

Stockholm School of Economics

Master Thesis in Finance (4305)

May 14, 2012

Forecasting the Steel Billet Price via Steel Futures Contracts and Convenience Yield:

An Empirical Study on the Price Behavior of London Metal Exchange's Steel Billet Product.

Author: Isaac Zhou[1]

Abstract

This thesis studies the price behavior of a highly important yet unexplored commodity: the steel. The author refers to the theoretical models that are commonly used in other commodity researches: the unbiased futures model (UFM) and the convenience yield model (CYM). After relevant adjustments, these models are ready to empirically test the London Metal Exchange's (LME) steel billet prices. Expectantly, the results indicate that in the UFM model, the steel futures price is neither an unbiased nor an efficient predictor for the steel spot price in the future. Surprisingly, the results generated from the CYM model suggest that the convenience yield is not a considerably better alternative than merely using the futures price to predict the steel prices in the future. The author explains that these two models fail for both internal and external reasons. Internally, there are limitations embedded in the model setup and availability of data series, whereas the external reason is due to the heavy speculative activities in the commodity market during recent years. This thesis further advises that by using dynamic programming techniques, the structural model is a potential alternative for future studies to capture the salient behavior of the steel prices.

Key Words: Steel, Commodity Study, Unbiased Futures, Convenience Yield, Price Behavior

Supervisor: Associate Professor Paolo Sodini

Discussant: Alexander Bergendorf

Acknowledgement

This thesis owes a great academic debt to my supervisor Dr. Paolo Sodini, Associate Professor from Department of Finance, Stockholm School of Economics. His encouraging comments, invaluable insights and brilliant suggestions have remarkably supported my research on this niche subject of the commodity market. Meanwhile, I am also very appreciative of my alma mater, Stockholm School of Economics. Her constant intellectual challenges, rigid academic training and abundant opportunities get me well prepared for the road ahead.

[1] **40149@student.hhs.se**

Table of Contents

List of Figures i
List of Tables i
1 Introduction 1
1.1 Commodity: a New Asset Class 2
1.2 A Brief Review on Commodity Taxonomy 4
1.3 Background of the Steel Market 5
1.4 Originality and Purpose 6
1.5 Organization 7
2 Literature Review 8
2.1 Synthesis of Past Studies 8
3 Theoretical Framework 10
3.1 Unbiased Futures Model (UFM) 11
3.2 Convenience Yield Model (CYM) 13
4 Hypotheses 19
4.1 Hypothesis for UFM 19
4.2 Hypothesis for CYM 19
5 Methodology 21
5.1 Methodological Approach for UFM 21
5.2 Methodological Approach for CYM 22
6 Data 25
6.1 Data Choice and Motivation 25
6.2 Source and Reliability of the Data 25
6.3 Data Description 26
6.3.1 LME Steel Billet Spot Price 26
6.3.2 Contango vs. Backwardation 27
7 Empirical Results and Analysis 29
7.1 Results for UFM 29
7.2 Results for CYM 33
7.2.1 Results for CYM 3 Months 34
7.2.2 Results for CYM 15 Months 34
7.3 Out of Sample Prediction 35
8 Limitation 38
8.1 Model Limitations 38

8.2 Data Limitations 38
9 Conclusion 40
9.1 Comments on Speculation 40
10 Suggestions for Future Research 42
11 References 43

List of Figures

Figure 1-1 Global Metal Consumption in 2011 1
Figure 1-2 Comparison between commodity indices and inflation index 3
Figure 1-3 Comparison Among Commodity, Stock and Fixed Income Returns 4
Figure 6-1 LME Steel Billet Spot Price Series 26
Figure 6-2 LME Steel Contango 27
Figure 7-1 LME Steel Billet Log Price 29
Figure 7-2 LME Steel Billet Log Returns 30
Figure 7-3 Plot of ACF to check serial correlation 31
Figure 7-4 Plot of the Convenience Yields 33
Figure 7-5 Three Periods Out of sample prediction 36
Figure 7-6 Three Periods Out of sample prediction 37

List of Tables

Table 3-1 Cash Flow Explanation 15
Table 7-1 Augmented Dickey Fuller Test 30
Table 7-2 UFM Model Regression 3-6 31
Table 7-3 Wald Test: Unbiasedness Test Hypothesis 1-a 31
Table 7-4 Testing the efficiency of the UFM model: 32
Table 7-5 Ljung-Box Test 32
Table 7-6 Integration of order 1 Test 33
Table 7-7 CYM for 3 Months Equation 34
Table 7-8 Ljung-Box Test for CYM 3 Months 34
Table 7-9 Engle-Granger cointegration test 34
Table 7-10 CYM for 15 Months 35
Table 7-11: Ljung-Box Test for CYM 15 Months 35

1 Introduction

This chapter first points out the lack of academic attentions on some important commodities such as steel and iron ore. Then it provides the economic evidence and motivates that more researches should be conducted on the steel given its importance in the industry. Then the chapter continues to give a brief review of the commodity market and explains why commodities should be regarded as a separate asset class. Due to the diversity in the commodity and the steel markets, the author introduces how to differentiate these products. The purpose and originality of this thesis are provided at the end of this chapter.

During the past decade, the commodity market experienced the most dramatic price fluctuations in the known trading history. During the past five years, many commodity prices reached their historical high, for instance the gold bullion peaked at 1,898 USD per troy ounce in September 2011; WTI crude oil also topped at 145 USD per barrel in July 2008… (DataStream, 2012) On the other hand, since late 2008, the prices of most consumption commodities plummeted with the financial melt-down. This price instability has not only caused the panic in the marketplace but also drawn attention from practitioners. Gradually, the commodity investment has become an increasingly important component for many investors to consider in their portfolios. However, unlike the other more popular asset classes, i.e. the equity and the bond, the academic research on commodity investment is relatively insufficient.

In the already limited commodity studies, some particularly important products such as the steel and its input raw material iron ore are completely neglected by scholars. This lack of attention does not match the importance of the steel in the global economy. Notably, the global consumption of the steel during the past two decades has dwarfed that of any other metals (Coats, et al., 2008).

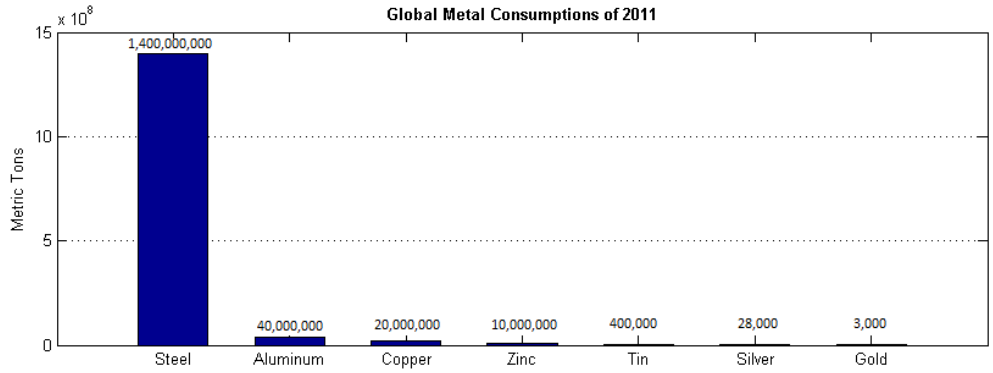


Figure 1-1 Global Metal Consumption in 2011

This figure compares the 2011 global steel consumption with that of other six important metals. It shows that steel consumption in 2011 is 1.4 billion Metric Tons, which literally "dwarfed" all the other six metals' consumption combined.

Figure 1-1 illustrates the 2011 global consumption of seven most important metals. As shown in the figure, the steel consumption is considerably greater than all the other six major metals combined! The historic mammoth steel consumption was mainly driven by the Japanese economic miracle in 1980s. Recently it is caused by the enormous infrastructure buildings in China and India. For instance, in 2008, the Chinese steel consumption is 58% of the global consumption. This percentage has grown ever since (Coats, et al., 2008). It is doubtful whether this trend will last concerning the recent slow-down of the Chinese economy.

Nonetheless, given its significant consumption level and daily transaction volumes, the failure to carefully study the steel price behavior will have profound consequences both in theory and practice.

On the other hand, the lack of academic research in steel related products is understandable: there are limited reliable data sets available. For example, up till 2012, there is still no futures market or standardized spot market for iron ore, an extremely important input raw material for steel production. At present, iron ore spot prices are primarily negotiated between buyers and oligopolistic miners such as BHP Billiton, Rio Tinto, Vale S.A. and so on. As a result, it is challenging to empirically study the real price of the iron ore in this opaque market. Fortunately, over the past few years, a variety of futures contracts for steel products were created. They were actively traded on major exchanges globally. Among them, the most prominent ones are London Metal Exchange's (LME) steel billet contracts and New York Mercantile Exchange's (NYMEX, now part of the Chicago Mercantile Exchange Group) hot rolled steel coil contracts. This thesis aims to pioneer at empirically studying the price behavior of the steel billet. I will test two most widely used commodity models on the LME steel billet price series. Before diving into the steel product study, I found it necessary to have a brief review on the commodity market in general.

1.1 Commodity: a New Asset Class

What exactly is a commodity, i.e. How to define it? People with different roles in society would have different answers to this question (Gemen, 2005). For example, economists would regard the commodity as a scarce resource whose allocation and exploitation will affect its supply-demand-inventory relationship, hence influence its equilibrium price; whereas bankers would focus more on the monetary numéraire features of the commodity. This feature explains why gold and silver are used to back currencies in the history, most notably the gold standard and the silver standard of the United States Dollars. Other participants of the society, such as engineers on an oil rig, politicians and ecologists…will have their own interpretations. This thesis is within the field of finance, so I lean toward the definitions provided by the economists and bankers: The commodity is a scarce resource with potential monetary values.

Conventionally, the commodity is not regarded as a financially friendly asset class, as opposed to other popular instruments such as a bond or a stock. One reason is that both the bond and the stock can be priced by the present value from a streams of *explicit* financial benefits, either interest or dividend payments. However this benefit stream is not so obvious for commodities. Accordingly, financial theories originally created for equity and fixed income investments are not directly applicable to commodities. For example, the Capital Asset Pricing Model (CAPM) claims that the investor's rewards are justified by the time value of the money and the risks embedded in his portfolio (Cochrane, 2005). However, this model cannot be extended to most of the commodity products without proper adjustments (Gemen, 2005). Many scholars believe that in order to correctly price commodities, researchers should focus on modeling the dynamic relationship among supply, demand and inventory (Pirrong, 2012). The theoretical frameworks and methodological approaches will be discussed in detail in following chapters. This section lays more emphasis on the commodity's cyclic features. Meanwhile, it helps the investors to better understand the commodity's economic natures, so that they can further utilize these attributes to optimize their portfolios for risk hedging and return maximizing purposes.

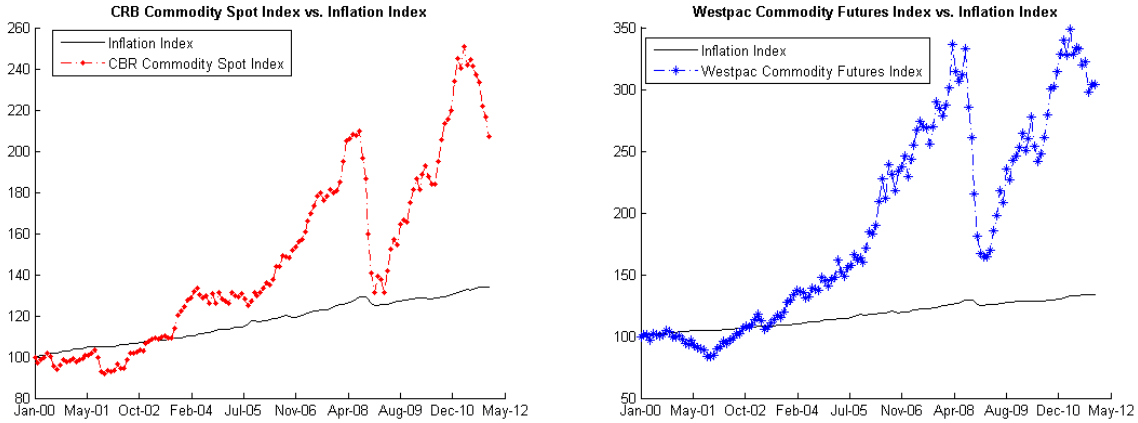


Figure 1-2 Comparison between commodity indices and inflation index

This figure compares the commodity spot and futures indices with inflation index during the past twelve years. It shows that both commodity indices surpass the inflation index considerably for the most of the time. Meanwhile, both commodity indices increase at much higher speed, however they are more volatile as well.

As shown in Figure 1-2, both commodity spot index (CBR Commodity Spot index, the dotted red line on the left part of figure 1-2) and commodity Futures index (Westpac Commodity Futures index, the dotted blue line on the right half of figure 1-2) have outstripped the inflation index (the solid black line), during the past decade. Plus, the inflation index is growing extremely slowly whereas the commodity spot and futures indices are increasing at a significant rate. This finding can be explained intuitively. During turbulent times, investors do not have enough confidence in the market; hence they prefer to either spend their money to consume more products or put their money in the investment commodities such as gold and silver. As a result, the commodity indices were driven up.

In practice, many fund managers have noticed these cyclic features of commodities and include them in their portfolios to hedge the risk exposures. During recent years, commodities have served mainly as a hedging alternative to provide diversifying benefits for other asset portfolios. Empirical evidences show that precious metals such as gold and silver are safe havens for the equity investments in times of financial turmoil. (Baur & Lucey, 2010) For hedge fund managers, merely hedging risk exposures is not enough; reaping lucrative profits from commodity trading is the goal.

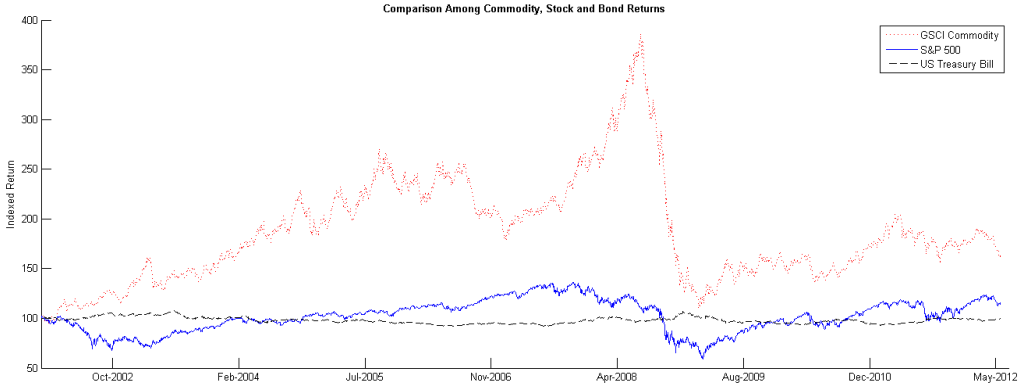


Figure 1-3 Comparison Among Commodity, Stock and Fixed Income Returns

This figure plots the indexed daily returns of three representative asset classes: commodity, equity and bond. The time spans from 2002 to 2012. It shows that commodity index (GSCI Index, the dotted red line) gains considerably higher return than the equity (S&P 500, the solid blue line) and bond (T-bond, the dashed black line) indices during the most recent decade.

