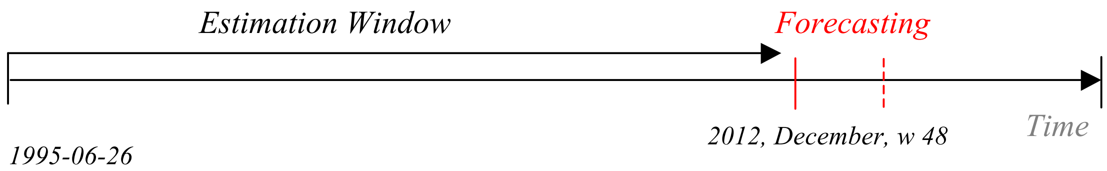
When forecasting in EViews, the observations in the in-sample-window are used to estimate the parameters. These are then used in order to predict future values of the dependent variable using estimated values for the explanatory variables in all of their selected lags. The forecasting period is known as the out-of-sample period due to the fact that these values are treated as unknown. The number of lags of the variables in the model are specified as the number of days/months/years back in time that are used when estimating the coefficients for the model that is used for forecasting.

Three different types of estimation methods are used; a recursive window (also called an expanding window), a rolling window and a fixed window. These are all presented because there are advantages and disadvantages of each method. Due to the lack of certainty that the model is being correctly specified, a limited forecasting method, such as rolling window, can be desired rather than the expanding window. This is because the old data, which is being excluded when rolling the window, might not be informative anymore, or stop the misspecified model from delivering proper result (Giacomini and White, 2006).

Below, there are a few figures that are presented in order to gain better understanding of each forecasting method. In the examples, the forecasting methods are used to estimate two weeks at a time, in December 2012. The first method is called recursive window where the start date is fixed at 1995-06-26. The end date is moved forward one observation at a time and therefore, the in-sample estimation window is expanding when calculating the parameter coefficients. Below, the recursive window method is displayed in the two figures. The first week, number 47, is estimated in (1) and in (2), the next week, number 48, is estimated.

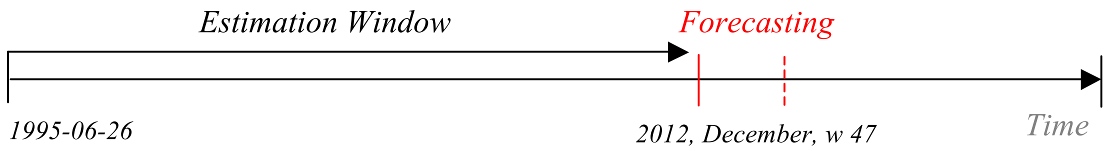


(1)

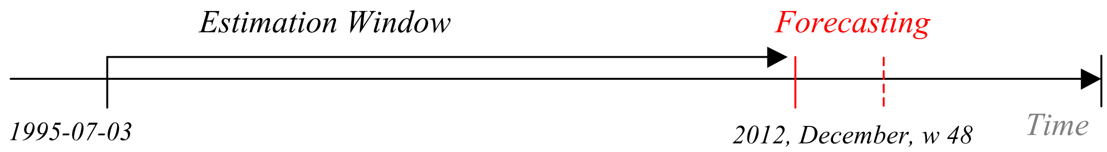


(2)

The second method, the Rolling Window uses a fixed number of observations as the estimation window with both the start- and end date is moving forward by one observation each time (Brooks, 2008). Therefore, the observations used to obtain the parameters are, as above, changing for each period that is forecasted. The first in-sample estimations are from 1995-06-26 until week number 46, in order to calculate week number 47, which is shown in the figure below (3). Week number 48 is estimated in (4), using observations from 1995-07-03 up to the week 47.

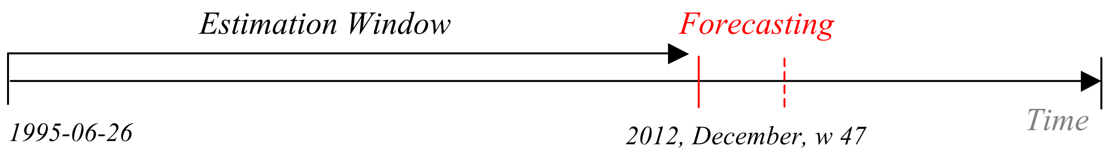


(3)

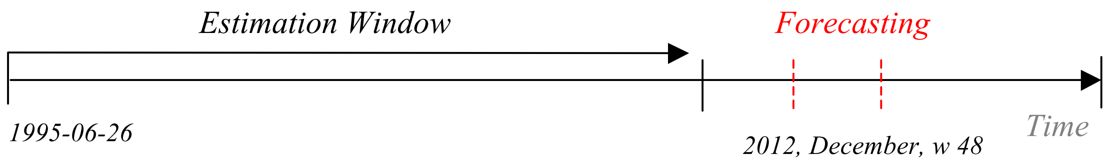


(4)

The fixed window always uses the same period as the estimation window in order to estimate the parameters in the model. This means that the estimated parameters are only estimated one time and used for all the out-of-sample forecasts (Giacomini and White, 2006). The in-sample estimation 1995-06-26 up to week 46, 2012 are used in both figures below when forecasting the two weeks. The same estimation period will apply to all weeks that are forecasted.



(5)



(6)

5.7 Benchmarks

To enable a forecasting model to be evaluated, it can be compared to a benchmark model. In order to be classified as a well-performing model it needs to beat the benchmark's prediction in a number of cases. The specific threshold depends on the confidence chosen, the power, by the implementer. The benchmarks that are chosen to be presented are an autoregressive benchmark and a random walk benchmark. To implement the evolutionary portion of an asset a random walk with drift is often most suitable, following the evolution below:

$$y_t = a + y_{t-1} + \varepsilon_t \qquad (5)$$

Where ε_t is a residual and for a > 0 the random walk will show an upwards trend, a < 0 a downwards trend and a = 0 will indicate no drift.

A random walk is used to see if the estimated model has better predictability than a model based on pure chance. If the estimated model does not surpass the random walk model the conclusion is that the model cannot predict the price evolution with a higher accuracy than a model based on randomness. The random walk model is simulated by calculating a residual based upon the cotton values, and then adding this residual to the previous value of the time-series. The simulated residual is calculated by taking the mean of the first differences of the actual cotton data series, added to the standard deviation of these differences times a random normally distributed variable. The simulated cotton price at time t is simply the actual cotton price at t-1 plus the simulated residual for time t. Then the actual cotton price is subtracted from the simulated cotton price in order to find the deviation from the correct value that the random walk produces. This error is the error used when evaluating the forecasting model.

The alternative benchmark is an autoregressive model using the values of the variable itself, lagged in time. An autoregressive model of order one, AR(1) model is defined as (Harvey, 1984):

$$y_t^* = \Psi y_{t-1}^* + \xi_t \tag{6}$$

where ψ is a measure of the weight put on the observation at t-1 and ξ is a residual.

When deciding on which model is better than the other, there are different forecast error measurements that can be used as evaluation tools. A measure that can be used is called the mean absolute error, MAE, which is also referred to as mean prediction error. It depends on which scale the dependent variable uses but it is not as sensitive to large residuals as the squared measures (McGee and Yaffee, 2000). The MAE shows the expected error on average, but the difficulty lies within the errors relative size. One can find it hard to tell if the error is small or large (Collingswoth, 2012). MAE is defined as:

$$MAE = \frac{1}{T}\sum_{t=0}^{T}(y_t - \hat{y}_t) \tag{7}$$

The second measure is called the root mean square error, RMSE, and is also dependent of the scale of the dependent variable. RMSE is calculated by taking the square root of the mean squared error, with the mean squared error defined as:

$$MSE = \frac{\sum_{t=0}^{T}(y_t - \hat{y}_t)^2}{(T-k)} = \frac{SSE}{(T-k)} \tag{8}$$

Where *SSE* is the sum of squared errors, *T* is the sample size and *k* is the number of parameters to be estimated. RMSE is useful when comparing different models using the same series and the model with the lowest RMSE will decide which is the better forecasting model. It allows for a fair evaluation when the errors are volatile but with a mean of zero, as positive and negative errors would cancel out using absolute errors while squared errors always result in a positive measure. The greater the measure, the greater the error. RMSE is defined as:

$$RMSE = \sqrt{MSE} \tag{9}$$

The mean absolute percentage, MAPE, is different from the other two measures because it does not depend on the scale of the dependent variable. However, there has been criticism against this measure due to the fact that small initial values will lead to asymmetry and instability. Considering this, MAPE is not suitable when comparing a model against a random walk without adjustments to the measurement (McGee and Yaffee, 2000).

$$MAPE = \frac{100}{T} \sum_{t=0}^{T} \frac{(y_t - \hat{y}_t)}{y_t} \tag{10}$$

When using these measures, the variance of the forecast error may vary over time if the model contains nonlinearities and the exogenous variables do vary.

