GLOBAL COMMODITY FUTURES MARKET MODELLING AND STATISTICAL INFERENCE

by

WEIQING TANG

A thesis submitted to the University of Birmingham for the degree of DOCTOR OF PHILOSOPHY

Department of Economics
Birmingham Business School
College of Social Science
University of Birmingham
This unpublished thesis/dissertation is copyright of the author and/or third parties. The intellectual property rights of the author or third parties in respect of this work are as defined by The Copyright Designs and Patents Act 1988 or as modified by any successor legislation.

Any use made of information contained in this thesis/dissertation must be in accordance with that legislation and must be properly acknowledged. Further distribution or reproduction in any format is prohibited without the permission of the copyright holder.
Abstract

This thesis first investigates the asset pricing ability of a new risk factor, namely Risk-Neutral Skewness (estimated based on option data) in the global commodity futures market. Skewness trading behaviour in the option market is attributed to heterogeneous belief and selective hedging concern. The negative (positive) the Risk-Neutral Skewness is accompanied with excess trading on put (call) option contracts, which leads to underlings’ over-pricing (under-pricing). Above results are robust to time-series and cross-sectional test and other alternatives.

Secondly, a new functional mean change detection procedure is proposed via the Kolmogorov-Smirnov functional form. Simulations indicate decent testing power under the alternative. An empirical test procedure is deployed for crude oil and gold futures price term structure, showing real market data change. The multivariate forecasting regression analysis uncovers trading behaviours behind the real-world change occurrence.

Lastly, the futures basis term structure is forecasted under the framework of the functional autoregressive predictive factor model with lag 1. By comparison, the new method outperforms other functional and non-functional methods, with maturities less than 10 months. The Model Confidence Set method statistically validate this result. A new variance minimization trading strategy is proposed and tested when the future futures basis is forecast and known.
Acknowledgements

First and foremost, I sincerely appreciate my supervisors, Prof Zhenya Liu and Prof David Dickinson, for their academic guidance and incredible support during my PhD study. Prof Zhenya Liu has introduced me a scientific world of doing research independently, critically and practically. Prof David Dickinson has made great support on my research activities, without his generous help, my research journey will be tougher. My special thanks are for Prof Lajos Horváth from University of Utah who gave me a strong guidance on statistical analysis and the newest statistical theory. Without his patient and considerate instructions, I will not put my thesis in the current standard. Moreover, I would like to take this opportunity to thanks my viva examiners, Dr Pei Kuang and Professor Ana-Maria Fuertes. It is their comments and suggestions making this thesis better and better.

My great appreciation goes to my parents who support me through my whole high degree pursuing. It is their unconditional love that makes me moving forward and challenging further. Additionally, my heartfelt thanks go to my girlfriend, Miss Jingyi Yang, who always trust on me with countless praise on my research. More importantly, it is her accompany lighting the dark road and encouraging me challenge every possibility.

Last, but not least, I am grateful for the financial support from the department of economics, business school, University of Birmingham.
2.3.1 Out-Performance of Risk-Neutral Skewness ........................................... 26
2.3.2 Pricing Mechanism of Risk-Neutral Skewness ...................................... 27
2.4 Data Description ....................................................................................... 30
2.5 Methodology ........................................................................................... 30
  2.5.1 Risk-Neutral Skewness .......................................................................... 31
  2.5.2 Commodity Market Variables ............................................................. 34
  2.5.3 Risk Factor Portfolio Construction ....................................................... 36
2.6 Empirical Results ..................................................................................... 37
  2.6.1 Summary Statistics for Return and Commodity Risk Factors ............... 37
  2.6.2 Trading Strategy Performance ............................................................. 39
  2.6.3 Risk-Neutral Skewness Characteristics .............................................. 44
  2.6.4 Time Series Analysis ........................................................................... 47
  2.6.5 Robustness Check ................................................................................ 50
  2.6.6 Cross Sectional Analysis ..................................................................... 54
2.7 Conclusion ................................................................................................. 59

3 Term Structure Mean Change Detection .................................................. 61
  3.1 Introduction ............................................................................................. 62
  3.2 Dynamic Nelson-Siegel Model Estimation .............................................. 67
  3.3 Functional Models and Projected Method ................................................. 70
  3.4 Simulation Study ...................................................................................... 78
  3.5 Empirical Finite Sample Analysis ............................................................ 86
    3.5.1 Data Description ................................................................................ 87
    3.5.2 Dynamic Nelson-Siegel model Estimation ....................................... 89
List of Tables