In Figure 1-3, I benchmark the daily return of three indices starting in 2002: GSCI Commodity Index, S&P 500 and T-Bond. The figure indicates that over the past ten years, the commodity index return (GSCI Index, the dotted red line) have been extremely high compared to other financial indices: S&P 500 (the solid blue line) and US Treasury Bond (T-bond, the dashed black line). Some scholars might claim that the relatively low performance of the equity market during the past ten years is doomed by the aftermath effects of the Internet bubble burst (2001 to 2002), subprime crisis (2007) and Eurozone debt crisis (2011 till now). Meanwhile, people would also argue that the high commodity return is a result of "illiquid premiums", resulting in the yield differences between illiquid assets and more liquid assets. This argument is to some degree reasonable but not completely true. The illiquid premium explains the low return of T-bond because it is one of the most liquid assets. However, it remains questionable to account the return gap between the equity investment and commodity investment based on liquidity differences only. Many commodities such as gold and crude oil are fairly liquid compared to most of the equity assets. Moreover, illiquid assets are not sustainable in the long run and will usually die out due to lack of transactions as time passes. This is not the case of for commodities because they have been active for decades and will keep growing in trading volumes.

In 2010, most of the commodity hedge funds, or "commodity pools" in practitioners' term, earned double digit returns. Even though top generic funds (funds that mainly invest in commodity indices) reported heavy losses for 2011, e.g. Blenheim Capital lost 17% and Clive Capital lost 11% (Blas, et al., 2012), I believe these losses are caused by external factors such as Eurozone debt crisis and instabilities resulted from the political movements in the Middle East. This argument on externalities' impact is reasonable, because more specific commodity pools did not have any loss, but what's more they proved to be consistently profitable. For example, Galena Metals (a commodity pool specialized in metals and minerals) reported 11% annual profit in 2011. Thanks to its increasing trading volumes and activities, the commodity investment deserves more attention in the academic community. This thesis meets this demand. Like those successful commodity pools, benefiting from focusing on their expertise in a specific sector, this thesis concentrates on the steel billet product only.

1.2 A Brief Review on Commodity Taxonomy

The commodity market is highly diversified. Any one-size-fit-for-all attempt will fail in the commodity studies. This section uses a top-down approach explaining the varieties of the commodity products. The first and foremost fundamental distinction among commodities is the storability. Common commodities such as the metals and agricultural products are storable. On the other hand, more unorthodox products, such as weather, freight and carbon permits, are non-storable. Apparently, the subject of this thesis, the steel billet, is a storable commodity.

Another important divide is the seasonality. Some commodities are produced and consumed on a continuous basis; hence they are not subject to seasonal conditions (e.g. temperature, rainfalls and other adverse weather conditions…) Metals and the crude oil products are continuous commodities. Many agricultural products, such as coffee and frozen concentrated orange juice fall into the category of seasonal commodities. The steel is a continuously produced and consumed product.

For most agricultural commodities, there is another crucial factor concerning the sustainability. For example, both the crop and coffee are seasonal commodities, but their pricing dynamics are essentially different because the production of crops depends on farmlands whereas coffees grow on trees. Once the coffee tree is damaged, a new tree has to be planted. It will take time for the new tree to mature and get ready for production again. As a result, land crops are more resilient and flexible than tree crops against the adverse conditions. Similar effect is also found in crude oil wells. Some wells can be closed and reopened again for future productions, but others, for instance, the stripper well, once closed, cannot be reopened (Litzenberger & Rabinowitz, 1995). For steel products, this factor does not affect their production directly, though it is a big issue for the raw materials such as iron ore and coking coal. However, as long as the established value chain is not completely damaged, the steel producers can conveniently switch to alternative raw material suppliers if such supply shock occurs. For my thesis, this factor will not be considered to price the steel.

Finally, the purpose is another important way to define the diversity of commodities. For example, metals can be further divided into two major sub-categories: products such as the steel, zinc and copper are mainly for manufacturing purposes. They are consumption metals. Whereas, gold, silver and platinum are held more for the purpose of investing, they are investment metals (Hull, 2012). The price behavior and economic implications between consumption assets and investment assets are remarkably dissimilar. For example, the gold is a precious metal. Its production is much less than its inventory level due to its scarcity and investment purposes. Whereas, the steel is a consumption product, its production is much higher than its inventory level due to heavy consumptions. Consequently, these two types of metals should be treated separately.

All of these diversities create the difficulties to price the commodities. They also imply the inappropriateness of studying different commodities as a whole. Varieties in the commodity products reflect different factors, which will fundamentally define the commodity pricing dynamics. The commodity diversity also imposes great challenges to set up theoretical models. Accordingly, the one-size-fit-for-all approach adopted by many existing papers is both incorrect and unattainable. Bearing this in mind, I will only study the price behavior of the steel billet traded on London Metal Exchange (LME).

1.3 Background of the Steel Market

Just as there are so many varieties in the commodity market, the diversification in steel products is also quite sizable. Most obviously, the steel products can be firstly grouped by their shapes: flat, long, tubes, pipes and semi-finished. Two widely traded products are London Metal Exchange's steel billet, which is a semi-finished product, and New York Mercantile Exchange's hot rolled coil, which belongs to the flat type. In industry, practitioners also pay great attentions to the products from different engineering processes. For example, the flat steel is further categorized into hot rolled coil (HRC), cold rolled coil (CRC), hot-dipped galvanized (HDG) and electro-galvanized (EG)… the engineering techniques are reflected in price differences as well. For instance, per metric ton, the HDG is usually 200 USD more expensive than the HRC, because the HDG product requires additional zinc layers surfaced on either side (single-sided HDG) or both sides (double-sided HDG). As this thesis is in the finance field, the engineering and technical details are beyond the scope of my research, but it is true that no two steel types are priced same.

Before a reliable and organized steel market was introduced, the steel prices were determined by the negotiation between steel buyers and oligopolistic steel mills such as Arcelor Mittal, US Steel, Nucor, Severstal etc. This creates significant discrepancy to the steel prices. Similarly, this inefficient situation still exists in other important commodity products, such as the iron ore. The creation of the steel spot and futures price markets brought price transparency and promoted virtuous competition in the industry. The London

Metal Exchange (LME) and New York Mercantile Exchange (NYMEX, part of CME Group) are two notable exchanges committed to this endeavor. The thesis choses to study the price behavior of LME steel billet, because it is the first steel futures contract in the market. A more detailed data explanation can be found in chapter 6.

In this section, I will conclude with a brief illustration of the vertical integration structure in the steel market. The vertical integration describes a firm's position in its value chain. As for the steel market, the steel manufacturers are fit in between their upstream suppliers and downstream buyers.

The upstream suppliers provide input raw materials such as the iron ore, coking coal, scrap, lime stones and so on. As explained before, this market, especially the iron ore marketplace, is highly muddy and inefficient. Heavy speculations will distort the steel price eventually. For example, during the past few years, China is alleged of engaging in overstocking the iron ore inventory. In 2011, the China's iron ore imports accounted for 75% of the global iron ore trade (Iron Ore Team, 2011). It is estimated that there are over ninety million metric tons of iron ore materials idly stacked in major Chinese sea ports. This is the main reason that the price of the iron ore was abnormally high during the past years. The anomalous price of the iron ore eventually drove up the prices of the steel products artificially. Given the fact that China is one of the largest steel producers, the high steel price is beneficial to the county's export. Hence, it is not hard to understand the incentive for China to manipulate the steel price via overstocking the iron ore products.

On the other hand, the downstream buyers create markets for the steel products. Most of the steel buyers are industrial manufacturers. One of the most prominent steel buyers is the automotive producers, such as Toyota, General Motor, BMW, Ford and so on. On average, steel takes around 50% of the car production, ranging from the car frameworks to the engine valves. The steel purchasing department is an extremely important division for an auto manufacturer both financially and strategically. Steel buyers from this department usually have dual missions: they need to effectively negotiate the steel prices with large producers such as Arcelor Mittal, ThyssenKrupp, Tata Steel etc. so that the annual purchase value will be within the allocated budget. Meanwhile, they should be tactical enough to not cut price too much and keep good relationships with the steel suppliers for sustainable collaborations in the long run.

1.4 Originality and Purpose

Steel is a stepchild in the commodity research. There are some empirical attempts to study the price behavior of the metals (Ng & Pirrong, 1994); however the steel products remain untouched. This obscurity reflects the niche role of the steel price studies as compared to other more popular commodity products researches. As mentioned before, the production and consumption of steel is enormous. The current volume is mainly driven by the rapid development and infrastructure buildings in emerging markets, particularly in Brazil, Russian, India and China (BRIC). Heretofore, no academic work is found to study the price behavior of the LME steel billet price series. Generally speaking, no quantitative empirical work in the steel product is ever published before this thesis. The author takes on two popular models in commodity research: the unbiased futures model (UFM) and the convenience yield model (CYM). Most existing commodity researches use only one model. They fail to apply both and compare with each other. My thesis will develop both models and match their performances at predicting steel spot prices. Meanwhile, most of the extant papers on the commodity are restricted by the data frequencies. Monthly and weekly data are most common frequencies. Low-frequent data are not convincingly sufficient at explaining the real price behavior. Unlike its lower frequency counterparts, this thesis distinguishes itself by using the high frequency daily price series.

Even though high frequency data usually shows volatility clustering and introduces more serial correlations, but it provides a holistic view and captures more features of the steel price trajectory.

From a more pragmatic perspective, the steel is becoming a rising star for global commodity investors. Major investment banks, such as JP Morgan Chase, Macquarie and HSBC devote heavily to their steel analyses from more practical perspectives. They have established the leadership in steel trading practices. As mentioned in section 1.2, commodity pools with specific knowledge in steel products, such as Glena metals, drew great attentions for their stellar performance during the crisis. Furthermore, consulting firms with an expertise in steel sourcing strategies gain great respects and businesses from major steel buyers such as automotive manufacturers, technology and construction companies. This thesis provides a theoretical guidance for steel practitioners to consider during their trading and industrial purchasing activities. Hence, I envision three types of audiences: commodity researchers, commodity pool traders and operational consultants dealing with the steel sourcing strategies.

1.5 Organization

The rest of the paper flows as this: chapter 2 provides a general literature review of the existing researches on commodities. From these studies, I extract two commonly used models: the unbiased futures model (UFM) and the convenience yield model (CYM). I further develop and adjust both theoretical frameworks to fit in the steel context in chapter 3. Chapter 4 states the hypotheses for both models. Chapter 5 explains in detail how the models are set up and implemented in Matlab. This chapter also reveals some econometric considerations and provides relevant solutions. Chapter 6 describes the data of choice: LME steel billet spot and futures price series of daily frequencies. Chapter 7 presents the empirical results and detailed analysis. Chapter 8 draws the conclusions and implications of the thesis. Chapter 9 articulate the limitations embedded in the models and data. Chapter 10 provides suggestions for further researches.

2 Literature Review

This section gives a summary of the existing works on commodity products. It first points out that the commodity research is not a focal point of the academic finance study. The chapter continues to present two distinctive spectrums of approaches: the theory of storage and the structural model approach. The historic developments and recent situations of both approaches are illustrated. This thesis will apply the theory of storage approach only.

It is a well-known fact that the commodity research is not a mainstay in academia. (Pirrong, 2012) Most of existing works are published on the less recognized journals of agricultural, industrial and energy economics. Only a few articles made it to more acknowledged finance and economics publications, such as *Journal of Economics* (Pindyck, 1994), *Journal of Political Economy* (Hong, 2000) or *Journal of Finance* (Jagannathan, 1985). The publication of commodity studies on more specific journals is reasonable; as for the most of cases, the one-size-fit-for-all attempt would fail, especially in the highly diversified commodity markets. However, the niche role of commodity researches does not grant the right to ignore its importance.

2.1 Synthesis of Past Studies

There are two schools of thinking in the commodity research. One is the *theory of storage*, which analyzes the incentives for commodity investors to hold inventory. It also explains the relationship between inventory and prices. This theory introduces the term of "convenience yield", which is a stream of implicit financial benefits to the commodity investors. Researchers from this school attempt to create the analogy between the commodity convenience yield and the dividend of a dividend-paying stock or the interest payment of a non-zero coupon bond. By doing so, commodity pricing can be fit into the present value model as used in the equity and fixed income studies. The other school takes on a structural model with dynamic programming techniques. Scholars from this school tend to model the pricing dynamics of a commodity and mimic its salient price behaviors in a highly theoretical context. In additon to these two schools of thinking, there is another way of thinking by default. It uses the futures price as an *unbiased* expected value for the commodity spot price in the future. This approach is not formulated as a school of theory due to its simplicity and straightforwardness, however it is the most direct and intuitive way to start, as the case of this thesis.

The theory of storage can be traced back to 1930s. Working (1933) and Working (1934) firstly study the price relationships among grain futures contracts of different terms traded in Chicago. He finds out that the spread between the spot and futures prices usually changes with the inventory level of the grains. Inspired by Keynes' *Treatise on Money*, Kaldor (1939) brought forth an early version of the storage theory based on his discovery that "those holding commodity stocks receive an implicit benefits from holding inventories." Working (1949) elaborated on this concept and introduced the term of "necessary return for storage," which can be regarded as the primitive version of the concept "convenience yield." This flow of study continues with works by Telser (1958), Williams (1986), Brennan (1991), Pindyck (1993) and most recently Knetsch (2007). The convenience yield is one of the most popular concepts in the commodity research.

The other school of research adopts the structural model. This type of school usually involves highly numerical methods. It models fundamental factors that will affect commodities's prices, such as supply, demand and inventory. Gustafson (1958) first attempts to gauge the optimal amount of inventory for agricultrual grains under the pricing constraints. Scheinkman & Schechtman (1983) further develop a partial equilibrium model to study the relationship between production and storage. Williams and Wright (1991) enrich the structural model approach by incorporating the numerical solutions to a wide range of storage problems in the commodity marketplace. However, the works on the structural model approach is far from perfection. As Pirrong (2012) points out, for this spectrum of thinking, the theory is relatively refined but the

empirical tests are heavily lagged. Due to its mathematical and computational limitations, this thesis will not consider this approach at this stage.

It is worth mentioning that there are some criticisms against the research with a focus on the convenience yield. For example, Williams and Wright (1991) argue that most of the papers using the convenience yield model ignore the market equilibrium dynamics of each individual commodities, i.e. the failure of one-size-fit-for-all approach. Pirrong (2012) believes that the convenience yield is a concept which usually fails to explain the salient commodity market price behaviors, and calls it a "dead end". Empirically speaking, it is true that convenience yield usually performs unsatisfactorily compared to the results from the structural model. More powerful method such as the Kalman Filter usually gives better performance at fitting models and predicting prices. However, I believe that the convenience yield is still valuable at explaining the price behavior from the theoretical perspective. Plus, convenience yield is a more accessible concept which is easier to understand and set up. On the other hand, the structural model approach adopting more advanced techniques needs to include highly relevant, yet extremely complicated, parameters. It is an utterly challenging task since most of the relevant parameters are not observable at all. (See Schwartz & Smith (2000)) Additionally, the calibration and simulation involved in the structural model approach usually come at a huge computational and time cost. These challenges are commonly referred as "the curse of dimensionality" by computer scientists. This thesis will focus on more fundamental models and empirically test them on the LME steel billet data. In next chapter, I will present two most popular theoretical models summarized from the received literatures.

3 Theoretical Framework

Following the theory of storage school of thinking, this chapter lays out the detailed theoretical foundations. Firstly, the chapter argues why random walk hypothesis is not appropriate for the commodity research. Then it presents in depth two theoretical models used in this thesis: the unbiased futures model (UFM) and the convenience yield model (CYM).