6. Choice of estimation method

Before the tests are executed, the choice of variables has to be set. A model using variables only for the sole purpose of providing a forecast can be estimated by including a huge amount of variables and then dropping them one by one until only variables with a high significance remains. The problem with a model created in that way is that the variables is not always based on fundamental values and therefore might produce a well-behaving model in one period but only due to correlation and chance. In order to create a consistent model each of the variables used in this thesis is included due to fundamental economical reasons as well as significance.

The endogenous variables that are used are cotton, cotton futures, maize, oil, sugar and wool. Because of GDP and S&P500, that are not determined within the model, they are treated as exogenous variables.

The choice between using a VAR- or a VEC model is based upon the result from several tests, including tests for stationarity and cointegration. After these tests have been conducted the choice of model can be established.

Seasonality effects are quite common in commodity time-series, resulting in trends and patterns of varying market activity and prices. Cotton is no exception, and these patterns are handled in a matter of ways. The season variation is limited, due to the worldwide production of the crop resulting in a spread out harvesting season over the whole annum. For the three largest producers of cotton the harvesting season is split as follows: the USA harvests in March until April, in China harvest occurs during the period of September until October, and India harvests from November until February. The subsamples that the data is split into are forecasting the cotton price during different months which helps reducing the seasonality further. Also, the forecasts are short-term, ranging from one day ahead for the static forecast and six days ahead for the dynamic forecast. With short out-of-sample periods and long in-sample periods the seasonality is not of a great importance neither in the estimation process of the coefficients nor in the forecasting process, in the case of cotton.

6.1 Stationarity

When testing for non-stationary variables, the augmented Dickey-Fuller (ADF) test is used instead of the Dickey-Fuller test. This choice is most appropriate because of the data used, which is more complicated than a simple autoregressive process of order one.

The test result from both the ADF- and the KPSS-test did show all variable being non-stationary apart from the sugar variable (see Appendix A-C). Sugar shows signs of stationarity during the KPSS test and is tangent to be stationary during the ADF test. There exists a risk when using one variable that is stationary while the rest are non-stationary, and first differences are being employed to rid these of their non-stationary behaviour. If a variable is stationary in level but is then differentiated there is a risk of the variable to become non-stationary instead. Manual differentiating does not risk this problem because the user can chose which variables to differentiate and which to leave in level, but employing an error correction model with automated differentiating does. Because of this, sugar is being tested for a unit root, both in level and in first differences in order to see if a first difference will in fact create a non-stationary behaviour in a variable originally stationary. In this case, stationary still persists after differentiating and no further procedures must be undertaken particularly for the sugar variable.

The conclusion is that all variables in the model are considered to be non-stationary due to the above information with the exception being sugar. Before taking any actions, a cointegration test has to be conducted to see if further manipulation is needed before proceeding with regressions that are required in order to estimate the VAR/VECM model.

6.2 Cointegration test

The result from the Johansen's Cointegration test are displayed in appendix D and state that there is only one cointegrating relation. As can be seen in the table, the reason is that the null hypothesis of no cointegrating relations is significantly rejected. The following null hypothesis of at least one cointegrating relations cannot be rejected and therefore the conclusion is that the data contains one cointegration relation. As mentioned before, when only one cointegration relation is detected, the VEC model is the most suitable choice. As a result of this the variables are used in first differences in the error correction model that reduces the non-stationary behaviour of the variables.

6.3 VAR Lag selection

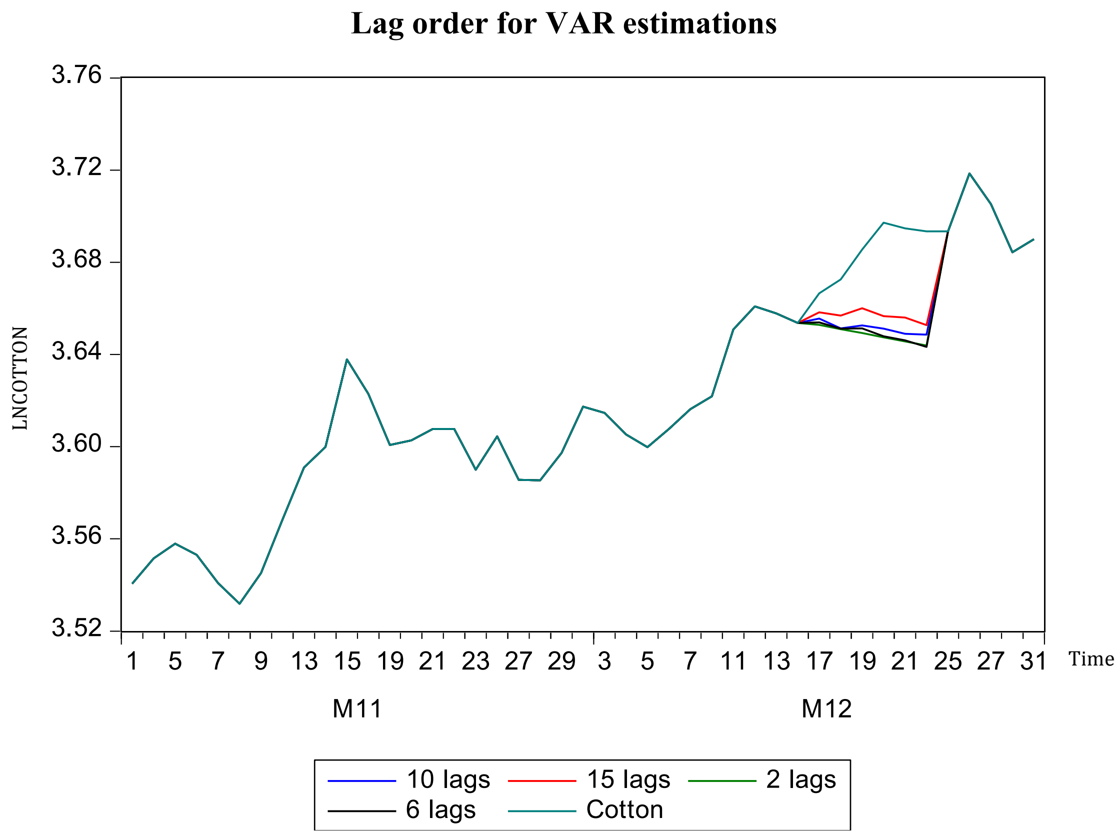
Before estimating a VAR model the appropriate number of lags has to be specified. Using the likelihood ratio-, Akaike- and Schwarz- information criteria the optimal lag length can be determined. In table III the results from these tests are shown. The greatest statistic is the number of lags recommended by the selected criteria.

VAR Lag Order Selection Criteria

Lag	Likelihood Ratio	Akaike	Schwarz
0	NA	-5.146	-5.120
1	124745.4	-32.504	**-32.428***
2	219.578	**-32.537***	-32.410
3	41.882	-32.530	-32.353
…	…	…	…
14	32.508	-32.463	-31.728
15	**51.132***	-32.459	-31.673

Table III - Lag order selection Criteria for VAR

The Akaike information criterion suggests two lags and the Schwarz information criterion suggests one lag, while the likelihood ratio criteria suggest 15 lags. It must be noted that these are suggestions based on different calculations, not rules. In order to determine the best lag length in this specific case forecast with two, six, ten and finally 15 lags were conducted to test the model. When all forecasts where graphed together, 15 lags were chosen due to the best predictability but also because of weekly differences in the amount of actual banking days. Considering holidays there can be a significant data loss in the lagged sample using fewer lags. Using even more lags than suggested will contribute to data loss by reducing the number of observations in the estimation sample. However, being on a daily basis, a reduction of 15 observations per variable is not to be considered unreasonable in a large data sample. From the testing of the different forecasts, the one with the 15 lags does capture volatility better by far than the shorter lag lengths while reducing the data material size in a reasonable magnitude and therefore providing a superior forecast. The different lag lengths when forecasting the last week, 2012-12-17 till 2012-12-24 using a recursive window, are shown in Graph IV.



Graph IV - The difference in forecast accuracy between different lag lengths.

6.4 VECM

Due to the result from the previous tests, the vector error correction model is chosen instead of a vector autoregressive model to cope with the non-stationary data and the lack of fully cointegrating vectors. The VECM is suitable for this type of data and is therefore superior to a standard VAR model. The VECM is then used in order to produce coefficient estimates for the forecasts.