2.1 Summary Statistics for Commodity Futures Return ................................. 38
2.2 Portfolio Performance Statistics for Commodity Risk Factors .................... 39
2.3 Pair-Wise Correlation Matrix of Risk Factors ........................................ 40
2.4 Risk-Neutral Skewness Quartile Characteristics ...................................... 45
2.5 Time Series Analysis – Portfolio Property based on Risk-Neutral Skewness .... 49
2.6 Risk-Neutral Skewness Time-Series Portfolio Performance Test .................. 53
2.7 Risk-Neutral Skewness Time-Series Portfolio Performance Test Continues ...... 54
2.8 Risk-Neutral Skewness Cross-Sectional Fama_Macbeth Regression ............... 57
2.9 Risk-Neutral Skewness Cross-Sectional Fama_Macbeth Regression Continues ... 58
3.1 Simulation Size and Power for DNS Projection Method .............................. 81
3.2 Simulation Size and Power for Full Functional Projection Method ................. 82
3.3 Summary Statistics for Gold Price Term Structure ................................... 94
3.4 Summary Statistics for Light Crude Oil Price Term Structure ...................... 95
3.5 Empirical Functional Change Point Detection Results ............................... 96
3.6 Pairwise Correlation Matrix for Gold CFTC Participators Position ............... 99
3.7 Pairwise Correlation Matrix for Light Crude Oil CFTC Participators Position .... 99
## LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.8</td>
<td>Gold Subsample Regression Analysis under Hedging Pressure Effect</td>
<td>100</td>
</tr>
<tr>
<td>3.9</td>
<td>Light Crude Oil Subsample Regression Analysis under Hedging Pressure Effect</td>
<td>101</td>
</tr>
<tr>
<td>3.10</td>
<td>Gold Subsample Regression Analysis under Managed-Money Trading</td>
<td>104</td>
</tr>
<tr>
<td>3.11</td>
<td>Light Crude Oil Subsample Regression Analysis under Managed-Money Trading</td>
<td>105</td>
</tr>
<tr>
<td>3.12</td>
<td>Gold Subsample Regression Analysis under Swap-Dealer Trading</td>
<td>107</td>
</tr>
<tr>
<td>3.13</td>
<td>Light Crude Oil Subsample Regression Analysis under Swap-Dealer Trading</td>
<td>108</td>
</tr>
<tr>
<td>4.1</td>
<td>Summary Statistics of Future Basis Term Structure on Individual Maturity Level</td>
<td>126</td>
</tr>
<tr>
<td>4.2</td>
<td>Subsample Forecasting Principals Comparisons</td>
<td>128</td>
</tr>
<tr>
<td>4.3</td>
<td>Futures Basis Term Structure Shape Preserving on Time Series Dimension</td>
<td>133</td>
</tr>
<tr>
<td>4.4</td>
<td>Performance of Backwardation Strategy on Individual Maturity Level</td>
<td>138</td>
</tr>
<tr>
<td>4.5</td>
<td>Model Confidence Set Testing with Root Mean Square Error Loss Function</td>
<td>141</td>
</tr>
<tr>
<td>4.6</td>
<td>Model Confidence Set Testing with Mean Absolute Error Loss Function</td>
<td>142</td>
</tr>
<tr>
<td>4.7</td>
<td>Model Confidence Set Testing with Maximum Drawdown Loss Function</td>
<td>143</td>
</tr>
<tr>
<td>4.8</td>
<td>Model Confidence Set Testing with Downside Deviation Loss Function</td>
<td>144</td>
</tr>
<tr>
<td>4.9</td>
<td>Performance of Trading Strategy on Individual Maturity Level</td>
<td>149</td>
</tr>
<tr>
<td>4.10</td>
<td>Performance of Trading Strategy with Transaction Cost on Individual Maturity Level</td>
<td>153</td>
</tr>
</tbody>
</table>
List of Figures