Similar to any other class of asset, the commodities' pricing practice should be defined by its fundamental economic dynamics from the market, such as the supply, demand and inventory level (Knetsch, 2007). Unlike other class of assets, e.g. stocks, bonds or derivative contracts, commodity pricing usually lays considerable amount of stress on physical storage. It is not hard to understand that the commodity price and inventory levels are dynamically intertwined. Essentially speaking, commodity price is a reflection of resource allocation decision, i.e. to store or to consume. Both the storage and the consumption of the commodity are associated with opportunity costs. Intuitively put, if an automotive manufacturer owns some metric tons of steel coils, it can either consume it today for production or keep it for future use. The opportunity cost of consumption today is the value of these steel coils in the future, which is uncertain in a volatile marketplace. If it decides to store, the manufacturer needs to reconsider the foregoing revenues to be generated (Provided that the manufacturer could produce and successfully sell the car). Plus, it should also consider the costs of storage, insurance, logistics and maintenance. The relationship between the steel price and steel inventory is a dynamic process. The steel price depends on how much the manufacturer consumes today; how much the manufacturer can consume today depends on the steel inventory level at present; the current inventory level depends on manufacturer's estimation of the steel price in the future; the steel price in the future depends on how much the manufacturer can consume in the future…this process goes on and on. Accordingly, the manufacturer needs to take a closer look at the inventory levels, opportunity costs and price estimations to have a better grasp of the development of steel prices.

In the received commodity pricing literatures, there are two distinct beliefs in forecasting commodity prices. The first is based on the random walk hypothesis, indicating that commodity market is not predictable at all, so the best shot to guess the commodity price is its current price level. For example, Larson (1960) tests Holbrook Working's (1958) thoery of anticipatory prices on the corn futures and implies that corn prices move randomly. Similar statement has been found that a random walk is consistent with the price behaviors of the commodities. (Cargill & Rausser, 1969) However, most of publications under this belief are quite old and become obsolete. The random walk hypothesis has been constantly challenged by modern scholars, who believe that the commodity prices are predictable as long as the market information is complete and can be properly processed. I believe that the random walk theory is not an appropriate hypothesis of commodity pricing for following three reasons:

Firstly, the theoretical assumptions for the random walk is too strong, i.e. the market should be efficient and comprised of well-informed rational participants. Plus, at any poin in time, the prices need to be set to accurately reflecting all the current information, which is "almost freely available to all participants" (Fama, 1965). Apparently, the commodity market fails to meet these rigid requirements in general.

Secondly, from an empirical perspective, the commodity market does not behave randomly, especially if we take into consideration of the speculations. If the random walk holds, then the past prices, even if they are speculative, do not impact today's prices, or the development of the future prices. As a result, if the random walk is true, then the commodity traders should develop their strategies without looking at historical price patterns, nor should regulators worry about the existence of speculations in the market because the past speculative price distortions will not impact the "fair" prices in the market today. This is both inappropriate

and incorrect. In the real world, professional traders, whether "fundamentalists" or "technicians", share the common belief that there are trends in the market (Sidney, 1961). Meanwhile, throughout 2008, most of the commodity prices went through their ceilings. The most well known fact negating the random walk theory in the real market is in July 2008, the WTI crude oil price was at historically high reaching over 145 USD per barrel and radically dropped to 33 USD per barrel in december 2008 (IndexMundi, 2012), which is a clear evidence of heavy speculation. The impact from this speculation lasts for months into 2009 as well. Similar speculative situation is discovered in the steel market too.

Thirdly, as Leuthold (1972), Stevenson and Bear (1970) point out that there are some major flaws in the statistical methods of many papers based on the random walks. Selection bias is one of them. For example, in Larson (1960)'s paper on the random process of the corn prices, he uses a whole ten-year period of monthly data. Doing so is equivalent of treating the whole price series, spanning over 10 years, without taking into accounts of business cycles, supply and demand shocks, seasonal effects of the agricultural products, droughts and adverse weather... So, Larson's(1960) model brings in substantial data selection biases. In my thesis, the steel data is daily and covers a four years time span from 2008 to 2012.

Contrary to the random walk hypothesis, this thesis will follow the path that commodity prices are predictable. Within this scope, I will test two popular models and study the price behavior of the London Metal Exchange's (LME) steel billets. The models are: the unbiased futures model (UFM) and the convenience yield model (CYM). In the rest of this chapter, I will explain them respectively.

3.1 Unbiased Futures Model (UFM)

Theoretically, the commodity futures price is considered as a forecasting indicator for spot prices in the future. The commodity futures price is supposed to provide fully informative evidences about the price behavior of its underlying. The futures price is also considered to provide indications about the future economic activities of the underlying asset. The reason is quite intuitive and straightforward: the futures prices reflect the price level on which both the buyers and sellers agree upon the delivery date of the commodity. Ideally speaking, steel futures prices should provide direct and complete information about steel market participants' expectations over the steel price in the future.

Unfortunately, as for other asset classes, using commodity futures prices as the only unbiased predictor for the future commodity spot prices always fail to generate optimal and convincing results. In reality, it is always the case that the contracted futures prices are either higher or lower than the commodity spot prices on the delivery date. This gap between the commodity's futures price and its future spot price is referred as the risk premium, which is caused by the pricing uncertainties. If the risk premium is persistent, then the futures price is not an unbiased, nor an optimal, predictor of the future spot price. In extant commodity researches, scholars attempt to form a widely accepted rejection against unbiasedness of the commodity futures price; however, the consensus is not reached yet.

The disputes are found most fierce in the energy commodity sector, where a number of empirical efforts are made to question the unbiasedness hypothesis of the futures price. For example, Moosa and Al-Loughani (1994) test on the monthly observations of the WTI crude oil prices and reject the hypothesis that futures prices are unbiased predictors of spot prices. Coimbra and Esteves (2004) also conclude that at least for short horizons (less than 12 months), crude oil futures prices do not perform well at forecasting oil prices. Similarly, Pagano and Pisani (2009) also detect serious forecasting errors in oil futures. However, these results are not aligned with the conclusions from others. For instance, Chernenko, et al. (2004) find little evidence to

reject the unbiasedness hypothesis of futures prices for crude oil and natrual gas. Chinn, et al. (2005) also fail to reject futures prices as unbiased predictors of crude oil, gasoline and heating oil. Wu (2005) suggests that raw futures prices are unbiased predictors of future oil prices.

One of the motivations of the unwillingness to reject the futures unbiasedness hypothesis is that futures price plays as an important role pertaining information for spot prices. Suggested by Hong & Yogo (2011), futures prices indeed carry important indications about the future economic activities of the underlying assets, thus should not be totally abandoned. In his paper, Grossman (1977) defines futures market as a platform for information exchange because futures prices reveal to less-informed market participants about the next period spot price expectations from more-informed traders.

However, "being unbiased" is different from "being important." Before setting up an unambiguous model, certain terms need to get clarified. The central concept is the unbiasedness. As described before, in an efficient market, the futures price should be used as the unbiased predictor of the future spot price. However, in the real world, there exist risk premiums, explaining the gap between the futures price and future spot price. This risk premium is caused by the forecasting errors from the pricing uncertainties of the underlying assets. In a risk-netural world, this risk premium is expected to be zero. So the unbiasedness of the futures price should be defined as:

$$E(S_{t+T}|I_t) = F(t,T) \tag{3-1}$$

In equation 3-1:

- $E(\cdot)$ is the expected value operator
- S_{t+T} is the commodity spot price at time $t+T$
- I_t is the set of all available information at time t
- $F(t,T)$ is the commodity futures price at time t for the delivery in T period, i.e. delivery at t + T

Assuming rational expectations, equation 3-1 can be transferred as:

$$S_{t+T} = E(S_{t+T}|I_t) + u_{t+T} \tag{3-2}$$

In equation 3-2:

- u_{t+T} is the white noise

Based on equation 3-1 and 3-2,

$$S_{t+T} = F(t,T) + u_{t+T} \tag{3-3}$$

In the real world, there are risk-averse investors. Those investors prefer an equally profitable investment at a lower risk. Accordingly, uncertainties caused by the forecasting errors should be included in the above system to reflect the risk premium. This can be done by adjusting equation 3-1 into:

$$E(S_{t+T}|I_t) = F(t,T) + \rho_t \tag{3-4}$$

In equation 3-4:

- ρ_t is the risk premium term, which varies through time

The risk premium term can be further defined as:

$$\rho_t = \rho + \varepsilon_t \tag{3-5}$$

In equation 3-5:

- ρ is a constant, representing the mean value of the risk premium.
- ε_t is a white noise.

Combine and refine equation 3-3, 3-4 and 3-5 to generate the unbiased futures model (UFM):

$$S_{t+T} = \rho + \beta F(t,T) + \epsilon_{t+T} \quad 3\text{-}6$$

In equation 3-6:

- ρ, is the risk premium term.
- β, is the unbiasedness coefficient.
- $\epsilon_{t+T} = u_{t+T} + \varepsilon_t$, is the error term. Since both u and ε are white noises, their sum ε is also a white noise. This term reflects possible price shocks caused by new information etc.

In equation 3-6, if the futures price is an unbiased indicator, then:

- $\rho = 0$, meaning there will be no risk premium, hence no uncertainty caused by the pricing risks.
- $\beta = 1$, meaning futures price will unbiasedly and completely predict the future spot price.
- $E(\epsilon_t \epsilon_{t+i}) = 0$, for any $i > 0$. This term reflects the efficiency of the UFM model. If all the future error terms are jointly expected to be zero, it means new information is not expected to cause the change of the future spot price.

The first two points define the unbiasedness of the futures price; the last point defines the efficiency. If all the above three requirements are fulfilled, future spot price upon delivery will be exactly the contracted futures price. As a result, UFM is an unbiased and efficient model to predict the future spot price of the commodity.

The focal point of UFM model is to question whether the futures prices are unbiased (and efficient) estimators of the future spot price, rather than cast doubt on the futures price's role as an important source of information for spot price forecasts. Rejection to UFM does not indicate that futures prices should be excluded from forecasting commodity spot prices in the future.

Heretofore, most of the research have predominantly tested similar unbiased futures model predominantly on enegery commodity products, as they are among the most liquid commodities. No such work has been done for the steel products. Moreover, most of the past empirical works adopt monthly data. This thesis fill the gap to both test the model on the steel billet and adopt the high frequency (i.e. daily) data.

3.2 Convenience Yield Model (CYM)

As discussed before, there is no widely acknowledged rejection against the unbiasedness for the commodity futures price to predict future spot price, however in academia UFM fails to fulfill the scholars' expectation on the theoretical level. A solid and well-polished theory should be able to generate the results that can satisfactorily capture and explain most features of the observed data from the real world. Despite of its useful and important implications, UFM is not a convincingly rock-hard theoretical model to serve as the monopolized approach to pricing storable commodities.

Noticing such phenomena, economists such as, Holbrook Working and Nicolas Kaldor introduced the concept of the convenience yield during 1930s and 1940s. The term "convenience yield" originates from the theory of storage, which aims to explain the difference between spot and futures prices by rationalizing the reasons to hold inventories. As explained in the introduction section of this chapter, the inventory level of the storable commodity has intertwined impacts on the trajectory of commodity price. On the other hand, the commodity price also provides guidance to the allocation of its inventory, which will in turn affect the price. The convenience yield reflects this relationship between the commodity price and its inventory level. However, it is necessary to point out that the proposed convenience yield model in this thesis is only applicable for those commodities that can be physically produced, stored and transported on a continuous basis, for example the subject of this thesis steel billet fulfills this requirement. Seasonally produced commodities such as agricultural products need adjustments to reflect their seasonal production scheduling. For details on how to price this type of commodity, refer to chapter six of Pirrong (2012).

Before setting up the model, let me define the convenience yield first. The intuition of the convenience yield is not hard to understand. The commodity holders receive a stream of "implicit benefits" by keeping a certain level of inventory. These benefits are reflected in multiple ways, such as to meet unexpected extra orders from the clients, to avoid unnecessary costs caused by constantly rescheduling the production agenda, to set aside enough stocks to avoid potential supply shocks caused by the political crisis in a major raw material country…However, the classic definition of the convenience yield is indirect and obscure. For example, the original and most referred definition comes from Kaldor (1939), who defines the convenience yield as benefits that "accrue to the owner of the physical commodity but not to the holder of a forward contact." This definition only describes the attributes of the convenience yield without giving it a definite and concise meaning. Other scholars attempts to creatively define the convenience yield as a "timing option", as it grants the commodity holders the right to put the commodity on the market when prices are high but not the obligation to sell them when prices are low (Brennan, 1958 and Telser, 1958). This farfetched analogy fails to deliver what exactly is the convenience yield. I define the commodity convenience yield as follow:

The convenience yield of a commodity at time t, CY_t, is the benefit gained from holding a marginal unit of this commodity from the beginning to the end of period t, minus the storage, insurance, maintenance and other associated costs.

It is intuitive to compare the convenience yield with the dividend (as that of a dividend-paying stock) or interest payment (as that of a coupon paying bond). This comparison is the fundamental idea of Pindyck's (1993) present value model, which is directly akin to the more famous Gordon (1958)'s *dividend discount model (DDM)*. Following this line of thinking, the second part of this thesis develops the convenience yield model (CYM) explaining the relationship between the convenience yield and the spot price:

$$S(t) = \theta \sum_{i=0}^{\infty} \theta^i E[CY(t+i)] \qquad 3\text{-}7$$

In equation 3-7:

- θ, is the discount factor, which is defined as $\theta = (1+\mu)^{-1}$, where μ is the commodity specific discount rate. μ is further defined as $\mu = r + \rho$, where r is the risk-less rate and ρ is the commodity specific risk premium. So $\theta = (1 + r + \rho)^{-1}$.
- $S(t)$ is the commodity spot price at time t.

- $CY(t+i)$ is the adjusted[2] one-period commodity convenience yield at time $t+i$, i.e. the benefit of holding one unit of commodity from the beginning to the end of the period $t+i$.

After defining the commodity convenience yield and revealing its relationship with the commodity spot price, the next step is to calculate the convenience yield, as it is not readily observable. Even so, given an efficient futures market, i.e. there is no arbitrage opportunity; we can conduct the calculation of the convenience yield by using spot and futures prices only. Fortunately, the steel billet futures traded on London Metal Exchange is a relatively complete and efficient market. As a result, the LME steel billet convenience yield can be measured from studying the spot and futures prices only, in other words, this thesis does not need any additional direct information about inventories level, production and demand level etc, they are already reflected in the futures and spot prices.

Equation 3-7 is actually the solution to recursively solve the following difference equation:

$$E[S(t+1)] = (1+\mu)S(t) - CY(t+1) \qquad 3\text{-}8$$

Because the futures market is efficient, there is no arbitrage opportunity, in order to meet the non-arbitrage requirement, the following relationship must hold:

$$CY(t,T) = (1+r_T)S(t) - F(t,T) \qquad 3\text{-}9$$

In equation 3-9:

- $CY(t,T)$, is the convenience yield time t to time t + T for T periods.
- r_T, is the riskless interest rate at time T.
- $S(t)$, is the commodity spot price at time t.
- $F(t,T)$, is the commodity futures price at time t for the delivery in T periods, at time t + T.

To clarify why equation 3-9 holds, think about the following situation:

At time t, a steel market participant purchases and holds one ton of steel product which is currently priced at $S(t)$. Meanwhile, he can also short a futures contract with a maturity of t+T. Assume there are no other cash flows during the time from t to t+T and a constant interest rate r_T, then at time t+T, this process is summarized in table 3-1:

Table 3-1 Cash Flow Explanation

Time	Cash Flow	Reason
t	$-S(t)$	Price paid out to purchase the steel
t+T	$CY(t,T)$	Convenience yield benefits to hold the inventory
t+T, Delivery	$F(t,T)$	Cash received by shorting the futures contract
From t to t+T	$CY(t,T)+F(t,T)-S(t)$	Total cash gain/loss
From t to t+T	$r_T \times S(t)$	Foregone interest rate, i.e. opportunity cost, if the participants did not buy the steel at t

This table gives the explanation to why equation 3-9 holds by investigating the cash flow situations from time t to time t+T.