6.5 Structural breaks

As mentioned earlier, by looking at the plotted cotton price one can suspect that there exist structural changes over time. An article by Baffes and Haniotis (2010) discusses the effect of the 2006/2008-commodity price boom and they find that price variability overwhelms price trends and therefore the outcome from testing for price trends depend on the time period. In the light of this result, the insight is used in this thesis to split the analysis into several different subsamples as well as a sample covering the whole period. This is done in order to capture the forecasting ability of the chosen variables in different settings. The Quandt-Andrew test is used instead of the Chow test due to the advantage of not needing to specify the exact suspected break points, which is allowing for subjective selections in the

Chow test. The test did first signal a breakpoint at 2010-02-16. After cutting the sample up to the first breakpoint, the second breakpoint indicated on 2003-12-01, resulting in three subsamples. Running the test one last time the decision about not including this breakpoint was regarding the number of observations in each sample that was not enough for proper parameter estimations. The following subsamples are the ones that are used throughout the thesis, containing the dates from the Quandt-Andrew test result (see appendix E for complete test results).

First subsample	Second subsample	Third subsample
1995 06 26 - 2003 12 01	2003 12 01 - 2010 02 16	2010 02 16 - 2012 12 24

Table IV – Resulting subsamples from the structural break analysis.

The structural break analysis signals when a break occurs in the sample. However, it does not signal when a break reverts to its mean, resulting in the possible inclusion of previous abnormalities in the subsample to follow. When deciding on how to treat the structural breaks, forecasts with different data-range were committed on subsample three, which contains the most volatile part of the three subsamples. The data included in the original subsample were ranging from 2010-02-16. In this sample the great price spike of 2010 was included in the information window, increasing the noise in the estimation of the coefficients. Therefore, another forecast was performed but instead pushing the starting date to 2011-07-15, effectively removing all traces of shock disruptions. The test results proved that this method did not produce statistically significant improvements over the sample where the shock was included, as can be seen in table V.

Information window	**Exclusion outperforms inclusion of peaks**
Static Window	2/5
Rolling Window	2/5
Expanding Window	2/5
Total	6/15

Table V – summarized results when comparing RMSE for subsample 3 including spikes and one that excludes spikes.

In total the estimation sample that were cut, only outperformed the complete subsample in 40 % of the cases, rendering it as unfit for prognostic use in the sense that it both produces poorer result as well as reducing the information provided by the sample by a significant degree. The conclusion of using the larger subsample, including the spike, is therefore the obvious choice.

When the data material is split into subsamples the parameter coefficients are estimated for each subsample. The model is the same, but the coefficient estimates will differ depending on the estimation period in each subsample. The purpose of this procedure is to rid the coefficient estimations from disturbances from abnormal events. By doing this, the model gets unique coefficient estimates for each subsample, allowing for more precise forecasts in each subsample.

6.6 Granger Causality test

A Granger Causality test was conducted but without contributing too much additional information due to the forecasting nature of this thesis. The variables in combination are shown to granger cause cotton. This result can be seen in appendix F, where the table shows that the variables in combination granger causes cotton, but their individual effect is limited.

6.7 Forecasting

The two different types of forecasting, static and dynamic, were both chosen to represent the forecasting method of the chosen model. Static forecasting only forecasts one day ahead and is therefore a suitable choice in order to find out if the model is working in the simplest case. Dynamic forecasting can predict the price the number of days ahead that is specified. The reason for choosing both methods is to be able to see if the price of cotton six days ahead down to one day is possible to predict.

The three estimation methods, rolling, expanding and fixed window, are all included in the forecasting result. This is because different methods do include both strengths and weaknesses. The decision whether one of them is more suitable than the other is difficult and therefore all of them are selected to enable the results from each method to help the decision of the better forecasting model.

The Expanding Window method is preferable in the case when there is additional information to be gained from a larger estimation window. It includes both the original start date of the estimation sample as well as a rolling end date, enlarging the estimation window over time. A potential threat using this method is that the start date only includes disturbances and noise and therefore a rolling window is preferable due to the fact that the rolling method only includes these observations once. A Fixed Window can outperform the other two methods in cases such as a period of calm before a short-term shock, and then a reversion to the quiet markets. With a fixed window, this shock term will not be included for later forecasts and the estimates will only use calm data. This is a double-edged sword however, if the fixed window happens to occur in the middle of the shock for example, then all estimations will be based on the approximated coefficients that are estimated during a highly volatile period.

6.8 Benchmarks

Models that are used as benchmarks enable conclusions to be made from the results that are obtained from the forecasts. In order to increase the reliability, a few actions are taken, for example testing the forecasts against more than one benchmark model, the random walk and the autoregressive model. Comparing 100 random residuals from a random walk with the forecasted errors instead of only comparing against one random walk does provide reliable results and the result are not based on one random occurrence. The motivation behind using an AR(1) and an AR(15) model is such that an AR(1) model is the most basic autoregressive model, displaying the immediate explanatory power of yesterday for today. In the estimated forecasting model, fifteen lags are used to capture weekly volatility differences and this deems a comparison to an autoregressive model lagged fifteen times suitable. The same logic is employed when estimating an AR(15) model as described earlier when estimating an AR(1) model.

The models are then compared and evaluated using the root mean square error measure because it is suitable when comparing different models. The power chosen for the evaluations is 10% for significant results to be achieved with some margin for random error. The measure MAPE that was mentioned was not chosen because it was not suitable to be compared with a random walk, which is used as one of the benchmarks. Another measure that was mentioned, MAE is not used here because the small errors are hard to distinguish from larger errors.

The following section shows the results from the forecasting model.

7. Results and Analysis

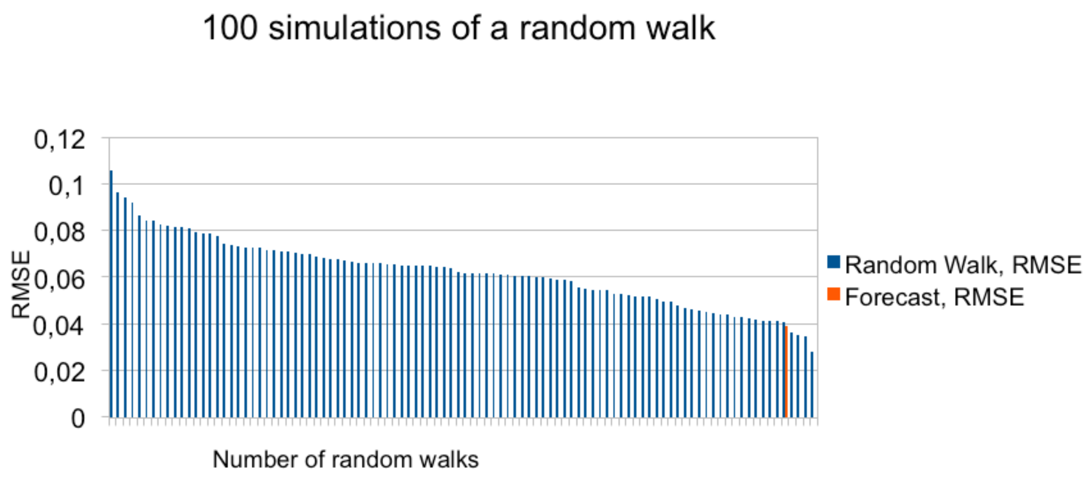
The purpose of this thesis was to investigate if it was possible for an estimated model using fundamental variables to achieve a better forecast than a random model or an autoregressive model. Due to the high degree of volatility in the sample range, the data was split into three subsamples in order to see if a well performing forecast in one subsample could be due to randomness and not to a well performing model per se. The result was also compared to that of the whole range of data.

The estimated static model is evaluated against an autoregressive model predicting prices one-step into the future. Both an AR(1)- and an AR(15)-process is estimated to test the static forecasting model both using one lag and 15 lags, with the latter being the same amount of lags as used when estimating the forecast model. The estimated model is superior to a simple autoregressive model in all of the observed subsamples. The errors from the autoregressive evaluation can be seen in table VI. The values are compared to the values achieved by the fixed forecast and in all of the cases the forecasted model beats the AR-process.

Subsample 1					
Sample week	1	2	3	4	5
AR(1)	0.045	0.044	0.065	0.066	0.081
AR(15)	0.052	0.047	0.068	0.087	0.093
Model (fixed)	0.020	0.019	0.026	0.029	0.032
Subsample 2					
Sample week	1	2	3	4	5
AR(1)	0.020	0.028	0.012	0.045	0.083
AR(15)	0.016	0.029	0.038	0.088	0.084
Model (fixed)	0.009	0.009	0.008	0.013	0.019
Subsample 3					
Sample week	1	2	3	4	5
AR(1)	0.030	0.032	0.025	0.031	0.022
AR(15)	0.024	0.029	0.017	0.033	0.021
Model (fixed)	0.009	0.014	0.009	0.012	0.008

Table VI – The root mean squared errors from the autoregressive model estimated. All the values of RMSE are larger than the RMSE from the estimated forecasts. The estimated model forecast RMSEs are for a fixed estimation window but the other methods does not differ significantly in size of RMSEs.