2.1 Cumulative Log Return of Risk Factors Sorted Portfolios ........................................ 41
2.2 Risk-Neutral Skewness based Quartile Portfolio Cumulative Log Return .................. 42
2.3 Skewness Comparison Analysis - Parametrization Problem .................................. 43
2.4 Extreme Quartile Commodity Assets Overlapping Analysis ................................ 46
2.5 Bootstrap Time-Series RNSK Regression Alpha Dynamics .................................... 51

3.1 Sensitivity of Dynamic Nelson-Siegel Model to Lambda Values .............................. 69
3.2 Sensitivity of Testing Power on Change Point Magnitude with Sample Size 250 ....... 84
3.3 Sensitivity of Testing Power on Change Point Magnitude with Sample Size 500 ....... 85
3.4 Projected Factor Dynamics for Assets Cross Financial Crisis Period ..................... 90
3.5 Price Term Structure 3-Dimensional Surface Plot ............................................... 93

4.1 Boxplot of Forecasting Error .................................................................................. 129
4.2 Maturity based Forecasting Error .......................................................................... 130
4.3 Forecasting Error Dynamics .................................................................................. 131
4.4 Forecasting Term Structure Shape Preserving Dynamics ....................................... 134
4.5 Strategy Performance based on Backwardation Theory for All Maturities ............... 137
4.6 Futures Contracts based New Trading Strategy Performance ............................... 151
Chapter 1

Introduction

This chapter firstly pin down the main commonalities across all the following chapters within this thesis, then provide a review on the global commodity futures market research background, both theoretical and empirical studies, and finally discuss the research motivation, research question, contribution, future potential improvements and works.
1.1 Research General Commonality and Framework

The first commonality comes out for organising these three separate chapters are on asset modelling, presented in different methods and perspectives. Generally, spanning across three chapters, global commodity futures modellings are investigated on the basis of: (1) asset return anomaly explanation via asset pricing framework, using multivariate risk factor approach, (2) futures contract price term structure modelling test via the functional project method with projected factors from Dynamic Nelson-Siegel (DNS) framework and (3) futures basis term structure fitting via the functional autoregressive model.

Another close link across these three parts is to understand futures market backwardation and contango: (1) documenting a new risk factor distinct from the term structure basis factor (not explained by the backwardation and contango theory), (2) discussing the term structure price data modelling and instability (market shift between backwardation and contango) and (3) fitting and forecasting the term structure futures basis (market backwardation and contango shape forecast).

Standing on the current literature of global futures market studies’ results, chapter two employs a new risk factor, the Risk-Neutral Skewness, to price the global commodity futures market return. This new risk factor shows successfully extraordinary performance and heterogeneity (less correlated with several traditional risk factors), which is distinct from the term structure basis factor and its corresponding theory behind: backwardation and contango. A clear pricing mechanism is also discussed on the basis of a strong theoretical background (option-based buying pressure from heterogeneous belief and selective hedging idea) to support empirical findings. Moreover, in order to provide some practical advantages as to why this new factor matters, the recent literature proposed realised skewness estimator (documented by the cumulative prospect theory) is also included to compare.

Different from single continuous futures data modelling, the third chapter considers more generic
case of modelling term structure price data and testing model stability. However, in literature, the
current term structure modelling studies mainly depend on the DNS model that is originally created
to deal with bond yield curve data. Consequently, chapter three will then explore whether the DNS
model is reasonable to accommodate the commodities futures term structure variation. Method used
for this analysis is concentrated on statistical data mean change detection, which incorporates DNS
model framework. If a mean change is detected on DNS implied factors, which implies that DNS
model is statistically not able to adjust data dynamics (mean level) with respect to corresponding
samples.

In most cases of mean changes, there are impacts on the term structure slope, indicating the
change on market shift between backwardation and contango. Therefore, by implementing this detec-
tion procedure, economic intuition for futures market characteristics behind this change are able to be
discussed. As to the testing details, a new detection procedure is proposed to handle this change de-
tection under the functional data analysis framework. The results are examined under both simulation
and empirical scenarios.

In terms of predicting term structure, the functional autoregressive method is employed in the
chapter four, which shows a good property on term structure data fitting and forecasting. The back-
wardation and contango scenarios, heavily dependent on the shape of term structure curve on the view
of data pattern, is also explored and discussed when using the functional autoregressive model.