At time t+T, the participant's gain from holding the steel should equal the opportunity cost from the foregone interest if he otherwise decided not to purchase or hold the inventory at time t. So:

$$CY(t,T) + F(t,T) - S(t) = r_T \times S(t) \qquad 3\text{-}10$$

[2] Here, "adjusted" means the additional costs of storage, transportation, maintenance and insurance etc. are deducted from the original convenience yield already.

In 3-10, by moving the futures and spot prices to the right side of the equation, we will get equation 3-9, describing the relationship among the convenience yield, futures price and spot prices. In 3-9 and 3-10, since the convenience yield has been adjusted for the storage, maintenance and other costs, there is no need to consider these costs additionally.

Alternatively, the non-arbitrage opportunity relationship can also help explain the why equation 3-9 holds, think about the following relationship after some mathematical reorganization from 3-9:

$$F(t,T) = (1 + r_T)S(t) - CY(t,T) \qquad 3\text{-}11$$

Provided that the relationship in equation 3-11 is violated, the market participants would detect clear arbitrage opportunities:

- If the current futures price is greater than the right side of the equation 3-11, at time t, commodity traders would sell the futures contract and borrow the money to buy the commodity. After adjusting the financing and other costs, the trader will still be able to make a positive benefit of $F(t,T) - [(1 + r_T)S(t) - CY(t,T)]$ in T periods at time t + T.
- If the current futures price is less than the right side of the equation 3-11, a reverse arbitrage opportunity would occur. The commodity traders should enter long position of the futures contact and wait for the delivery, because the futures are relatively cheaper.

With more and more traders engage in exploiting such opportunities, these two parts will converge, resulting the equilibrium relationship (Hull, 2012), as described in 3-11 and indicated in equation 3-9.

Up till now, I have established the relationship between the commodity spot price and convenience yield in equation 3-7; the commodity convenience yield can be calculate from spot and futures prices as illustrated in equation 3-9. The only missing piece is the commodity specific, i.e. steel specific, risk premium ρ. The estimation of this risk premium can be attained by running a cointegrating regression by following Campell & Shiller (1987) path of thinking.

First, if both the commodity spot price $S(t)$ and convenience yield $CY(t)$ are integrated of order 1, i.e. I(1), then the equation 3-8 suggests that they are cointegrated and the cointegrating vector is:

$$\begin{pmatrix} 1 \\ -\dfrac{1}{r+\rho} \end{pmatrix} \qquad 3\text{-}12^{3}$$

From the cointegrating vector 3-12, I will define a convenience yield spread derived from the commodity spot price and convenience yield:

$$Sp(t) = S(t) - (\frac{1}{r+\rho})CY(t) \qquad 3\text{-}13$$

In equation 3-13,

- $Sp(t)$, is the convenience yield spread at time t
- $S(t)$, is the observed steel spot price at time t
- $CY(t)$, is the convenience yield calculated from equation 3-9

[3]Because $S(t+1) - S(t) \sim I(0), so \begin{pmatrix} 1 \\ -\frac{1}{r+\rho} \end{pmatrix} \cdot (S(t), CY(t)) \sim I(0)$

- r, is the risk free rate
- ρ, is the steel specific risk premium

Given that $S(t)$ and $CY(t)$ are both integrated of order 1, i.e. I(1), if they are cointegrated, equation 3-13 means the convenience yield spread $Sp(t)$ is I(0), i.e. stationary. As a result, I can theoretically estimate the commodity specific return $r + \rho$ through running the cointegrating regression between spot price and the convenience yield. Plug equation 3-13 into equation 3-7, the relationship between the convenience yield spread and the commodity spot price variation can be generated:

$$E[\Delta S(t+1)] = (r+\rho)Sp(t) \qquad 3\text{-}14$$

In equation 3-14:

- $\Delta S(t+1)$, is the one-period steel spot price change from t to t+1

Equation 3-14 also implies a regression to indirectly[4] test the interaction between the convenience yield and spot price:

$$\Delta S(t) = \alpha + \beta Sp(t-1) + \varepsilon_t \qquad 3\text{-}15$$

In equation 3-15:

- α , is the convenience yield spread premium. It is expected to be zero if equation 3-14 holds
- β , is the convenience yield spread coefficient. It is expected to be $r + \rho$ if equation 3-14 holds
- ε_t , is the white noise

The regression 3-15 can be run only under one condition that both steel spot price and convenience yield are integrated of order 1, i.e. I(1). However, results from many researches on other commodities suggest that the commodity spot price is integrated of order 1, whereas sometimes it is not the case for the convenience yield. For example, Pindyck (1993) discovers that the convenience yields of copper, heating oil, lumber and gold are not integrated of order 1. Similarly, if the steel convenience yield fails to be integrated of order 1, I cannot run either the cointegration regression or the regression 3-15 on the steel price series. Alternatively, after some maths manipulation, I am able to establish another regression.

First of all, let me define a new modified convenience yield spread $Sp(t)^*$:

$$Sp(t)^* = (r+\rho)Sp(t) \qquad 3\text{-}16$$

Plug equation 3-13 into equation 3-16:

$$Sp(t)^* = (r+\rho)S(t) - CY(t) \qquad 3\text{-}17$$

Plug equation 3-9, where I define convenience yield, into equation 3-17:

$$Sp(t)^* = (r+\rho)S(t) - (1+r)S(t) + F(t,T) \qquad 3\text{-}18$$

After some reorganization, I will get:

$$Sp(t)^* = (\rho-1)S(t) + F(t,T) \qquad 3\text{-}19$$

Combine equation 3-16 and equation 3-19:

[4] Because the independent variable of this regression is the convenience yield spread not convenience yield per se.

$$Sp(t) = \frac{(\rho - 1)}{(r + \rho)} S(t) + \frac{1}{(r + \rho)} F(t, T) \quad 3\text{-}20$$

Based on the relationship illustrated in 3-14, it is obvious that:

$$(\rho - 1)S(t) + F(t, T) = E[\Delta S(t + 1)] \quad 3\text{-}21$$

Equation 3-21 suggests the alternative regression is:

$$\Delta S(t) = \alpha^* + \beta_1 S(t - 1) + \beta_2 F(t, T) + \varepsilon_t^* \quad 3\text{-}22$$

In equation 3-22:

- α^* , is the adjusted convenience yield spread premium. It is expected to be zero.
- β_1 , is the one-period lagged spot price coefficient. It is expected to be $(\rho - 1)$.
- β_2 , is the futures price coefficient. It is expected to be 1.
- ε_t^* , is the white noise.

Equation 3-22 proves my point that the convenience yield can be proxied by one-period lagged spot price and futures price to predict the changes of the commodity price. Now I can set out to test the convenience yield model by take the following steps:

1. Calculate the convenience yield from equation 3-9;
2. If both the convenience yield and steel spot price are integrated of order 1, I run a cointegrating regression of the spot price and the convenience yield to get the estimate of the steel specific risk premium ρ;
3. Then I calculate the convenience yield spread from equation 3-13 and test regression 3-15
4. If the convenience yield is not integrated of order 1, I test regression 3-22. It's worth noting that in this case, I cannot calculate the estimation of the steel specific risk premium ρ through cointegrating regression, but I can still run the regression 3-22 and use the resulted coefficient β_1 plus 1 as an estimate of the steel specific premium ρ.

4 Hypotheses

This chapter motivates and presents the hypotheses for both UFM and CYM models. Each hypothesis is divided into two parts. The first hypothesis involves joint tests on the unbiasedness and efficiency of the futures contracts. The second hypothesis is also further divided into two parts, which are mutually exclusive. Which sub-hypothesis to test is contingent on whether the convenience yield is integrated of order 1.

4.1 Hypothesis for UFM

As discussed in section 3.1, the futures price reflects the investor's expectation of the commodity price in the future at the delivery date. The futures price contains the information how the commodity spot price would evolve as it reaches to the maturity date specified in the futures contract. As a result, in the UFM, the futures price is used as an unbiased and efficient estimator of the future spot price of the commodity. This reasoning and the theoretical discussions in section 3.1 lead to the hypothesis of the first model:

Hypothesis 1: LME steel billet futures price series are unbiased and efficient predictor of the LME steel billet spot price series.

This is the most intuitive and direct method to forecast the steel spot price in the future. As described in section 3.1, equation 3-6 presents the unbiased futures model econometrically. To test the unbiasedness, I need to compare the risk premium term with zero and compare unbiased coefficient with one. To test the efficiency, I will test whether the error terms jointly equal zero. Given this, hypothesis one can be further dissected into two sub parts:

Hypothesis 1-a: In equation 3-6, the risk premium term ρ *equals zero, at the same time the unbiasedness coefficient* β *equals one.*

Hypothesis 1-b: In equation 3-6, the error terms are jointly expected to be zero. $E(\epsilon_t \epsilon_{t+i}) = 0$ *for any i>0.*

Samuelson (1965) discovers that futures prices tend to become less volatile as the contracts approach to maturity (i.e. Samuelson Effect). Bearing this Samuelson Effect in mind, I extend the analysis and look at futures price series with different terms. Futures price with longer maturity term tends to be less volatile. Accordingly, I would expect long term futures will have less impact on expected changes of the spot price; whereas futures with shorter term should contribute more to the variations of the steel spot price series.

4.2 Hypothesis for CYM

In the CYM model, the steel spot price change in the future can be forecasted by two types of methods, and both are associated with the convenience yield. First, if both the convenience yield and spot price are integrated of order 1, I can run a cointegrating regression of steel spot price and steel convenience yield to get the steel specific risk premium. Then I can calculate the steel convenience yield spread, which is used to forecast the steel spot price change of the next period. This is the case of equation 3-15. If the convenience yield is not integrated of order 1, then I can use the convenience yield, proxied by only steel spot and futures prices, to forecast the steel spot price change of the next period. This is the case of equation 3-22. These two situations are mutually exclusive. They are given as:

Hypothesis 2-a: If the steel convenience yield is integrated of order 1, the steel convenience yield spread can predict the future steel spot price change as indicated in equation 3-15, where $\alpha = 0$ *and* $\beta = r + \rho$. *r is the risk free rate and* ρ *is the steel specific risk premium.*

Hypothesis 2-b: If the steel convenience yield is not integrated of order 1, the steel convenience yield, proxied by steel spot and futures prices, can predict the future steel spot price change as indicated in equation 3-22.

In hypothesis 2-b, since the steel specific risk premium cannot be attained by running the cointegrating regression. I just run the regression and analyze the empirical results. The coefficients from the regression can be used to estimate the steel specific risk premium.

5 Methodology

This chapter explains how to implement both models and articulates the exact regressions to test. Econometric concerns such as unit root, serial correlation and heteroskedasticity are involved in each model discussion as well. In addition, their solutions, for instances first order difference and GARCH, are offered as well. Statistics tests used to verify these models are established; meanwhile the chapter also includes corresponding commands in Matlab.

5.1 Methodological Approach for UFM

In the unbiased futures model (UFM), the hypotheses aim to establish an unbiased and efficient relationship between LME steel billet spot price and futures prices of two different terms: three months and fifteen months. Since commodity prices can never be negative, natural logarithmic scale was chosen and will be used throughout the majority of this thesis. It implies that steel spot prices are lognormally distributed in dollar terms. However this conversion has some consequences on the risk premium term of the model.

As discussed in section 3.1, the risk premium is the price difference between futures and future spot price in dollar terms. Once the log price is applied, the risk premium becomes the log of the price difference rather than the difference of the log prices. Based on equation 3-4 and 3-6, in the log terms, I need to compare the risk premium with $\log(S-F)$ rather than 0. This imposes a mathematic challenge to the regression 3-6: The number in log operator can never be zero, i.e. $\log(0)$ is not defined in mathematics. Even so, I will still test the risk premium with zero. The reason lies in finance. If in the dollar terms, the steel future price is expected to equal the future steel spot price, then in log terms, this relation still holds. In other words, there should be no price difference in either case. So it is legit to compare the risk premium with zero in log terms.

There are two regressions to evaluate:

- 3-Month futures price

$$s_t = \rho_{3M} + \beta_{3M} f(t-3M, 3) + \epsilon_{t,3M} \quad 5\text{-}1$$

- 15-Month futures price

$$s_t = \rho_{15M} + \beta_{15M} f(t-15M, 15) + \epsilon_{t,15M} \quad 5\text{-}2$$

In equation 5-1 and 5-2:

- s_t is the log return of the LME steel spot price
- ρ_{3M} is the risk premium for the 3 months futures price series
- β_{3M} is the unbiasedness coefficient for the 3 months futures price series
- $f(t-3M, 3)$ is the log return of the LME steel 3 months futures price three months ago
- ρ_{15M} is the risk premium for the 15 months futures price series
- β_{15M} is the unbiasedness coefficient for the 15 months futures price series
- $f(t-15M, 15)$ is the log return of the LME steel 15 months futures price fifteen months ago
- $\epsilon_{t,3M}$ and $\epsilon_{t,15M}$ are error terms, they are white noises.

The log return was chosen because all the level values of price series are non-stationary, exhibiting a unit root in their time series. If non-stationary variables are included, it can lead to spurious regression with no economic meaning; as a result the inferences drawn from such analysis are invalid (Granger & Newbold, 1974). The first difference usually removes the unit root non-stationarity (Hamilton, 1994).

After these initial data manipulation, I plot the return of these price series to graphically present that they are stationary. In addition to the visual presentation, I further test the unit root stationarity by using the

Augmented Dickey-Fuller test (ADF) (Dickey & Fuller, 1979) (Dickey & Fuller, 1981). The unit root stationarity is of great importance to the analysis of the unbiased futures model. If a unit root is detected in the data series, shocks will have persistent effects on the process and will not dissolve away. Furthermore, I create Autocorrelation Function (ACF) plots to check that no serial correlation is detected. This is particularly important because the regression coefficients otherwise can be inefficient, forecasts may be suboptimal, hence the tests will be invalid (Granger & Newbold, 1974).

I expect to see volatility clustering in the plots of the return series, i.e. the volatility of these series should vary over time. As suggested by Campbell, et al. (1997), it is neither consistent nor statistically efficient to use volatility measures for the homoskedastic data on the price series showing certain level of heteroskedasticity. Hence, a simple regression model that fails to encompasse the heteroskedasticty will not convincingly capture the real attributes of the return series. Accordingly, I aim at fitting a Generalized Autoregressive Conditional Heteroskedasticity (GARCH) model to the regression in order to capture the persistence of the conditional volatility. Developped by Bollerslev (1986), GARCH is an extension of the ARCH model and captures the volatility clustering that is expected to present in the LME steel spot and futures return series (Hamilton, 1994). In order to estimate the parameters p and q in the GARCH (p, q) model, I use the Beyesian Information Criteria (BIC) to determine the value of p and q parameters. This will enable to select the appropriate model for the thesis considering the behavior of the observed steel return series. I will use the resulting GARCH model to generate the coefficients and statistics that are needed for the analysis. The abovementioned process can be implemented by using *garchset* and *garchfit* commands in Matlab.

Both hypothesis 1-a and 1-b are joint tests, i.e. both test more than one conditions simultaneously. Hypothesis 1-a can be done by testing restrictions on the coefficients of equation 6-1: $\rho_{3M} = 0$ and $\beta_{3M} = 1$; similarly for equation 6-2: $\rho_{15M} = 0$ and $\beta_{15M} = 1$. The Wald test (*waldtest* in Matlab) can be used for this purpose. Hypothesis 1-b can be tested by conducting an *F-test*.