When evaluating the model against a random walk, a hundred random walks are simulated to prevent the RMSE of the forecasts to be compared with one RMSE of a random walk, which is a value determined by chance. If only one value were to be simulated the probability of beating the model by pure chance would be to great. With a hundred simulations this chance is reduced significantly.



Graph V – Plotted random walks versus forecasted values during week three of subsample 2.

Graph V plots a hundred estimated random walks and displays the estimated root mean squared errors of the forecast model for week three, during subsample 2, using an expanding window method. The rolling and fixed window estimation is tangent to the expanding window and therefore they are not seen in the plot. The RMSE from the forecast is marked in red while the simulated random walks have blue colour. As can be seen, the RMSE is located in the right tail of the distribution making it significantly better than the random walk in this specific instance. It must be noted that this is the case for one week only, and for some weeks, the forecasted RMSE underperform the random walk. A summary of the results can be found in Table VII.

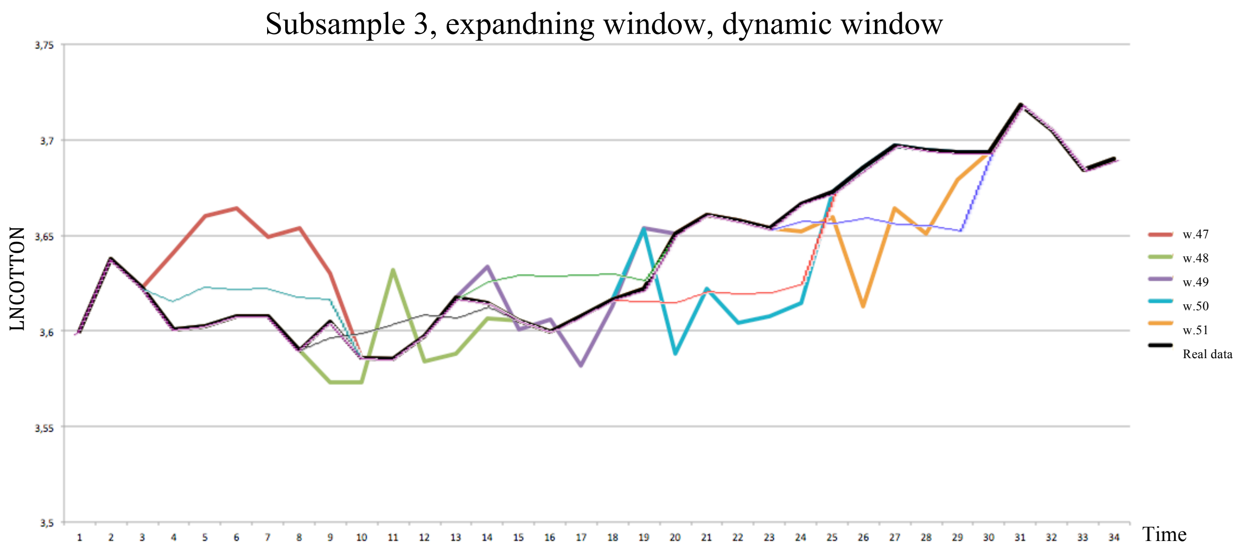
	Rolling	Expanding	Fixed
Static Method			
Sample 1	417/500	417/500	414/500
Sample 2	470/500	471/500	470/500
Sample 3	480/500	467/500	481/500
Whole Range	463/500	463/500	462/500
Dynamic Method			
Sample 1	327/500	337/500	309/500
Sample 2	371/500	376/500	378/500
Sample 3	386/500	386/500	382/500
Whole Range	338/500	344/500	342/500

Table VII – summarized results showing how many instances the estimated model beats a random walk in each sample. Each sample contains five weeks, and 100 random walks are estimated per week, which sums to 500 random walks per sample and estimation method.

As can be seen in Table VII the static forecasting method is statistically significantly better than a random walk in all but subsample 1 using a power level of 10 %. The results from the dynamic forecasts can be concluded that the performance is neither significantly better nor worse than that of a random model. If the requirement were to be relaxed to 15% significance would be found in both subsample 2 and 3 using any estimation method. Subsample 1 is a period characterised by a relatively quiet and calm cotton market where estimates tend to gain from predicting a mean trend evolution of zero, as a random walk is designed to do. In a quiet market, fundamental models cannot estimate parameters better than the random walk. This is because the mean of the cotton price is stable during the period and there is no need for estimated coefficients affecting the predicted price trend. The purpose of the estimated coefficients is to capture volatility and trends, which there are none in a quiet market.

The choice to include structural breaks and to split the sample into subsamples proved to be fruitful. Both subsample 2 and 3 provides more accurate results than the model estimated using the whole range. The single exception is subsample 1 that is the subsample that is characterised by the least volatile market, as argued. Therefore it is not surprising that the estimated model performs worst during this subsample. The best performing sample is subsample 3 that is also the sample that contains the largest deviations regarding the cotton price. The estimated model surpasses a random walk when it comes to predicting changing price trends, positive as well as negative.

Generally, the estimated model outperforms the random walk model when the market is volatile. A random model tends to stick to its mean value, by definition making a more stable model in a quiet market. The estimated model uses values for the variables looking back fifteen observations forcing it to include the previous volatility of the price. The model estimated using a VEC model does not fail during quiet markets but in some scenarios it tends to overestimate previous movements. In Graph VI the forecasts are plotted against a random walk forecast to show how they relate. However, it is of great importance to notice that the plot only shows one simulation of a random walk. Due to its nature it will change for every iteration and outperform the estimated dynamic model in some cases unless the estimated model is superior in 100/100 iterations, a case most unlikely and not observed in this thesis. Therefore the graph will tend to display different behaviour over each iteration, but the average tendencies will remain true.



Graph VI – Plotted forecasts vs. random walk, with actual values in black, forecasted values in thin lines and random walk in thick lines.

The static forecast is outperforming the dynamic model in all scenarios. This is not surprising because the static forecast only predicts one-step ahead while the dynamic forecast tries to predict values six days ahead. The static forecast model beats the benchmarks in a significant amount of cases providing a model that is robust in the short run, while the performance of the dynamic forecast is unstable in a greater degree than the static forecast.

The fact that the dynamic model could not beat a random walk in a significant amount of cases is not surprising because of the fact that price prediction is difficult. The short run model aims to forecast prices before the market can correct for mishaps. This explains in part how the model can beat the benchmarks in several cases still, even if it does not do so in a significant manner.

As the static model predicts the price fairly well, a sign predictor test is performed in addition to an actual price predictor test. If the model would be used for speculative purposes a one-day forecast where only the sign of the day-to-day difference is observed allows for short-term holdings. An investor would go short if the model predicts a decreasing price for the next day (a negative sign) and would go long if the model predicts an increasing price (a positive sign). The level of the price itself would not be of interest, only the direction of the evolution. The test that was conducted gave some interesting results. Subsample 1 is the subsample where the sign prediction shows a correct prediction in the highest amount of cases, followed by subsample 3 and subsample 2. The Whole Range proves to be suboptimal when forecasting for the direction of the price.

Subsample 1 is the sample that showed the poorest performance against the benchmark but it is the best performer in the sign prediction test. This could be explained by the construction of the tests. The benchmark test is sensitive to the size of the error whereas the sign prediction test only observes if the predicted price has the same sign as the actual value of the cotton price. When observing the plotted static price and forecast in subsample 1, it can be seen that the predictions are relatively far from the true value but it does follow it, providing a negative trend when a negative trend is present and vice versa (see appendix G). The forecast period in subsample 2 is highly volatile, but it contains a very stable estimation period. This gives the benchmark test a stable coefficient estimation that on average follows the trend of the price better than a random walk, but it misses the specific changes in direction in number of cases. The Whole Range sign test proves that coefficients using data including all the spikes and shocks in the sample will not predict the direction well at all.

The test shows that the static model can be used with some success for short holdings of cotton certificates and warrants. The method requires proper trimming of the estimation period used, so that it only includes relevant volatility and not historical data that is not significant any longer. This has been shown with the subsamples outperforming the whole range sign prediction in a decisive manner.

	Rolling	**Expanding**	**Fixed**
Sample 1	21/30	21/30	21/30
Sample 2	18/30	18/30	17/30
Sample 3	17/30	18/30	19/30
Whole Range	9/30	7/30	7/30

Table VIII - The results from the sign prediction test are displaying in how many instances the forecast model predicts the correct direction of the cotton price evolution one day ahead.

The results from the test proves that the model is fairly accurate when predicting the direction of the overall cotton price as long as the data is split into well motivated samples and not treated as a whole mass. A success rate of 50 % would be sufficient to deem a sign predictor as successful. The model has an average success rate of 63% over the three subsamples (not including the whole range), which is good.

If the model were to be revised using intraday frequency, the method of sign prediction could be used in day trading where several deals are struck within the same banking day using marginal changes in price in order to capture profits. This would increase the use of the model as a short-term speculative tool.

Different result could have been obtained by using other evaluation measures. In this thesis, RMSE is used because of its convenience when comparing different models. Though, the accuracy of the different evaluation measures such as MAE or MAPE, could have been included here as well. Apart from this, the benchmarks used are the most basic benchmarks and the results would possibly differ if more sophisticated benchmarks were used, such as an ARMA benchmark for example.

The dynamic forecast performs moderately well and with some adjustments, it could become a suitable model for weekly predictions. Minor alternations to the dynamic forecast model could improve the performance of the forecasts making it significantly better than the benchmarks, such as the inclusion of other variables unavailable at the time or using another frequency.

8. Conclusion

After evaluating the forecasts against the chosen benchmarks, no significant results were found concerning the dynamic, multi-period forecast. Having stated this, the dynamic forecast did not underperform either and with a slightly relaxed power requirement the results would have been significant. The static forecast results in a significantly better forecasting performance over the benchmarks, providing a model that can in some degree deliver a forecast over the one-day ahead price level of cotton. Using this result the model can be used in further applications such as short-term holdings of commodity derivatives for strict speculative purposes. The models use for commercial farmers when predicting future prices to reduce uncertainty is limited though, due to the failing performance of longer horizons. The choice of a VAR forecast proved to be motivated because of the sheer amount of coefficients estimated using lagged variables. The VAR model offered a suitable way to manipulate and arrange the chosen data without the need for a greater amount of repetitive actions. A sign prediction test was also conducted in order to see if the model could offer speculative traders a short-term model for day trading. This test proved that the model would be useful for predicting the direction of the cotton price for a one-day horizon, with proper choice of data range to exclude historical data that is no longer significant.

9. Future Research

It would be interesting to monitor the development of the cotton price a decade from now, considering the research that is taking place in the form of the recycling of textiles. Among others, Ekström (2012) states how the research in the field of transforming used textiles into ethanol and biofuels develops rapidly, and Jeihanipour and Taherzadeh (2009) described how it is possible to convert cotton-based waste to ethanol but how the recycling process needs pre-treatment in order to produce sustainable results. If this research would prove to be a success it would be of interest to observe how the acre allocation would evolve. If the transition to maize and sugar production would be reversed due to the higher supply of ethanol and greater demand for cotton that this would imply.

The method that is used when forecasting the cotton price in this thesis has been a linear estimation method through a vector error correction model. It would be of interest to analyse if a non-linear model would provide better and more accurate parameter estimates followed by a better forecast.

Ethanol is a fuel that has taken a much larger market share in recent years, partly due to strong governmental incentives from the US and EU and partly due to increasing fuel prices. In turn this leads to a change where farmer crop focus has shifted to ethanol producing crops (Mutuc et al., 2010). The production of various biofuels is a land-intensive production process that is in direct competition with cotton in the form of acre allocation. However, the trade of ethanol on commodity markets did not take place on a major scale until it was quoted on CBOT in 2005, and even then the liquidity is far too low to allow the variable to be included in our model (Dahlgran, 2010).

It would be of great interest to include some variables that has recently shown to be of great importance in the commodity markets. Variables of interest such as polyester and ethanol cannot be included in the analysis due to data shortage. Polyester futures are traded in the form of PTA and these contracts are fairly limited as they are traded on the China Chemical & Fiber Economic Information Network (CCFEI) and have been traded for just over two years, being introduced in 2011. Ethanol does not have a much longer lifespan being introduced on the Chicago Board of Trade in 2005.

10. References

Adom, P. K. (2013). Time-varying analysis of aggregate electricity demand in Ghana: a rolling analysis. *OPEC Energy Review*. **37**(1), 63-80.

Amann, G., Lehecka, V., Schmid, E. (2012) Does speculation drive agricultural commodity spot prices? *Jahrbuch der ÖGA* **22**

Andrews, D. W. K. (1993). Tests for Parameter Instability and Structural Change with Unknown Change Point. *Econometrica* **61**(4), 821-856.

Baffes, J. (2011). Cotton Subsidies, the WTO, and the 'Cotton Problem'. *The World Economy.* Washington DC: World Bank.

Baffes, J., and Haniotis, T., (2010). Placing the 2006/08 Commodity Price Boom into Perspective. Working paper No. 5371. Available at The World Bank: http://www-wds.worldbank.org/external/default/WDSContentServer/IW3P/IB/2010/07/21/000158349_20100721110120/Rendered/PDF/WPS5371.pdf

Brooks, C. (2008). *Introductory econometrics for finance*. (2nd edition). Cambridge England ; New York, Cambridge University Press.

Camacho, M., Pérez-Quirós, G. (2013). Commodity Prices and the Business Cycle in Latin America: Living and Dying by Commodities? CEPR Discussion Paper No. DP9367. Available at SSRN: http://ssrn.com/abstract=2224292

J. Campbell and Perron, P. (1993). A note on Johansen's cointegration procedure when trends are present. *Empirical Economics* **18**(4), 777-789.

Carter, C. A., G. C. Rausser, Smith, A. (2011). Commodity Booms and Busts. *Annual Review of Resource Economics* **3**(1). 87-118.

Carter C. A., Janzen J. P., (2009). The 2008 cotton price spike and extraordinary hedging costs. *ARE Update* **13**(2)**,** 9–11.

Chan, K., Hameed, A., Sie Ting, L. (2003). What If Trading Location Is Different from Business Location? Evidence from the Jardine Group. *Journal of Finance.* **58**(3), 1221-1246.

Cheung, Y., & Lai, K. S. (1993). Finite-Sample Sizes of Johansen's Likelihood Ration Tests for Conintegration. *Oxford Bulletin Of Economics And Statistics*, **55**(3), 313-328

Cipan, J. and Woshnagg, E. (2004). Evaluating Forecast Accuracy. *UK Ökonometrische Prognose no. 406347*. University of Vienna

Collingsworth, B. (2012). Using mean absolute error for forecast accuracy. Retrieved 2013-05-06 from: http://blog.canworksmart.com/predictive-analytics/using-mean-absolute-error-forecast-accuracy/

Commodity Research Bureau (U.S.) (2010). *The CRB commodity yearbook 2010*. Wiley; John Wiley [distributor].

Dahlgran, R. A. (2010). *Ethanol Futures: Thin but Effective? — Why?* Proceedings of the NCCC-134 Conference on Applied Commodity Price Analysis, Forecasting, and Market Risk Management. St. Louis, MO. Retrieved from http://www.farmdoc.illinois.edu/nccc134.

Danthine, J.-P. and J. B. Donaldson (2005). *Intermediate financial theory.* (2nd edition). Oxford, Elsevier Academic Press.

Datastream International. (2013). Available: Datastream International.

Dillinger, W. (1998). Brazil's State Debt Crisis: Lessons Learned. *Economica (National University of La Plata).* **44**(3), 109-143.

Dreman, D. N. and M. A. Berry (1995). Overreaction, Underreaction, and the Low-P/E Effect. *Financial Analysts Journal.* **51**(4), 21-30.

Ekström, K. M. (2012). Nätverk, Trådar och Spindlar, Samverkan för ökad återanvändning och återvinning av kläder och textil. *Vetenskap för profession.* **22** 2012

Enders, W. (2010). *Applied econometric time series*. (3rd edition) Hoboken, NJ, Wiley.

Fama, E. F. (1965). The behavior of stock-market prices. *Journal of Business* **38**, 34-105.

Fox, J. (2009). *The myth of the rational market : a history of risk, reward, and delusion on Wall Street*. (Reprint edition) New York, Harper Business.

Froot, K. A. and E. M. Dabora (1999). How Are Stock Prices Affected by the Location of Trade? *Journal of Financial Economics* **53**(2), 189-216.