In UFM model, I also assume that the error-terms are mere white noises, i.e. i.i.d following Gaussian with the distribution of N(0, 1). The Ljung-Box test (*lbqtest* in Matlab) is included to check how close the error terms to a white noise. This is important as residuals are not suppose to show any form of correlation or pattern. I conclude the unbiased futures model by conducting a 3 periods (days) out-of-sample forecast for the log return of the LME steel billet spot price series.

5.2 Methodological Approach for CYM

The first job for the convenience yield model is to calculate convenience yields for both futures price series. Based on the relationship described in equation 3-9, the convenience yield can be calculated from the LME steel billet spot and futures price series, together with the risk-free interest rate. I will use the short term (4-weeks) US Treasury bill rates (US Treasury, 2012) as risk-free interest rate. I can get two different convenience yields:

- 3-Month futures price:

$$CY(t, 3M) = (1 + r_{3M})s_t - f(t, 3M) \quad 5\text{-}3$$

- 15-Month futures price:

$$CY(t, 15M) = (1 + r_{15M})s_t - f(t, 15M) \quad 5\text{-}4$$

In equation 5-3, 5-4:

- $CY(t, 3M)$ is the 3-Month convenience yield and $CY(t, 15M)$ is the 15-Months convenience yield.

- r_{3M} and r_{15M} are the risk-free rates reflecting the financing costs.
- s_t is the log steel spot price.
- $f(t,3M)$ is the log 3-month futures price and $f(t,15M)$ is the log 15-months futures price.

The next step is to check whether these convenience yields are integrated of order 1. This can be alternatively done by checking whether the first difference series of these convenience yields are integrated of order 0. To implement this integration test, use *i10test* command in Matlab. Once the integration test results are attained, I can decide which hypothesis to test: hypothesis 2-a or 2-b.

If the convenience yield is integrated of order 1, I can run a cointegrating regression between the steel spot price and the convenience yield to calculate the steel specific premium. This can be done by conducting an Engle-Granger cointegration test (Engle & Granger, 1987). In Matlab, the command is *egcitest*. I use the cointegrating vector from cointegration to calculate the estimate of the steel specific risk premium ρ_{3M} and ρ_{15M}. Then I can define the steel convenience yield spread based on equation 3-13:

- 3-Month futures price:

$$Sp(t,3M) = s(t) - (\frac{1}{r_{3M} + \rho_{3M}})CY(t,3M) \quad 5\text{-}3$$

- 15-Month futures price:

$$Sp(t,15M) = s(t) - (\frac{1}{r_{15M} + \rho_{15M}})CY(t,15M) \quad 5\text{-}4$$

After calculating the convenience yield spread, I will conduct the unit root stationariy test on this spread series for the same econometric reasons explained in section 6.1. I will also fit a GARCH model to account for the heteroskedasticity. Afterwards, I will estimate the following regression models explained in equation 3-15:

- 3-Month futures price:

$$\Delta s(t)_{3M} = \alpha_{3M} + \beta_{3M} Sp(t-1,3M) + \varepsilon_{t,3M} \quad 5\text{-}7$$

- 15-Month futures price:

$$\Delta s(t)_{15M} = \alpha_{15M} + \beta_{15M} Sp(t-1,15M) + \varepsilon_{t,15M} \quad 5\text{-}8$$

In equation 5-7 and 5-8:

- $\Delta s(t)_{3M}$ and $\Delta s(t)_{15M}$ are steel spot log returns.
- α_{3M} and α_{15M}, are convenience yield spread premiums, which are expected to be zero.
- β_{3M} and β_{15M} are the convenience yield spread coefficients, $\beta_{3M} = r_{3M} + \rho_{3M}$ and $\beta_{15M} = r_{15M} + \rho_{15M}$.
- $\varepsilon_{t,3M}$ and $\varepsilon_{t,15M}$ are white noises.

However, if the steel convenience yields are not integrated of order 1, then cointegration cannot be run. As a result, I need to run another regression as suggested in equation 3-22:

- 3-Month futures price:

$$\Delta s(t)_{3M} = \alpha_{3M}{}^{*} + \beta_{3M,1} s(t-1) + \beta_{3M,2} f(t,3M) + \varepsilon_{t,3M}{}^{*} \quad 5\text{-}9$$

- 15-Month futures price:

$$\Delta s(t)_{15M} = \alpha_{15M}{}^{*} + \beta_{15M,1} s(t-1) + \beta_{15M,2} f(t, 15M) + \varepsilon_{t,15M}{}^{*} \quad 5\text{-}10$$

In equation 5-9 and 5-10:

- $\alpha_{3M}{}^{*}$ and $\alpha_{15M}{}^{*}$, are convenience yield spread premiums, which are expected to be zero.
- $\beta_{3M,1}$ and $\beta_{15M,1}$, are one-period lagged spot price coefficients. These coefficients are used to estimate the steel specific premiums.
- $\beta_{3M,2}$ and $\beta_{15M,2}$, are future price coefficients. $\beta_{3M,2} = 1$ and $\beta_{3M,2} = 1$.
- $\varepsilon_{t,3M}{}^{*}$ and $\varepsilon_{t,15M}{}^{*}$ are white noises.

Finally, I conclude the convenience yield model by conducting a 3 periods (days) out-of-sample forecast for the log return (Δs) of the LME steel billet spot price series.

6 Data

This chapter motivates the choice of data. It also discusses the source and reliability of the data. At the end of this chapter, it plots the LME steel billet data visually and discovers that the LME steel billet price performs contango rather than backwardation.

The analysis of this thesis use these data series: daily LME steel billet cash settlement (as steel billet spot price), daily LME steel billet 3 Months settlement, daily LME steel billet 15 Months settlement and Treasury bill rates (as the riskless interest rates).

6.1 Data Choice and Motivation

Starting in 2008, LME initiated first ever steel futures contracts that are globally traded on an organized clearing house. Introduced during the most volatile time of the steel marketplace, the LME steel billet contracts were designed and created to help steel market participant hedge their risk exposures more effectively (Steel Times International, 2008). The idea to bring together steel product contracts on an exchange can be traced back to 2005, when LME announces its decision on steel futures creation (Steel Times International, 2005). After LME steel billet contracts, NYMEX constructed the futures contract for hot rolled steel coil, another important type of steel product. These are two major exchanges actively engaged in steel product transaction on a daily basis. In addition, there are a number of contracts for other types of steel products, mainly forwards and swaps, traded either over-the-counter or on much smaller exchanges focused on regional transactions. This thesis chooses LME steel billet products for several reasons:

1. LME steel billet contracts are the first ever created and publicly traded steel futures in a well-recognized exchange with global impact. Compared to other steel product contracts, the LME steel billet price series cover the longest trading period and contain the most complete observations. Even though some experienced steel participants usually argue that LME is not the first exchange engaged in trading steel contracts. For example, back in 1990s, a small number of steel futures were created in China and three small exchanges were created to trade those products, however, they were all shut down in 1994 due to poor regulation and lack of liquidity. (Steel Business Briefing, 2012)
2. LME steel price series are published daily, meeting the need of the high-frequency data as requested in this thesis.
3. LME steel data is one of the most influential and standardized international steel benchmark to which all the steel market participants pay attention.
4. The quality of the LME steel billet is clearly defined and strictly regulated by its trading specifications. For example, the lot size of the steel billet is 65 metric tons, with a tolerance of +/-3.5% of the size. The delivery has to take place in one of the appointed facility locations that is approved for the LME steel billet storage. (London Metal Exchange, 2008) The complete rules and guidance for LME steel billet can be found on the steel section of the LME's webpage. Regulation concern is another reason for this thesis to shy away from poorly supervised regional exchanges in China.

6.2 Source and Reliability of the Data

All the three price series and US Treasury bill rates are traced and collected from Datastream, a holistic finance and economics database provided by Thomson Reuters. Even though Datastream is not the direct source of the data (LME and US Department of Treasury are responsible for the publications for steel price series and Treasury bill rate series respectively), there are several benefits for using Datastream:

1. Datastream is a frequently used financial database covering a wide range of price series with all levels of frequencies;

2. Datastream is synchronized with real time data and its sources are well cited, enabling users to trace back to the very original publication of the data if it is still available and accessible;
3. Datastream is a highly regarded and dependable database for both industry and academic works. The reliability is guaranteed;
4. The Datastream can be built in spreadsheet packages such as Excel for easy data downloads and process.

For above reasons, this thesis adopts Datastream as the primary and main source of the data. The trustworthy of the database also minimize the observation errors of this paper.

6.3 Data Description

First of all, I plot the price trajectories of the LME steel billet spot price series. The LME steel billet cash settlement is used as the spot price series. The other two LME steel price series are used as 3-Month futures and 15-Month futures price series. All the price series in this thesis are of daily frequency.

6.3.1 LME Steel Billet Spot Price

In April 2008, when LME first launched the steel billet futures, there were two region-specific contracts created, Mediterranean and Far East steel billet. On July 27th 2010, LME merged the two contracts into a single, global contract. From historical trading volume perspective, Mediterranean contracts are considerably larger than the Far East ones. So the "merger" of the two contracts is actually the "acquisition" of Far East by Mediterranean. Thus, this thesis will use the Mediterranean contract price before July 27th, 2010 and adopt single global contract price after this date. Plus, the price difference between these two contracts is negligible. For the steel spot price, the active trading time started from July 24th 2008, depicted in

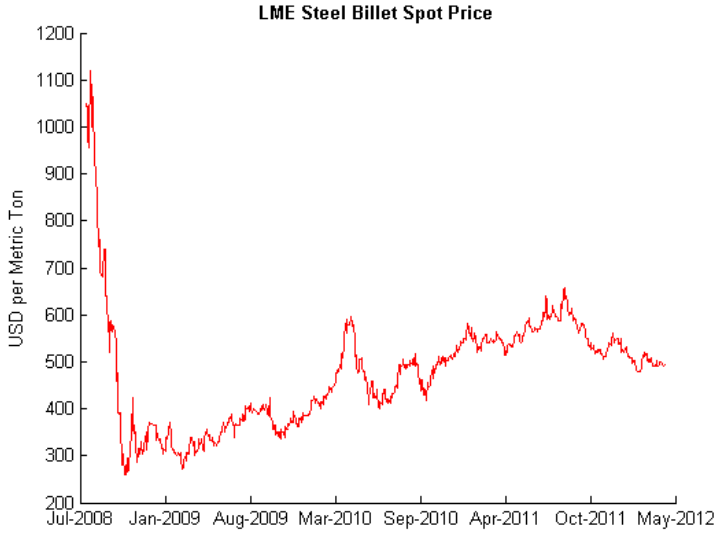


Figure 6-1 LME Steel Billet Spot Price Series

This figure plots steel billet spot price from July 2008 to May 2012. It shows that the steel spot price has experienced great volatility throughout its trading history.

As shown in Figure 6-1, the steel price fluctuates over the time. The price was particularly volatile during the introduction period, which justified LME's initial idea of creating steel futures contracts as a tool for market participants to hedge their risk exposures. From the mid of 2008 to the beginning of the 2009, the

steel spot price plunged sharply from all-time high, over 1100 USD/MT, to all-time low, around 250 USD/MT. Just as the case of other commodities during this time, this dip was mainly caused by external factors such as global financial crisis. The price picked up again in March 2010, then dipped again and presented heavy volatility ever since.

6.3.2 Contango vs. Backwardation

Commodities usually perform different patterns in their spot and futures prices. Traders typically use the terms *contango* and *backwardation* to describe the situation of commodity spot price relative to its futures price. One way to define them is based on the relationship between futures price and the current spot price. Contango is the case that the futures price is above the spot price, whereas backwardation occurs if the futures price is below the spot price (Hull, 2012). As depicted in figure 6-2[5], for the most of time, the LME steel billet performs contango rather than backwardation.

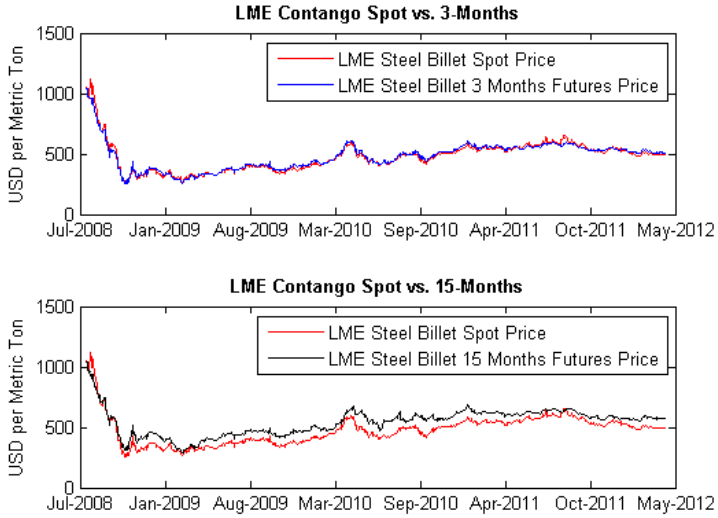


Figure 6-2 LME Steel Contango

This figure plots and compares the level value of three price series from July 2008 to May 2012. In the upper plot, the red line is the spot price and the blue line is the 3 months futures price. The price difference is not noticeable. In the lower half, the red line is the spot price and the black line is the 15 months futures price. For the most of the time, contango is detected.

In Figure 6-2, except for some periods when backwardation was detected, most of the time the steel spot price is higher than the steel futures prices. The red line is the steel spot price, the blue line is the 3-Months Futures price and the black line is the 15-Months Futures price. For most of the period, at any given time, the futures prices are higher than the spot price. Contango is especially obvious when I compare the spot price with the 15 months futures price. Based on these observations, the steel price behavior is quite different from other commodities which show backwardations, such as in crude oil (Litzenberger & Rabinowitz, 1995) (Gibson & Schwartz, 1990) (Gabillon, 1991); wheat, cotton and corn (Telser, 1958); even within the metal family, gold and silver are backwarded (Gemen, 2005). This discovery is very interesting for commodity research. Contrary to Keynes' (1930) argument that "the spot price must exceed the forward price…to avoid

[5] Due to the fact that the LME steel billet spot price was not created until July 2008 and futures contracts were traded in April, I align 3-Months and 15-Months futures with spot price series and get rid of the excess data.

the risk of price fluctuations… in normal conditionals the spot price exceeds the forward price, i.e. there is a backwardation," the steel market performs a contango instead. It means that the traditional assumption of normal backwardation for commodities in general is not appropriate in this thesis. Nor is its implication such as futures prices are downward-biased estimation of the future spot prices. Hence, it is necessary to conduct an empirical study in the steel market and test the unbiasedness of the steel futures price to predict the future steel spot price as described in the UFM model.

In this section, in order to clearly describe the steel price level, I use the USD per Metric Ton unit in section 6.3. However, as indicated in the method part, for the rest of UFM model, I will adopt the log price instead.

7 Empirical Results and Analysis

This chapter presents the empirical findings of both models. It is separated in two sections: results for UFM and results for CYM. For UFM, both the unbiasedness hypothesis and efficiency hypothesis are rejected. For CYM, the hypothesis is partially rejected. The out of sample forecast suggest that neither models is effective or accurate at predicting the steel prices in the future.