Gellings, C. and K.E. Parmenter. 2004. Energy Efficiency in Fertilizer Production and Use. *Efficient Use and Conservation of Energy, Encyclopedia of Life Support Systems*. Eolss Publishers, Oxford, UK

Gerik, T. J., Faver, K. L., Thaxton, P.M., and El-Zik, K. M. (1996) Late Season Water Stress in Cotton: I. Plant Growth, Water Use, and Yield. *Crop Science* **36:** 914-921

Gil-Diaz, F. (1998). The Origin of Mexico's 1994 Financial Crisis. *Cato Journal* **17**(3), 303-313.

Gillson, I., Poulton, C., Balcombe, K., and Page S. (2004). Understanding the impact of Cotton Subsidies on developing countries. MPRA Working paper No. 15373. Overseas Development Institute

Giacomini, R. and H. White (2006). Tests of Conditional Predictive Ability. *Econometrica* **74**(6), 1545-1578.

Graham, B. (1965). *The intelligent investor : a book of practical counsel*. (3rd edition) New York, Harper & Row.

Hamilton, J. D. and National Bureau of Economic Research. (2009). Causes and Consequences of the Oil Shock of 2007-08. NBER working paper series no. w15002. Cambridge, Mass., National Bureau of Economic Research**:** Electronic resource.

Hertel, T. W., J. Beckman, et al. (2011). Commodity Price Volatility in the Biofuel Era An Examination of the Linkage Between Energy and Agricultural Markets. NBER working paper series no. w16824. Cambridge, Mass., National Bureau of Economic Research**:** Electronic resource.

Holmstrom, B. and J. Tirole (1993). Market Liquidity and Performance Monitoring. *Journal of Political Economy* **101**(4), 678-709.

Intercontinental Exchange (2012). Cotton No. 2. Retrieved 2013-04-11 from https://www.theice.com/publicdocs/ICE_Cotton_Brochure.pdf

International Cotton Advisory Committee (ICAC). (2012). Production and Trade Policies Affecting The Cotton Industry. International Cotton Advisory Committee.

Jansson, M. (2013). Guld och olja är alla indikatorer man behöver. *Fokus Råvaror*, Handelsbanken**:** **1** 2013.

Janzen, J. P., Smith, A. D., and Carter. C. A., 2012. Commodity Price Comovement: The Case of Cotton. Proceedings of the NCCC-134 Conference on Applied Commodity Price Analysis, Forecasting, and Market Risk Management. St. Louis, MO. [http://www.farmdoc.illinois.edu/nccc134].

Jeihanipour, A. & Taherzadeh, M.J. (2009). Ethanol production from cotton-based waste textiles. *Bioresource Technology*. **100**: 1007- 1010.

Keynes, J. M. (1936). *The general theory of employment, interest and money.* London, Macmillan.

Kurozumi, E., Tanaka S. (2010). Reducing the size distortion of the KPSS test. *Journal of Time Series Analysis*. **31** 415–426

MacKinnon, J. G. (1991). Critical values for cointegration tests. In R. F. Engle and C. W. J. Granger (eds), *Long run Economic Relationships: Readings in Cointegration*, **Ch. 13**: 267–76. Oxford: Oxford University Press.

MacKinnon, J. G. (1996). Numerical distribution functions for unit root and cointegration tests. *Journal of Applied Econometrics.* **11**: 601–618.

McGee, M. and Yaffee, R. A. (2000). *An Introduction to Time Series Analysis and Forecasting: With Applications of SAS and SPSS*. New York, Academic Press.

McKibbin, W., Martin, W. (1999). The East Asian Crisis: Investigating Causes and Policy Responses. Policy Research Working Paper 2172. Retrieved from The World Bank

Mutuc, M., Pan, S., Hudson, D., (2010) Response of Cotton to Oil Price Shocks. The Southern Agricultural Economics Association Annual Meeting, Orlando, FL. February 6-9, 2010.

Caffrey, J. Interviewed by Navarro, B., J. CNBC Halftime Report. CNBC, 2013. Web. Mon. 15 Apr. 2013.

Owsley, F. L. and H. C. Owsley (1959). *King Cotton Diplomacy* . (2nd edition). Revised by Harriet Chappell Owsley. Chicago, University of Chicago Press.

Pindyck, S. R., and Rotemberg, J. J., (1988). The Excess Co-Movement Of Commodity Prices. Working paper No. 2671. Retrieved from National Bureau of Economic Research http://www.nber.org/papers/w2671

Raıssi, H. (2012). Comparison of procedures for fitting the autoregressive order of a vector error correction model. *Journal of Statistical Computation and Simulation.* **83**(10), 1517–1529

Seung, B., Ahn, S.K. And Zadrozny, P.A. (2012). Estimation of vector error correction models with mixed-frequency data. *Journal of time series analysis.* **34**, 195-205

Sharma, R. (2012). *Breakout nations : in search of the of the next economic miracle*. London, Allen Lane.

Shiller, R. J. (2000). *Irrational exuberance*. (1st edition) Princeton, N.J., Princeton University Press.

Sørensen, P. B. and H. J. Whitta-Jacobsen (2010). *Introducing advanced macroeconomics : growth and business cycles.* (2nd edition). London, McGraw-Hill Higher Education.

Tse, Y. (2012) The Relationship Among Agricultural Futures, ETFs, and the US Stock Market. University of Texas at San Antonio, retrieved from umsl.edu: http://business.umsl.edu/seminar_series/Spring2012/DBA20.pdf

Van Aarle, B., Boss, M., Hlouskova, J. (2000). Forecasting the Euro Exchange Rate Using Vector Error Correction Models. *Weltwirtschaftliches Archiv*, **136**(2)**:** 232-258

Yafa, S. H. (2005). *Big cotton: how a humble fiber created fortunes, wrecked civilizations, and put America on the map*. New York, Viking.

Östensson. O. (2012). The 2008 commodity price boom: did speculation play a role? *Miner Econ* **25:** 17-28.

11. Appendix

Appendix A – Unit Root Test, Augmented Dickey-Fuller

Augmented Dickey-Fuller test

Null Hypothesis: Unit root (individual unit root process)
Automatic lag length selection based on SIC: 0 to 2

Method	Statistic	Prob.**
ADF - Fisher Chi-square	30.523	0.016
ADF - Choi Z-stat	-2.760	0.003

** Probabilities for Fisher tests are computed using an asymptotic

Chi-square distribution. All other tests assume asymptotic normality.

Table A.1 – Augmented Dickey-Fuller test. The column named Prob. displays the probability to that the variable is stationary.

Series	Prob.	Lag	Max Lag
LNCOTTON	0.121	0	31
LNCOTTON_FUTURE	0.243	0	31
LNGDP	0.360	1	31
LNMAIZE	**0.070**	**1**	**31**
LNOIL	0.233	0	31
LNSP500	**0.052**	**2**	**31**
LNSUGAR	**0.064**	**1**	**31**
LNWOOL	0.410	1	31

Table A.2 – Augmented Dickey-Fuller test. The column named Prob. displays the probability to that the variable is stationary. The bold rows indicate variables that reject the null hypothesis with a power of 10% but none at the 5% level.

Appendix B – Stationarity Test, KPSS

KPSS test

Null hypothesis: Variable is stationary
Bandwidth: 53 (Newey-West automatic) using Bartlett kernel

Variable	KPSS-stat
LNCotton	4.443
LNCotton Future	5.950
LNGDP	8.240
LNMaize	0.793
LNOil	5.775
LNS&P500	2.228
LNSugar	**0.462****
LNWool	1.980

Table B.1 – KPSS test. The column named KPSS-stat displays the statistic for evaluating if the variable is non-stationary when compared to the critical values in table B.2. The bold rows indicate variables that accept the hypothesis of non-stationarity with a power of 5% for two stars.

Asymptotic critical values*:	1% level	0.739
	5% level	0.463
	10% level	0.347

Table B.2 – KPSS critical values. The table shows the values, which the KPSS-statistics are to be compared to.

Appendix C – KPSS- and ADF tests

Stationarity tests

Null Hypothesis: Unit root (individual unit root process)
Automatic lag length selection based on SIC: 0 to 1

Method	Statistic	Prob
ADF - Fisher Chi-square	137.298	0.000
ADF - Choi Z-stat	-9.968	0.000

Table C.1 – Summary ADF-test for the included variables in first differences

ADF-test

Series	Prob.	Lag	Max Lag
DLNCOT	0.0001	0	31
DLNCOTFUT	0.0001	0	31
DLNGDP	0.0153	0	31
DLNMAIZE	0.0001	0	31
DLNOIL	0.0001	1	31
DLNSP500	0.0001	1	31
DLNSUGER	**0.0001**	**0**	**31**
DLNWOOL	0.0001	0	31

Table C.2 – Individual ADF-test. The interesting variable to observe is sugar in order to see if the non-stationary that is being displayed when Sugar was examined in level still consists in first differences. This is indeed the fact allowing for the variable to be included in first difference in a VEC model along with the other variables without further manipulations.

KPSS test statistic:		**0.048*****
Asymptotic critical values:	1% level	0.739
	5% level	0.463
	10% level	0.347

Table C.3 –KPSS test for the sugar variable in first differences.

Appendix D - Johansens Cointegration Test

Unrestricted Cointegration Rank Test (Trace)

Hypothesized No. of CE(s)	Eigenvalue	Trace Statistic	0.05 Critical Value	Prob.**
None *	***0.012***	***118.605***	***95.754***	***0.001***
At most 1	0.005	65.476	69.819	0.105
At most 2	0.003	41.615	47.856	0.169
At most 3	0.003	27.603	29.797	0.087
At most 4	0.002	15.440	15.495	0.050
At most 5 *	***0.001***	***5.654***	***3.8415***	***0.017***

Trace test indicates 1 cointegrating eqn(s) at the 0.05 level

* denotes rejection of the hypothesis at the 0.05 level

**MacKinnon-Haug-Michelis (1999) p-values

CE(s) = Cointegration equations

Table D.1 – Trace test. The bold rows indicate significant results regarding the number of cointegrating vectors in the data material. No specifications regarding which variables that are cointegrated are given. The test indicates that there is one cointegrating relation because the null hypothesis of no cointegrated relations are significantly rejected but the null hypothesis of one cointegration relation cannot be rejected.

Unrestricted Cointegration Rank Test (Maximum Eigenvalue)

Hypothesized No. of CE(s)	Eigenvalue	Max-Eigen Statistic	0.05 Critical Value	Prob.**
None *	***0.011***	***53.129***	***40.078***	***0.001***
At most 1	0.005	23.86	33.877	0.465
At most 2	0.003	14.01	27.584	0.821
At most 3	0.003	12.16	21.132	0.531
At most 4	0.002	9.786	14.265	0.226
At most 5 *	***0.001***	***5.655***	***3.841***	***0.017***

Max-eigenvalue test indicates 1 cointegrating eqn(s) at the 0.05 level

* denotes rejection of the hypothesis at the 0.05 level

**MacKinnon-Haug-Michelis (1999) p-values

Table D.2 – Maximum Eigenvalue test. The bold rows indicate significant results regarding the number of cointegrating vectors in the data material. No specifications regarding which variables that are cointegrated are given. The test indicates that there is one cointegrating relation because the null hypothesis of no cointegrated relations are significantly rejected but the null hypothesis of one cointegration relation cannot be rejected.

Appendix E – Andrew-Quandt Breakpoint test

Test for Whole Range

Null Hypothesis: No breakpoints within 15% trimmed data
Varying regressors: All equation variables
Equation Sample: 6/26/1995 12/31/2012
Test Sample: 1/21/1998 5/12/2010
Number of breaks compared: 3211

Statistic	**Value**	**Prob.**
Maximum LR F-statistic (2/16/2010)	463.466	0.000
Maximum Wald F-statistic (2/16/2010)	3707.724	0.000

Structural break detected at: 2/16/2010

Table E.1 – Andrew-Quant test for Whole Range. The statistic shows that the greatest F-statistic is given at 2/16/2010 indicating the greatest significance for a structural break in the sample. The data is then split before conducting the next breakpoint test. The sample is trimmed by 15%, removing the outer 15 % limits.

Test for subsample prior to 2/16/2010

Null Hypothesis: No breakpoints within 15% trimmed data
Varying regressors: All equation variables
Equation Sample: 6/26/1995 2/16/2010
Test Sample: 8/18/1997 12/04/2007
Number of breaks compared: 2687

Statistic	**Value**	**Prob.**
Maximum LR F-statistic (12/01/2003)	300.434	0.000
Maximum Wald F-statistic (12/01/2003)	2403.475	0.000

Structural break detected at: 12/01/2003

Table E.2 – Andrew-Quant test for subsample. The statistic shows that the greatest F-statistic is given at 12/01/2003 indicating the greatest significance for a structural break in the sample.

Appendix F – Granger Causality Test

Granger Causality test

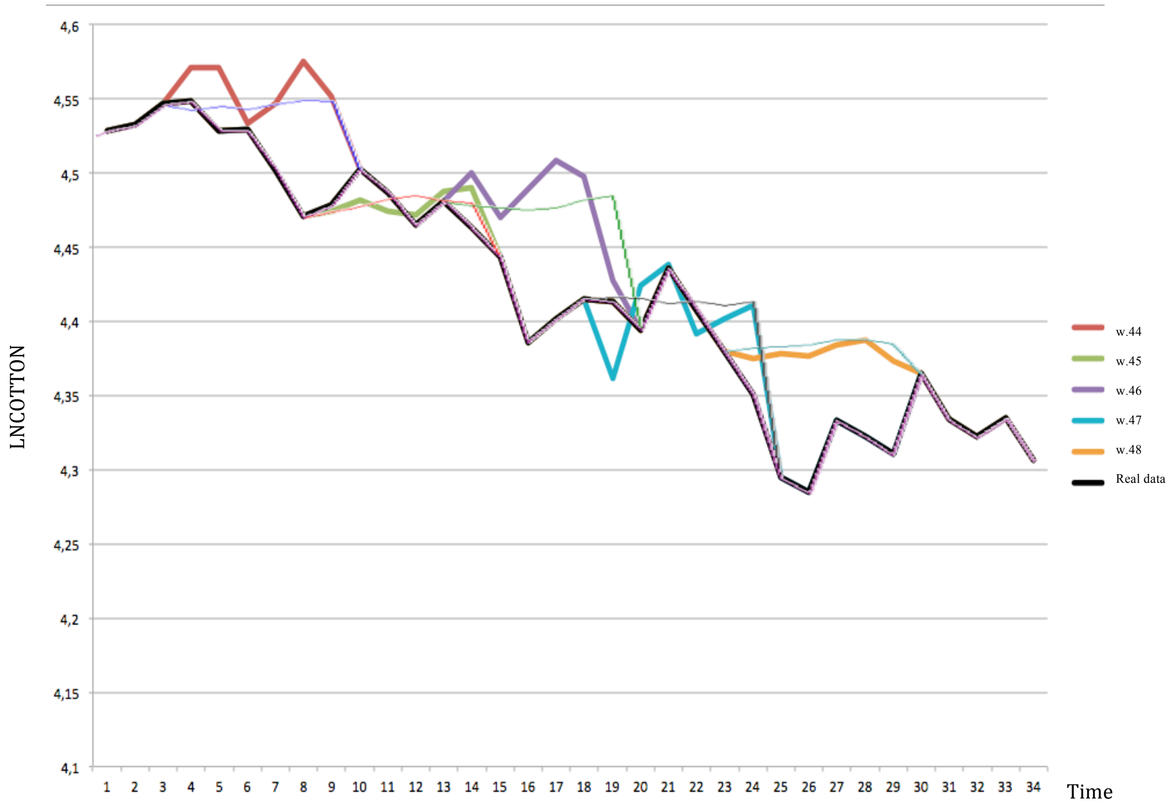
Sample: 6/02/1995 3/25/2013
Included observations: 4584
Dependent variable: LNCOTTON

Exogenous: GDP, S&P500

Variable	**Chi-sq**	**df**	**Prob.**
LNCOTTON_FUTURE	10.361	2	0.006
LNMAIZE	2.362	2	0.307
LNOIL	5.781	2	0.056
LNSUGAR	4.054	2	0.132
LNWOOL	1.284	2	0.526
All	**21.520**	**10**	**0.018**

Table F.1 – Granger Causality Test. The column named Prob. Displays the probability to accept the null hypothesis, that the variable Granger causes cotton. The variables as a group significantly Granger causes cotton.