7.1 Results for UFM

For the UFM model, the first task is to check the unit root stationarity of the steel price series by plotting the log price for all the three series:

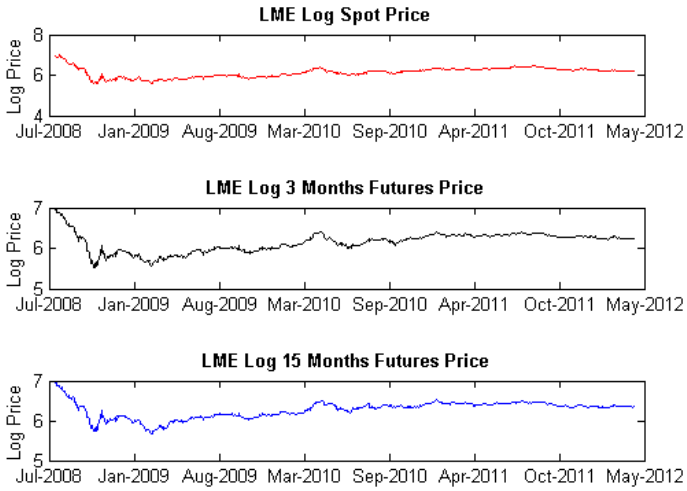


Figure 7-1 LME Steel Billet Log Price

This figure plots the log value of three price series. The red line is the log spot price, the black line is the log 3 months futures price and blue line is the 15 months futures price. The time is from July 2008 to May 2012.

As depicted in figure 7-1, the log price series do not provide too much information. All three series seem to follow random paths. To solve this, Hamilton (1994) suggests taking the first difference. It will both eliminate the unit root and give finance meaning to the generated series: the log return. The log return plots:

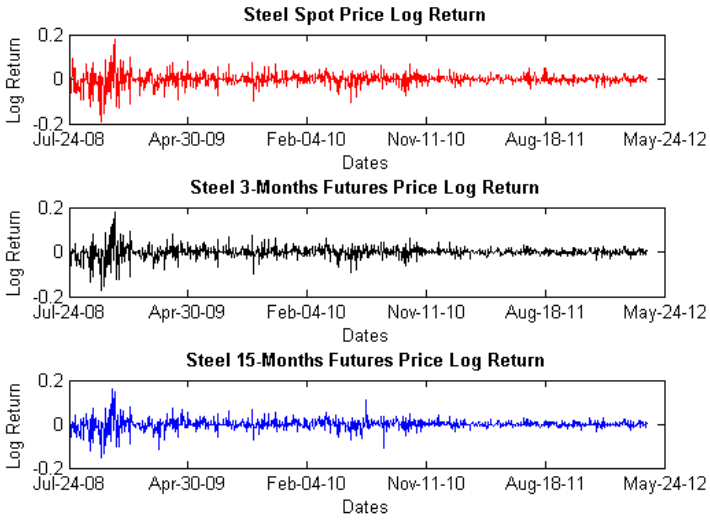


Figure 7-2 LME Steel Billet Log Returns

This figure plots the log returns of three series. The time is aligned from July 2008 to May 2012. It shows the tendency of unit root stationarity and volatility of clustering.

As shown in Figure 7-2, these steel return series are not stationary. It is consistent with the results of many other commodities. However, it displays the tendency of unit root stationarity through time. Figure 7-2 is no more than a visual presentation; a formal test is needed to determine this tendency. As discussed in section 5.1, I turn to run an Augmented Dickey-Fuller (ADF) test on the log return series to test the unit root stationarity. This can be done via the *adftest* command in Matlab. The results are illustrated in table 7-1:

Table 7-1 Augmented Dickey Fuller Test

Log Return Series	Rejection	Test Statistic	p-Value
Spot Return	1	-32.7267	0.001
3 Months Futures Return	1	-34.3298	0.001
15 Months Futures Return	1	-29.0471	0.001

0 = Do not reject the null hypothesis of unit root;
1 = reject the null hypothesis of unit root
This table presents the results of ADF test to see if price series contain unit root.

Based on table 7-1, on the 99.9% confidence level, all the unit root hypotheses are rejected. This is consistent with the results shown from the visual presentation of the time series plotted in figure 7-2, all three steel return series are indeed unit root stationary.

In figure 7-2, I also sensed volatility clustering for all steel returns, indicating that they might be serially correlated (i.e. autocorrelated). Hence, I plot autocorrelation functions (ACF) to check serial correlations.

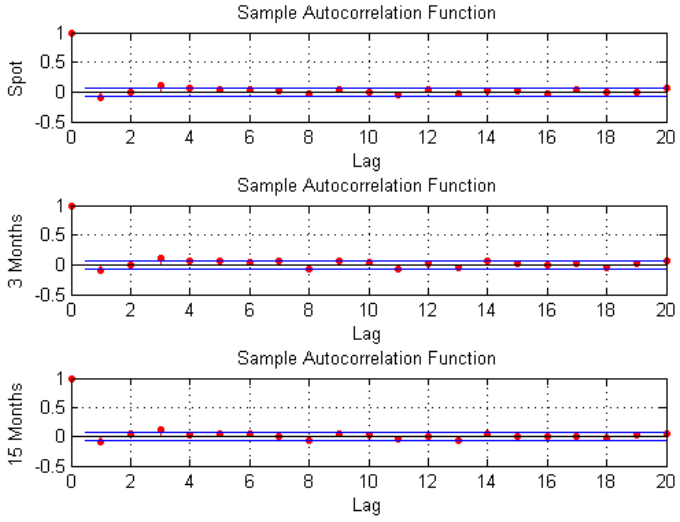


Figure 7-3 Plot of ACF to check serial correlation

This figure presents the results of sample autoregressive function of three price series to detect autocorrelation.

Based on figure 7-3, the results are not absolutely obvious. Most of the lags are either on or close to the significance level boundaries. The only lag slightly out of the significance boundaries is lag 3.

The above results suggest the variances are changing with time. Accordingly, I will fit a GARCH model to account for the heteroskedasticity in the return series. First I need to estimate the p and q parameters to be used in the GARCH (p, q) model. Based on the Bayesian Information Criteria (BIC), the GARCH (1, 1) model is identified. In table 7-2, I summarize the results of the GARCH (1, 1) model with the 3 months and 15 months lagged steel futures log returns as independent variables and the spot steel log return as independent variables, as described in equation 5-1 and 5-2 in section 5.1.

Table 7-2 UFM Model Regression 3-6

	Spot and 3-Months (Equation 5-1)		Spot and 15-Months (Equation 5-2)	
Coefficient	$\rho_{3M} = 0.00022661$	$\beta_{3M} = -0.0662$	$\rho_{15M} = 0.00040361$	$\beta_{15M} = -0.0420$
Standard Error	0.00056334	0.0296	0.00063471	0.0253

This table presents the coefficients and standard error of the UFM models after fitting them with GARCH (1, 1)

Table 7-2 shows that the steel risk premium terms from both regressions are very small and both unbiasedness coefficients are negative. However to test the unbiasedness hypothesis, a joint test is needed. As described in section 5.1, Wald test (*waldtest* in Matlab) is needed for hypothesis 1-a:

Table 7-3 Wald Test: Unbiasedness Test Hypothesis 1-a

	Spot and 3-Months	Spot and 15-Months
Rejection	1	1
Wald Statistic	5.9915	3.4917

0 = Do not reject the hypothesis 1-a that $\rho = 0$ and $\beta = 1$;
1 = reject the null hypothesis 1-a

This table provides the result of the Wald test for the joint hypothesis about unbiasedness

In table 7-3, both Wald tests strongly suggest that the unbiasedness hypothesis, i.e. hypothesis 1-a, should be rejected. This result indicates that UFM model is not unbiased at predicting future spot prices for steel billet; therefore in the steel billet context, the futures price cannot be used as an unbiased predictor of future spot price.

After testing the unbiasedness of the UFM model, there is an additional aspect of UFM model to test: efficiency, i.e. hypothesis 1-b. Described in section 5.1, I test whether the error terms are jointly zero, meaning whether new coming information would have an impact on the steel price changes in the future. If error terms are jointly zero, then new information will not affect steel price, hence this model is efficient. Explained in 5.1, this can be tested via an F-test.

Table 7-4 Testing the efficiency of the UFM model:
$s(t) - f(t,T) = \alpha + \beta[s(t-1) - f(t-1,T)]$
$T = 3M, 15M$

Coefficient	Spot & 3-Months	Spot & 15-Months
Rejection	1	1
α	-0.0010919	0.0015159
β	-0.0464929	-0.01341987
R^2	0.0022	0.00166
F-statistic	2.09	2.95

0 = Do not reject the efficiency hypothesis 1-b;
1 = reject the hypothesis 1-b.
This table presents the results from the F-test on the efficiency of the UFM models

In table 7-4, both F-statistics are significant on a 5% level, meaning that the error terms are not jointly zero and the efficiency hypothesis of UFM model should be rejected. This result indicates UFM is not an efficient model for the steel billet. In the steel billet context, the futures prices are not efficient predictors for future spot price neither. Additional information will have impacts on the steel billet spot price.

So far, I can reject both hypothesis 1-a and 1-b, hence I can conclude that the UFM model is neither unbiased nor efficient at predicting future spot prices for the steel billet. Then, I will run the Ljung-Box tests to verify whether the error terms are white noise or not, this approach will help to prove the appropriateness of the model. The command for Ljung-Box test in Matlab is *lbqtest*.

Table 7-5 Ljung-Box Test

UFM	3 Months	15 Months
Lags	1	1
Rejection	1	1
p-Value	0.0141	0.0139

0 = Do not reject the null hypothesis that error terms are white noises;
1 = reject the null hypothesis
This table gives the results whether the error terms from the UFM models are close to white noises.

Table 7-5 means that null hypothesis of UFM's regression error terms are white noises should be rejected. The results indicate that for the steel billet price, the UFM model suggested by equation 3-6 is not the most suitable model. One of the possible explanations is that some steel related independent factors are neglected.

To recap, the results of the UFM model are not satisfactory at explaining steel billet price behavior, but there is some level of explanatory power in the steel futures series to predict the steel billet's spot price in the future. This inference is not surprising, as the tone of this thesis is to negate the unbiasedness of the steel futures price rather than its explanatory power. Steel futures price should always be included in models to predict spot price; furthermore, other steel relevant factors should be explored and incorporated into the predicting model.

7.2 Results for CYM

Now, I will present results from CYM model. First, according to equation 5-3 and 5-4, I calculate steel billet convenience yields from both the 3-Month and 15-Month futures price series. Plot convenience yield:

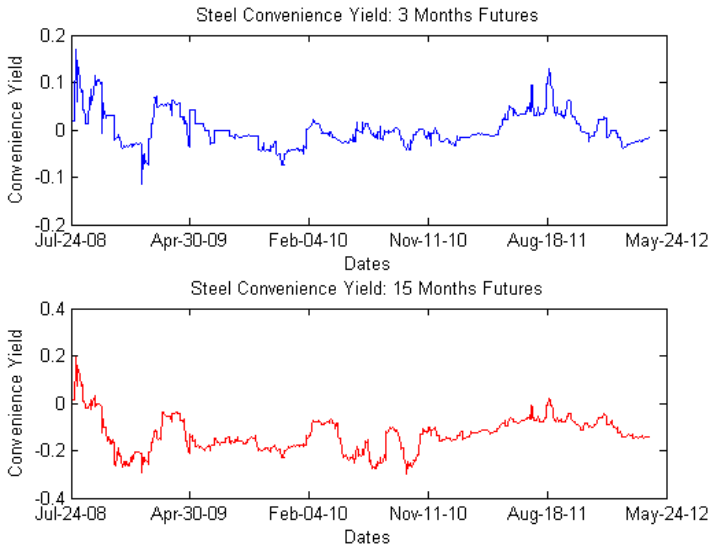


Figure 7-4 Plot of the Convenience Yields

This figure presents the convenience yields calculated from the 3 months futures and 15 months futures respectively. The time spans from July 2008 to May 2012.

Figure 7-4 depicts that the steel convenience yields seem to follow random paths. As suggested in section 5.2, the next step is to check whether the convenience yields are integrated of order 1. This step is crucial because it determines which regression to use: equation 5-7, 5-8 or equation 5-9, 5-10 and which hypothesis to test: hypothesis 2-a or hypothesis 2-b.

Table 7-6 Integration of order 1 Test

	Steel Billet Convenience yield 3-Months	Steel Billet Convenience yield 15-Months	Steel Spot Price
Rejection	1	0	0
p-Value	0.0010	0.1472	0.2921

0 = Do not reject the integration hypothesis;
1 = reject the null hypothesis that convenience yield is integrated of order 1.
This table presents the results of integration result of convenience yields and steel spot prices.

Interestingly, table 7-6 suggests that the steel billet convenience yield calculated from 3-Months futures are not integrated of order 1, whereas the convenience yield calculated from 15-Months futures are integrated of order 1. Conducting the same integration test, steel spot price is verified to be integrated of order 1. That

means that I can use equation 5-8 to test hypothesis 2-a for 15 months convenience yield. As for 3 months convenience yield, I resort to use equation 5.9 to test hypothesis 2-b.

7.2.1 Results for CYM 3 Months

For the same econometric concerns on heteroskedasticity, I fit a GARCH model for equation 5.9 first. After running the *garchfit* in Matlab, the BIC suggests a GARCH (1, 1) model. In table 7-7, I summarize the results of fitting the GARCH (1, 1) model on equation 5-9. Unlike the previous UFM model, this regression and hypothesis 2-b do not require joint statistic test.

Table 7-7 CYM for 3 Months Equation			
Coefficient	$\alpha_{3M} = -0.0402$	$\beta_{3M,1} = 0.7257$	$\beta_{3M,2} = -0.7165$
t-Statistic	3.2845	-53.9506	53.3316
This table presents the coefficients and t-statistic of the CYM 3 after fitting the GARCH(1, 1) model			

Table 7-7 shows that both the steel convenience yield spread premium term and the 3 Months convenience yield coefficients, proxied by one day lagged steel spot price and 3 Months futures price, are significant at 5% confidence level. So these variables succeed in explaining some of the variations in steel billet price change. Now, I validate the GARCH (1, 1) by conducting a Ljung-Box test to see if the error terms in equation 5.9 are indeed white noises.

Table 7-8 Ljung-Box Test for CYM 3 Months	
Rejection	1
P-Value	0.0047
T-Static	40.2441
0 = Do not reject the null hypothesis that error terms are white noises; 1 = reject the null hypothesis This table gives the results whether the error terms from the CYM 3 model are close to white noises.	

In Table 7-8, for the CYM 3 Months model, the test shows that on a 99.5% confidence level, I reject the null hypothesis favoring that in equation 5.9 error terms are white noises.

7.2.2 Results for CYM 15 Months

For the case of using 15 month futures price series, I first conduct an Engle-Granger cointegration test between the steel spot price and the steel convenience yield calculated from the 15 months futures price.

Table 7-9 Engle-Granger cointegration test	
Rejection	0
Steel-specific risk premium	3.2447
T-Static	-3.1234
Critical Value	-3.3426
0 = Do not reject the null hypothesis that steel spot prices and convenience yield calculated from 3 months contracts are integrated; 1 = reject the null hypothesis This table gives results from the cointegration test between the steel spot prices and convenience yield to get the steel specific risk premium.	