Sample 1 - Dynamic Forecast Method

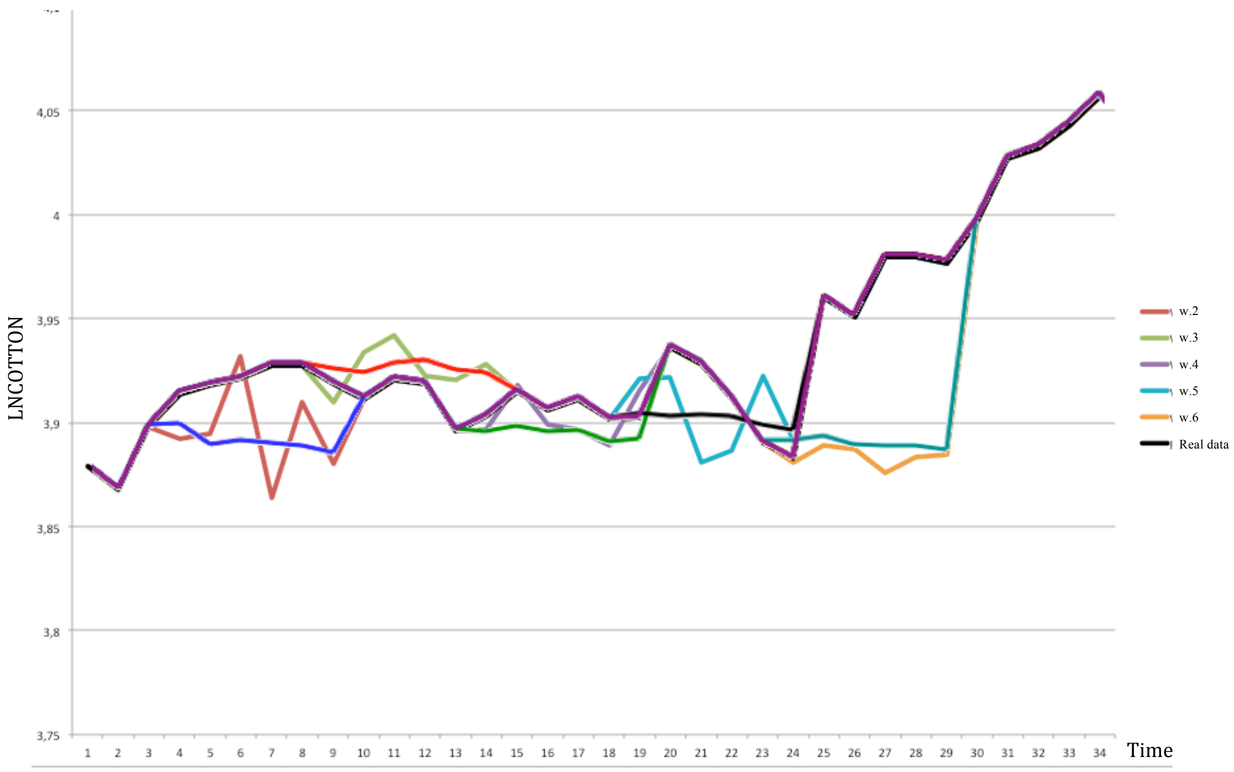


Dynamic forecast, estimation window: 1995 - 2003

	w. 44	w. 45	w. 46	w. 47	w. 48
Rolling	59/100	89/100	74/100	75/100	30/100
Expanding	59/100	90/100	82/100	79/100	27/100
Static	59/100	88/100	69/100	71/100	22/100

Graph and Table G.1 – The results from a dynamic forecast during subsample 1 using an expanding window, compared to the results from a random walk and the actual observed values. The thick lines represent the random walk, the thin line represents the result from the model, and the black line represents the actual data. The table shows in how many instances the estimated model beats the random walk model.

Sample 2 – Dynamic Forecast Method

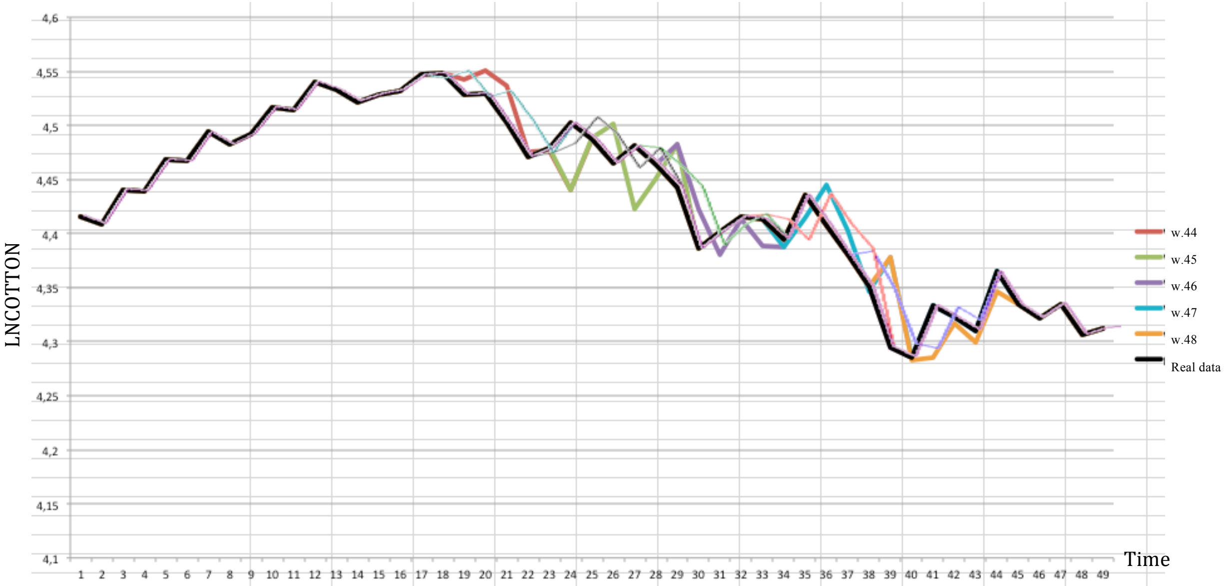


Dynamic forecast, estimation window: 2003 – 2010

	w. 2	**w. 3**	**w. 4**	**w. 5**	**w. 6**
Rolling	45/100	96/100	90/100	94/100	46/100
Expanding	45/100	96/100	90/100	94/100	51/100
Static	45/100	95/100	92/100	93/100	53/100

Graph and Table G.2 – The results from a dynamic forecast during subsample 2 using an expanding window, compared to the results from a random walk and the actual observed values. The thick lines represent the random walk, the thin line represents the result from the model, and the black line represents the actual data. The table shows in how many instances the estimated model beats the random walk model.

Sample 1 – Static Forecast method



The static forecast method is much more in line with actual observed values due to its iterated estimation process allowing it to in practice predict one day ahead. Therefore the graph gets quite messy and only sample 1 is displayed.

Static, estimation window: 1995-2003

	w. 44	w. 45	w. 46	w. 47	w. 48
Rolling	74/100	96/100	93/100	92/100	62/100
Expanding	74/100	96/100	93/100	92/100	62/100
Static	74/100	96/100	91/100	92/100	61/100

Graph and Table G.3 – The results from a static forecast during subsample 1 using an expanding window, compared to the results from a random walk and the actual observed values. The thick lines represent the random walk, the thin line represents the result from the model, and the black line represents the actual data. The table shows in how many instances the estimated model beats the random walk model.

Sample 3 - Dynamic Forecast Method, estimation window: 2010 - 2012

	w.47	**w.48**	**w.49**	**w. 50**	**w. 51**
Rolling	83/100	94/100	51/100	77/100	81/100
Expanding	83/100	94/100	48/100	85/100	76/100
Static	83/100	94/100	34/100	88/100	83/100

Table G.4 – displaying the amount of cases where the forecast beats the random walk in subsample 3, dynamic forecast.

Whole range – Dynamic Forecast method, estimation window: 1995-2012

	w.47	w.48	w.49	w. 50	w. 51
Rolling	71/100	79/100	89/100	57/100	42/100
Expanding	71/100	79/100	90/100	58/100	46/100
Static	71/100	79/100	90/100	58/100	44/100

Table G.5 - displaying the amount of cases where the forecast beats the random walk in the whole range sample, dynamic forecast.

Sample 2 – Static Forecast method, estimation window: 2003-2010

	w. 2	w. 3	w. 4	w. 5	w. 6
Rolling	98/100	100/100	100/100	100/100	72/100
Expanding	98/100	100/100	100/100	100/100	73/100
Static	98/100	100/100	100/100	100/100	72/100

Table G.6 - displaying the amount of cases where the forecast beats the random walk in subsample 2, static forecast.

Sample 3 -Static Forecast method, estimation window: 2010-2012

	w.47	**w.48**	**w.49**	**w. 50**	**w. 51**
Rolling	98/100	94/100	100/100	90/100	98/100
Expanding	98/100	80/100	99/100	92/100	98/100
Static	98/100	93/100	99/100	93/100	98/100

Table G.7 - displaying the amount of cases where the forecast beats the random walk in subsample 3, static forecast

Whole range – Static Forecast method, estimation window: 1995-2012

	w.47	w.48	w.49	w. 50	w. 51
Rolling	85/100	92/100	99/100	89/100	98/100
Expanding	85/100	92/100	99/100	89/100	98/100
Static	85/100	91/100	99/100	89/100	98/100

Table G.8 - displaying the amount of cases where the forecast beats the random walk in the whole range sample, static forecast