In Table 7-9, I am confirmed that the steel spot price and the steel convenience yield are indeed cointegrated. The steel-specific risk premium is calculated from the resulting cointegrating vector. Then, as indicated in equation 5-4, I can calculate the convenience yield spread, which is used as the independent variable of the regression in equation 5-8. Similarly, I fit GARCH (1, 1) model for equation 5-8:

Table 7-10 CYM for 15 Months		
Coefficient	$\alpha_{15M} = -0.0190$	$\beta_{15M} = 0.0031$
T-Statistic	-0.9373	0.9514
This table presents the coefficients and t-statistic of the CYM 15 after fitting the GARCH(1, 1) model		

Table 7-10 shows that neither convenience yield spread premium term nor the convenience yield coefficient is statistically significant. Hence, the 15 month convenience yield spread does not successfully explain the variations in the steel billet spot price changes. I suspect this is caused by Samuelson Effect, indicating that steel futures price with longer term reflects more about the steel equilibrium price in the long run rather than its current price level.

Now, I validate the GARCH (1, 1) model by conducting a Ljung-Box test to see if error terms in equation 5-8 are indeed white noises.

Table 7-11: Ljung-Box Test for CYM 15 Months	
Rejection	1
P-Value	0.0024
T-Statistic	24.7443
0 = Do not reject the null hypothesis that error terms are white noises; 1 = reject the null hypothesis This table gives the results whether the error terms from the CYM 15 model are close to white noises.	

In Table 7-11, for the CYM 15 Months model, the test results indicate that on 99.76% confidence level, I reject the null hypothesis favoring that in equation 5-8, error terms are not white noises. That means some relevant steel factors are ignored in equation 5-8, when CYM adopts longer term futures price series. For CYM, the model derived from 3 Months futures price has more explanatory power than that derived from 15 months futures. There are two possible explanations. First, from the finance perspective, the 3 Months futures contain more relevant information to the steel spot price than 15 Months futures. 3 Months futures price is a near term estimation of the spot price, which lay more focus on near term steel price development, whereas the 15 Months lay more emphasis on the long term equilibrium price. Second, from the econometric perspective, equation 5-9 directly uses one-period lagged steel spot price and 3 months futures price to forecast steel billet spot price changes; however, in equation 5-8, the steel specific risk premium is just an estimate from the cointegrating regression. As a result, equation 5-8 would contain more estimation errors.

7.3 Out of Sample Prediction

I conclude this chapter by conducting three periods out of sample predictions for both models and both future price series. When I implemented the above models in Matlab, I intentionally left out last three days data for out-of-sample prediction purpose. The correspondent prediction command in Matlab is *garchpred*:

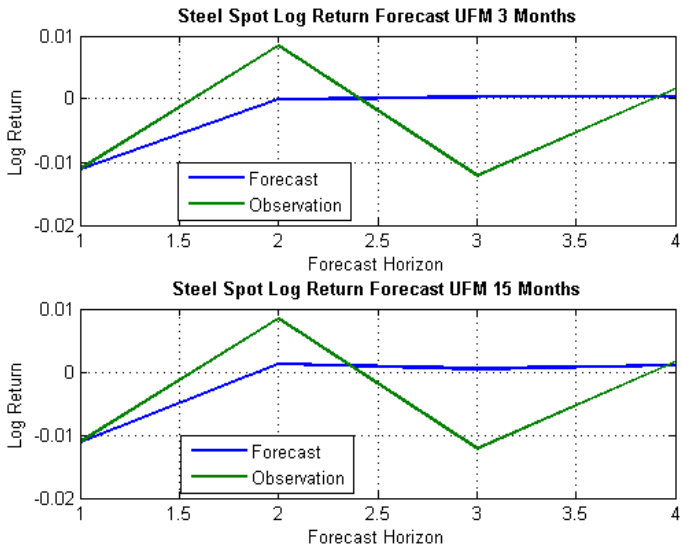


Figure 7-5 Three Periods Out of sample prediction

This figure presents the result of the 3 days out of sample prediction by using UFM model. It shows that the model derived from 15 months futures contract perform slightly better than the model derived from 3 months futures.

Figure 7-5 illustrates the out of sample steel spot price prediction by using UFM models. The green line is the observed LME steel billet spot price and blue line is the projection from the calculation. As shown in the figure, the fit is really bad for both future series, though the price predicted by the 15 month future series performs slightly better than the price calculated from the 3 month futures price. Given this 3 period forecast horizon, I calculate the Root Mean Squared Error (RMSE), which is usually used to measure the average magnitude of the error. The RMSE for the UFM-3 Months model is 0.0087 and RMSE for UFM-15 Months model is 0.0084, which proves that by using 15 months futures, the UFM model performs a slightly better fit than using 3 months futures. One possible explanation is that with the introduction of the steel futures prices, the steel billet market becomes more transparent. Compared to the situation in the past, the spot price is getting a little bit closer to the equilibrium price determined by the steel market forces. On the other hand, the futures contract with longer maturity reflects investors' expectation of the steel billet's equilibrium price in the long run. Hence, the futures contract with longer term performs relatively better at predicting spot price.

Similarly, the 3-period out of sample prediction of the spot log returns by using CYM:

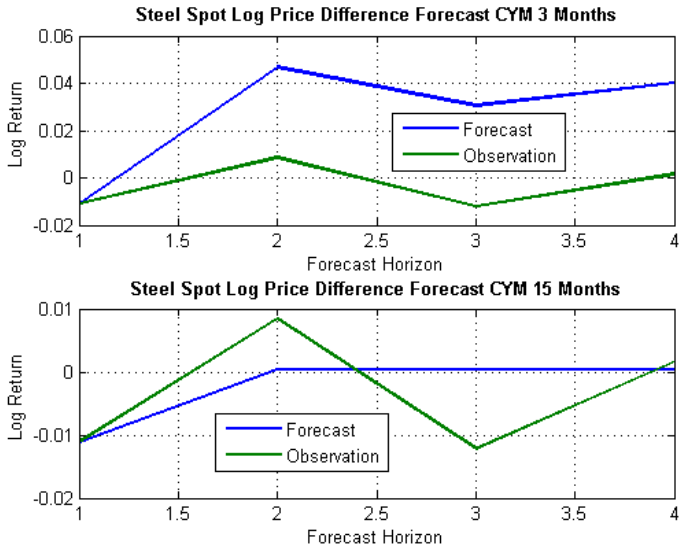


Figure 7-6 Three Periods Out of sample prediction

This figure presents the result of the 3 days out of sample prediction by using CYM model. It shows that the model derived from 3 months futures contract correctly predicts the change of direction, and the slopes of the last two periods are close to the observed value, but the difference between the level values is significant. The model derived from 15 months is slightly better than the model derived from UFM models.

In Figure 7-6, even though the CYM 3 Months model successfully captures the direction movement, its prediction on the level of future steel spot price change is completely off the chart. The RSME for CYM 3 months model is at 0.0345, the worst of all four predictions. CYM 15 Months model follows a similar path of UFM models, with only slightly better RSME value at 0.007. This section indicates that within the steel billet context, the convenience yield model is not always a better fit than using the futures price as an unbiased predictor. This is especially the case when short term futures price is used. The CYM with long term futures just performs slightly better than the UFM models.

Generally speaking, most of the forecasts have certain good estimates of the log returns towards the end of each forecasting horizon segment. But the results in between are considerably poor. UFM 3 months, UFM 15 months and CYM 15 months fail to capture the changing directions of the returns for the last two forecasting horizons. The CYM 3 months captures this feature; however the predicted return level is significantly different from the observed value. The confidence intervals of these price forecasts would be considerably wide, hence fail to provide reliably accurate results considering the amount of volatility displayed in the changes of steel billet spot price levels. Further, for the UFM model, the choice of different steel futures with different terms only slightly contribute to the quality of the forecasts. For the CYM model, steel futures price with longer term gives lower RSME value hence performs better at predicting future steel spot prices.

8 Limitation

In this chapter, the author explains the reasons that models fail in the empirical tests. There are undiscovered variables not included in the model. The futures price series are only limited to two terms.

The author is well aware of his research shortcomings and includes delimitations of the thesis in this chapter. The most interesting question is why two theoretically sound models, UFM and CYM, fail empirically in the steel billet market? I believe there are two major limitations that can provide some insights.

8.1 Model Limitations

First of all, both models have built-in disadvantages. For UFM, the goal is to explore the relationship between the steel billet futures price and spot price; CYM aims to examine the interaction between the price changes and convenience yield spread, which is calculated by futures and lagged spot price. Other factors such as the risk-free interest rate and the steel specific risk premium are involved too. However, according to the statistical results, there are neglected variables that might contribute to the price changes. For example, Hong and Yogo (2011) suggest that open interest, i.e. the amount of outstanding futures contracts (Hull, 2012), contributes to explain the spot price changes of some precious metals such as silver. But as explained before, silver is a metal mainly used for investment purposes, its price behavior will be different from steel. The impact from inclucing open interest remains to be tested. Plus, including more factors is not necessarily an effective or efficient solution. For example, Koekebakker & Ollmar (2005) include 10 factors to explain the price changes in the Nordic electricity market. According to Principal Component Analysis (PCA), it is a well known fact that three risk factors are statistically sufficient for most of the cases (Abdi & Williams, 2010). A model with too many risk factors is both cumbersome to implement and the model can be easily distracted by factors that actually matter less. In this thesis, the UFM model lost the merits by just using futures price as its only independent variable. However, as explained before, the focal point of this thesis is to test the unbiasedness hypothesis. Including more variables and more lags will divert the test focus on the explanatory power of the futures price rather than unbiasedness or efficiency.

Another drawback in the thesis lies in the CYM model, where I assume that the risk-free interest rate is not stochastic. This assumption is convenient in two ways. First, the forward and futures contracts are interchanable under the constant interest rate conditon (Hull, 2012). Second, the modeling of the convenience yield is simple. Nevertheless, to better capture the salient features of the steel price behavior, a more powerful model should have adopted the stochastic process, which is either assumed to be Geometric Brownian or Ornstein-Uhlenbeck. (see Merton (1973) and West & Harrison (1996)).This change requires solving stochastic different equations, which will considerably increase the complexity of the problem.

8.2 Data Limitations

The biggest challenge of my thesis data is the time span. LME did not launch the steel billet contracts until 2008. As described in section 6.31, there are quite recognizable abnormalities of the steel prices during the period of 2008 and 2009, which is alleged to be caused by heavy speculations in the whole commodity world. The speculation issue will be discussed further in section 9.1. Here I just stress the data limitations. For more active metal commodities with longer history, there are more than just two futures series. For example, for coppers traded on LME, there are futures with different terms from 3 months up to 27 months. However, new products like steel billet only have two futures: 3 months and 15 months. If there had been futures contracts with more terms, I could conduct more analyses and compare these models developed from futures with different maturities.

Up till now, there are futures contracts for only two steel products: billet and coil. Granted the importance of the steel billet, it is after all a semi-finished product. The steel industry calls for the creation of futures contracts for other important steel products such as HDG and EG. In addition, there is no option contract for steel products traded on a global exchange at present. The involvement of the options will further vitalize the trading activities in the steel market, as well as adding more complexities in the practice and research on steel pricing.

Finally, steel is an interim product of the value chain. In the downstream, there are terminal products such as automobiles, building skeletons, heavy mechanics etc; in the upstream, steel is directly related to a prominent yet academically neglected raw material: iron ore. One possible solution to optimize proposed theoretical models is adding iron ore price series as an independent variable for the regression. However, there is neither futures data nor reliable spot price data for iron ore yet. The genesis of an internationally organized and transparent iron ore futures market remains to be initiated.

9 Conclusion

This chapter reiterated the results from the empirical findings. Plus, the author also gives discussion why results from different futures terms will be slightly different. Moreover, in the second half of this section, the author discusses the heavy speculations in the commodity market and their impact on distorting the fair market prices.

In this chapter, I conclude the findings of my thesis. First of all, in the UFM model, I did not see general patterns and explanatory power of the LME steel billet futures price over the steel billet spot price. For the CYM model, the convenience yield for the 3 months futures is not cointegrated with the steel billet spot price. The convenience yield for the 15 months futures is cointegrated with the spot price. However, the statistical inference is more significant for the 3 months CYM than that of the 15 months CYM model. When it comes to the out of sample forecasts, the results are unsatisfactory for both models. They either fail to capture the levels of steel price changes or become incorrect at predicting the direction of price movements.

To sum up, as for the hypothesis 1, I reject both parts and confirm that: within the context of the steel billets, the futures price contracts are neither unbiased nor efficient predictors for the future spot prices. As for the hypothesis 2, I am able to reject it partially. For hypothesis 2-a, I reject, i.e. if the steel convenience yield is integrated of order 1, the steel convenience yield spread fail to predict the future steel spot price change. For hypothesis 2-b, the situation is a little bit tricky. I cannot completely reject it. In other words, if the steel convenience yield is not integrated of order 1, the convenience yield, proxied by the one-period lagged spot price and futures price have statistically significant power explaining the price variations of the LME steel billet spot price in the future. I want to make it clear that even though I reject most of the hypotheses, it does not mean that steel billet price follows a random walk. Flaws in the model setup and data do not infer failure in theory.

Another main conclusion from this thesis is to reiterate that even though both models generate biased prediction results, futures price series still contain important information for future spot price prediction. Plus, futures price help to derive other important pricing components such as the convenience yield and steel-specific risk premium.

9.1 Comments on Speculation

In addition to the limitations of the model and data explained in chapter 8, I believe there is another important source of disturbance, which seems external: heavy speculations. This external disruption offers another explanation why scholars tend to shy away from the research of commodities. In many occasions, the commodity markets are severely condemned for active speculations, especially in 2008 when prices skyrocketed for most products. The price upheavals during the last three years are accounted as the evidence of heavy speculation, which is also the reason why most of the theoretical models fail empirically. Indeed, speculators distort the normal and fair prices in the market. As a result, the relationship between futures prices and spot price is twisted. But, in order to identify distortions and understand how the price behaves with speculative disturbances, it is necessary to start with understanding how normal undistorted prices should have behaved theoretically.

This thesis, even though failed to establish the merits of either UFM or CYM model, takes the initiative to test the steel billet price empirically. It also implements the process to model some economic measures, such as the convenience yield and the steel specific risk premium, which are relevant to the further researches on similar products. In the long run, the steel prices are determined by the equilibrium of the consumption, resource exploitation, production and inventory. Moreover, innovations in exploration and extraction

methods, geopolitics, cartels and environmental regulation also have impacts on the prices. This thesis focuses on the price behavior of the steel, so only spot and futures prices are of interest in this paper.

10 Suggestions for Future Research

This chapter briefly brings forth an alternative structural approach using the dynamic programing techniques. The author believes that this approach can better grasp the stochastic behavior of the steel billet prices for future researches. However the disadvantages of this approach: over-complexity and curse of dimensionality are also illuminated.

As explained in chapter 2, there is another spectrum of thinking in commodity research. Scholars from this school of thinking usually study commodity price behavior by adopting the structural model approach and using dynamic programming techniques to model variables that will affect commodity prices, directly or indirectly. Several scholars pioneer at researches in this field. Most notably, Pirrong (2012) makes an extensive study on many commodity types from continuously produced and consumed products to seasonal products, from storable goods to non-storable energy products. Similar endeavors can also be found in Schwartz & Smith (2000) and Ncube, et al. (2010). Given the complexity of this approach, the maths and econometric requirements on the researchers are demanding. For example, most of these received literatures include recursively using advanced time series technique such as the Kalman Filter (See Pirrong, 2012). In addition, for most of the occastions, the original Kalman Filter model cannot be used directly. For instance, if the price trajectory is nonlinear, which is true for most cases, an Extended Kalman Filter needs to be developped to meet the unlinearity criteria.

In addition to the mathematical challenges, the structural model approach also involves more computational costs, for example, most commodities' daily price series usually contain thousands of observations. A full run on these commodities can take a modern computer (2.5 GHz CPU, 2012) several hours to finish necessary simulation and calibration for parameter estimation. This is an important reason that this thesis had to let go of the structural model. However, the author believes in the high-end of the finance research, this promising approach should be considered as the future alternative to study the commodity prices.

Last but not least, as mentioned multiple times, some important steel raw materials such as the iron ore has no futures market yet. LME, NYMEX or exchanges alike should work with oligopolistic raw material suppliers such as BHP Billiton, Alcoa, Newmont Mining, etc. They might reach a collaboration to create futures market for the iron ore. Then, there will be more transparency in both the steel and the iron ore market. Hence, scholars can study the relationship between these two important commodity price behaviors and establish models with better price fit and more power to predict future prices.

11 References

Abdi, H. & Williams, L., 2010. Principal Component Analysis. *Wiley Interdisciplinary Reviews: Computational Statistics,* 2(4), pp. 433-459.

Adams, W. & Dirlam, J. B., 1964. Steel Imports and Vertical Oligopoly Power. *American Economic Review,* 54(5), p. 626.

Baur, D. G. & Lucey, B. M., 2010. Is Gold a Hedge or a Safe Haven? An Analysis of Stocks, Bonds and Gold.. *Financial Review,* 45(2), pp. 217-229.

Baur, D. G. & McDermott, T. K., Aug 2010. Is Gold a Safe Haven? International Evidence. *Journal of Banking & Finance,* 34(8), pp. 1886-1898.

Bessembinder, H., Coughenour, J. F., Seguin, P. J. & Smoller, S. M., 1995. Mean Reversion in Equilibrium Asset Prices: Evidence from the Futures Term Structure. *Journal of Finance,* 50(1), pp. 361-375.

Blas, J., Farchy, J. & Meyer, G., 2012. *Bruising year for commodities hedge funds.* [Online] Available at: http://www.ft.com/intl/cms/s/0/9179ec9a-4c2f-11e1-bd09-00144feabdc0.html#axzz1uAQbXDLa [Accessed 5 May 2012].

Bollerslev, T., 1986. Generalized Autoregressive Conditional Heteroskedasticity. *Journal of Econometrics,* pp. 307-327.

Brennan, M., 1958. The Supply of Storage. *American Economic Review,* Volume 48, pp. 50-72.

Campbell, J. Y., Lo, A. W. & Mackinlay, A. C., 1997. *The Econometrics of Financial Markets.* Princeton: Princeton University Press.

Campell, J. Y. & Shiller, R. J., 1987. Cointegration and Tests of Present Value Models.. *Journal of Political Economy,* 95(5), p. 1602.

Cargill, T. F. & Rausser, G. C., 1969. *Futures Prices Behavior as a Stochastic Process.* New York, Paper presented at the American Statistical Association Meetings.

Carpantier, J.-F., 2010. *Commodities Inventory Effect.* Louvain, Belgium: Center for Operations Research and Econometrics, Discussion Paper.

Chernenko, S. V., Schwartz, K. B. & Wright, J., 2004. *The Information Content of Forward and Futures Prices: Market Expectation and the Price of Risk,* s.l.: International Finance Discussion Paper 808. Board of Governors of the Federal Reserve System.

Chinn, M. D., LeBlanc, M. & Coibion, O., 2005. The Predictive Content of Energy Futures: An Update on Petroleum, Natural Gas, Heating Oil and Gasoline. *National Bureau of Economic Research Working Papers,* Issue 11033.

Coats, A., Kang, D. & Lyssogorskaya, V., 2008. *Emerging Markets Now Drive Global Steel Demand,* Hong Kong: HSBC Metals and Mining.

Cochrane, J. H., 2005. *Asset Pricing.* Revised ed. Princeton: Princeton University Press.

Coimbra, C. & Esteves, P. S., 2004. Oil Price Assumptions in Macroeconomic Forecasts: Should We Follow Futures Market Expectations?. *OPEC Review: Energy Economics & Related Issues,* 28(2), pp. 87-106.

DataStream, 2012. *Thomson Reuters,* s.l.: Thomson Reuters.

De Long, B. J., Shleifer, A., Summers, L. H. & Waldmann, R. J., 1990. Noise trader risk in financial markets.. *Journal of Political Economy,* 98(4), p. 703.

Deaton, A. & Laroque, G., 1992. On the Behavior of Commodity Prices. *Review of Economic Studies,* Volume 59, pp. 1-23.

Deaton, A. & Laroque, G., 1995. Estimating a Nonlinear Rational Expectations Commodity Model with Unobservable State Variables.. *Journal of Applied Econometrics,* Volume 10, pp. 9-40.

Deaton, A. & Laroque, G., 1996. Competitive Storage and Commodity Price Dynamics. *Journal of Political Economy,* Volume 104, pp. 896-932.

Deng, S., 1999. Stochastic Models of Energy Commodity Prices and Their Applications: Mean Reversion with Jumps and Spikes. *Georgia Institute of Technology Working Paper.*

Dickey, D. A. & Fuller, W. A., 1979. Distribution of the Estimators for Autoregressive Time Series with a Unit Root. *Journal of the American Statistical Association,* pp. 427-431.

Dickey, D. A. & Fuller, W. A., 1981. Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root. *Econometrica,* pp. 1057-1072.

Duffi, D., 1992. *Dynamic Asset Pricing Theory.* Princeton, NJ.: Princeton University Press.

Engle, R. F. & Granger, W. J., 1987. Co-Integration and Error-Correction: Representation, Estimation, and Testing.. *Econometrica,* Volume 55, pp. 251-276.

Eydeland, A. & Geman, H., 1995. *Asian Options Revisited: Inverting the Laplas Transform.* s.l., RISK.

Fama, E. F., 1965. Random Walks in Stock Market Prices. *Financial Analysts Journal,* 21(5), p. 56.

Gabillon, J., 1991. The Term Structure of Oil Futures Prices. *Oxford Institute for Energy Studies (Working Paper),* Issue 17.

Geman, H., 2005. *Commodities and Commodity Derivatives: modeling and pricing for agriculturals, metals, and energy.* Chichester, West Sussex, England: John Wiley & Sons Ltd.

Geman, H. & Nguyen, V.-N., 2005. Soybean Inventory and Forward Curve Dynamics. *Management Science,* 51(7), pp. 1076-1091.

Gemen, H., 2005. *Commodities and Commodity Derivatives: modeling and pricing for agriculturals, metals, and energy.* Chichester, West Sussex, England: John Wiley & Sons Ltd.

Gibson, R. & Schwartz, E., 1990. Stochastic Convenience Yield and the Pricing of Oil Contingent Claims. *Journal of Finance,* Volume 45, pp. 959-976.

Gordon, M. J., 1958. Dividends, Earnings and Stock Prices. *Review of Economics and Statistics,* 41(2), pp. 99-105.

Granger, C. W. & Newbold, P., 1974. Spurious Regressions in Econometrics. *Journal of Econometrics,* pp. 111-120.

Grossman, S. J., 1977. The Existence of Futures Markets, Noisy Rational Expectations and Informational Externalities.. *Review of Economic Studies,* 44(138), p. 431.

Gustafson, R., 1958. *Carryover levels for grains: a method for determining amounts that are optimal under specified conditions.* 1178 ed. s.l.:United States Department of Agriculture Technical Bulletin.

Hamilton, J., 1994. *Time Series Analysis.* Princeton: Princeton University Press.

Hong, H., 2000. A Model of Returns and Trading in Futures Market. *The Journal of Finance,* LV(2).

Hong, H. & Yogo, M., 2011. What does Futures Market Interest Tell Us about the Macroeconomy and Asset Prices?. *National Bureau of Economic Research, working paper.*

Hull, J. C., 2012. *Options, Futures and Other Derivatives.* 8th ed. Harlow, Essex, England: Pearson Education Limited.

IndexMundi, 2012. *IndexMundi.* [Online] Available at: http://www.indexmundi.com/commodities/ [Accessed 29 4 2012].

Jagannathan, R., 1985. An Investigation of Commodity Futures Prices Using the Consumption-based Intertemporal Capital Asset Pricing Model. *Journal of Finance,* Issue 40, pp. 175-191.

Kaldor, N., 1939. Speculation and Economic Stability. *The Review of Economic Studies,* Volume 7, pp. 1-27.

Keynes, J., 1930. *The Applied Theory of Money.* London: Macmillan & Co..

Knetsch, T., 2007. Forecasting the Price of Crude Oil via Convenience Yield Predictions. *Journal of Forecasting,* Issue 26, pp. 527-549.

Koekebakker, S. & Ollmar, F., 2005. Forward Curves Dynamics in the Nordic Electricy Market. *Managerial Finance,* Volume 31.

Kumar, S., Ghildayal, N. & Ostor, C., 2008. A Systems Approach in Examining Optimization Opportunities and Dynamics of the Global Steel Industry. *Information Knowledge Systems Management,* Volume 7, pp. 401-427.

Larson, A. B., 1960. Measurement of a Random Process in Futures Prices. *Food Research Institute Studies ,* Volume 1, pp. 313-324.

Lee, J. W. & Zhang, Y., 2009. Evidence on Normal Backwardation and Forecasting Theory in Futures Market. *Journal of Derivatives & Hedge Fund,* Issue 15, pp. 158-170.

Leuthold, R. M., 1972. Random Walk and Price Trends: The Live Cattle Futures Market.. *Journal of Finance,* 27(4), pp. 879-889.

Litzenberger, R. & Rabinowitz, N., 1995. Backwardation in Oil Futures Market: Theory and Empirical Evidence. *Journal of Finance,* 50(5), pp. 251-255.

Litzenberger, R. & Rabinowitz, N., 1995. Backwardation in Oil Futures Markets: Theory and Empirical Evidence. *Journal of Finance,* 50(5), pp. 1517-1545.

London Metal Exchange, 2008. *LME Steel Billet.* [Online] Available at: http://www.lme.com/steel/ [Accessed 3 May 2012].

Mankiw, G., 2002. *Macroeconomics.* 5th ed. New York: Worth Publishers.

Merton, R., 1973. An Intertemporal Capital Asset Pricing Model. *Econometrica,* Volume 41, pp. 867-887.

Moosa, I. A. & Al-Loughani, N. E., 1994. Unbiasedness and Time Varying Risk Premia in the Crude Oil Futures Market. *Energy Economics,* 16(2), pp. 99-105.

Ncube, M., Doran, J. & Srivastava, A., 2010. *Calibration of Reduced-Form Models; Implementation of a Kalman-Smoother Expectation Maximizaiton Procedure.* Miami, Florida: Florida State University.

Ng, V. & Pirrong, C., 1994. Fundamentals and Volatility: Storage, Spreads and the Dynamics of Metal Prices. *Journal of Business,* Volume 34, pp. 203-230.

Pagano, P. & Pisani, M., 2009. Risk-Adjusted Forecasts of Oil Prices. *B.E. Journal of Macroeconomics: Topics in Macroeconmics,* 9(1).

Pindyck, R. S., 1979. The Cartelization of World Commodity Markets.. *American Economic Review,* 69(2), pp. 154-158.

Pindyck, R. S., 1993. The Present Value Model of Rational Commodity Pricing. *Economic Journal,* 103(418), pp. 511-530.

Pindyck, R. S., 1994. Inventories and the short run dynamics of commodity prices.. *RAND Journal of Economics,* 25(1), pp. 141-159.

Pindyck, R. S., 2001. The Dynamics of Commodity Spot and Futures Markets: A Primer.. *Energy Journal,* 22(3).

Pindyck, R. S., 2004. Volatility and Commodity Price Dynamics. *Journal of Futures Markets,* 24(11), pp. 1029-47.

Pindyck, R. S. & Rotemberg, J. J., 1990. The Excess Co-Movement of Commodity Prices. *Economic Journal,* 100(403), pp. 1173-1189.

Pirrong, C., 2012. *Commodity Price Dynamics: A Structural Approach.* 1st ed. Cambridge, New York, Melbourne, Madrid, Cape Town, Singapore: Cambridge University Press.

Rogers, C. D. & Robertson, K., 1987. Long term contracts and market stability: The case of iron ore. *Resources Policy,* 13(1), pp. 3-18.

Routledge, B., Seppi, D. & Spatt, C., 2000. Equilibrium Forward Curves for Commodities.. *Journal of Finance,* Volume 55, pp. 1297-1338.

Samuelson, P. A., 1965. Proof That Properly Anticipated Prices Fluctuate Randomly.. *Industrial Management Review,* 6(2), p. p41.

Scheinkman, J. & Schechtman, J., 1983. A Simple Competitive Model with Production and Storage. *Review of Economic Studies,* 50(3), pp. 427-441.

Schneeweis, T. & Spurgin, R., 1997. Comparisons of Commodity and Managed Futures Benchmark Indexes. *Journal of Derivatives,* 4(4), pp. 33-50.

Schwartz, E., 1995. Review of Investment Under Uncertainty. *Journal of Finance,* Volume 50, pp. 1924-1928.

Schwartz, E. & Smith, J., 2000. Short-Term Variations and Long-Term Dynamics in Commodity Prices. *Management Science,* 46(7), pp. 893-911.

Sidney, S. A., 1961. Price Movements in Speculative Markets: Trends or Random Walks. *Industrial Management Review,* 2(2), p. 7.

Steel Business Briefing, 2012. *Steel risk management tools - are they new?*. [Online] Available at: http://www.steelbb.com/steelfutures/ [Accessed 4 May 2012].

Steel Times International, 2005. LME announces its decision on steel futures.. *Steel Times International*, Nov/Dec, p. 6.

Steel Times International, 2008. LME Steel first day.. *Steel Times International*, May/June, p. 4.

Stevenson, R. A. & Bear, R. M., 1970. Commodity Futures: Trends or Random Walks. *Journal of Finance,* 25(1), pp. 65-81.

Tcha, M. J. & Wrigth, D., 1999. Determinants of China's import demand for Australia's iron ore. *Resources Policy,* 25(3), pp. 143-149.

Team, I. O., 2011. *CISA Finds Price Manipulation in Iron Ore Market, Suggesting National Reserve Strategy*. [Online] Available at: http://www.ironoreteam.com/content/cisa-finds-price-manipulation-iron-ore-market-suggesting-national-reserve-strategy
[Accessed 3 June 2012].

Telser, L., 1958. Futures Trading and the Storage of Cotton and Wheat. *Journal of Political Economy,* Volume 66, pp. 233-255.

Treasury, U. D. o. t., 2012. *Daily Treasury Bill Rates*. [Online] Available at: http://www.treasury.gov/resource-center/data-chart-center/interest-rates/Pages/TextView.aspx?data=billrates [Accessed 14 May 2012].

Tse, Y., 2002. Residual-Based Diagnostics for Conditional Heteroscedasticity Models. *Economics Journal ,* 5(2), pp. 358-373.

West, M. & Harrison, J., 1996. *Bayesian Forecasting and Dynamic Models*. 2nd ed. New York: Springer-Verlag.

Williams, J. & Wright, B., 1991. *Storage and Commodity Markets*. Cambridge: Cambridge University Press.

Working, H., 1933. Price Relations Between July and September Wheat Futures in Chicago since 1885. *Wheat Studies,* Issue 9, pp. 186-238.

Working, H., 1934. Price Relations Between May and New-Crop Wheat Futures at Chicago Since 1885. *Wheat Studies,* X(5), pp. 184-228.

Working, H., 1949. The Theory of the Price of Storage. *American Economic Review,* Volume 39, pp. 1254-1262.

Working, H., 1958. A Theory of Anticipatory Prices. *American Economic Review,* Volume 63, pp. 188-199.

Wu, T., 2005. Do Oil Futures Prices Help Predict Future Oil Prices. *FRBSF Economic Letter*, 30 12, p. 38.