All current empirical findings and related contributions are summarised in the last chapter. Fol-
lowing the summary, current research limitations and future research possibilities is also discussed.
CHAPTER 1. INTRODUCTION

1.2 Understanding Global Futures Market Modelling

The first question initializing the futures market research is whether the futures market is integrated or segmented with the stock market? Answers for this question is converted to whether CAPM model (Sharpe, 1964; Lintner, 1965 and Mossin, 1966) works in the futures market and whether practical investors can benefit from futures contracts’ trading because of portfolio risk diversification (if two assets are segmented). In this sense, the existence of market co-movement between futures and equities, and assets correlation becomes the research point.

Investing and trading in commodity futures become more attractive with a considerable compound return 9.85% from 1950 to 1976, while trading in equity market yields 9.81% annual return (Bodie and Rosansky, 1980). At the same time, the buy-and-hold strategy on the Goldman Sachs Commodity Index (GSCI) has been outperforming the buy-and-hold strategy on S&P 500 by 1% since 1969, recorded at a 12.2% annual return. Moreover, these two assets indices are negatively correlated at -0.03, implying large diversification benefits by building portfolio for both asset classes (Erb and Harvey, 2005). Gorton and Rouwenhorst (2004) and Bhardwaj et al. (2015) document that the risk premium (excess return deviating from its average value) generated from trading in commodity futures is quite close to trading in the stock market, with 4.95% and 5.91% annually from 1959 to 2014.

Regarding the market integration test, the CAPM model is the best candidate using the equity market portfolio as a benchmark to capture the co-movement with the global commodity futures market. Consistent with the law of one price theory, a factor (stock market portfolio return) that can explain cross sectional the average return in equity should be able to explain the average return of the global commodity futures as well. However, there is no significance found via the CAPM model, extended CAPM model and traditional equity motivated approach, which states the heterogeneity of the commodity futures return (Jagannathan, 1985 and Erb and Harvey, 2005).
CHAPTER 1. INTRODUCTION

Dusak (1973), Kolb (1992) and Bessembinder (1992), among others, point out that commodities like wheat, corn and soy bean do not generate risk premium based on the CAPM asset pricing model. A similar result is also confirmed by Ehrhardt et al. (1987) who use a two-factor model under the framework of arbitrage pricing theory (APT). Based on the weighting index method by Marcus (1984), Baxter et al. (1985) and Black (1976) find that the systematic risk premium does not exist.

However, it is worth mentioning that the above result does hold only for agriculture products such as wheat, corn and soy bean. By allowing the existence of speculators’ net long and short position in futures market as well as inclusion of commodity index in market portfolio, Carter et al. (1983) document the market portfolio beta significance. This is consistent with Bodie and Rosansky (1980) who use more commodity products and longer testing samples and Fama and French (1987) who find time-varying risk premium for most commodity products.

The most comprehensive market segmentation test is proposed by Daskalaki et al. (2014), who among others, offer a large comparison study by employing the CAPM extended series model (CCAPM, MACPM, MCCAPM,...), macro-economic factors (consumption growth, money growth, FX factor...) and equity-motivated factors (FF, Carhart, LFF and LCarhart). The results indicate that there is no statistically transmission from the equity market to the global commodity futures market. Up to now, findings on the CAPM model based test is still mixed and ambiguous.

1.3 Theoretical Background

Motivated by the market segmentation evidence above, pricing theories for futures market are needed. In the following sections, the most widely cited traditional theories applied in the global commodity futures market are introduced and discussed. In addition, theories that are proposed in other asset markets are also referred for the sake of explaining new findings in this thesis,
1.3.1 Storage, Backwardation and Hedging Pressure

The early theories on commodity futures mainly refer to storage cost. The theory of storage is initially introduced by Working (1949) and Kaldor (1939), stating that the cost of carrying the underlying physical products needed to be priced (compensated). On the contrary, potential profits by selling physical products at a higher price under the scenario of commodity supply scarcity are also needed to be taken into account. These two directions formally shed light on a positive (negative) relation between commodity return and storage cost (convenience yield).

Moving forward and standing on the side of the production firm, Pindyck (1990) uses the decomposed cost of the holding inventory and states that a convex function shows a good fit on the convenience yield with respect to inventory level. Recently, linear regression (with inclusion of the squared convenience yield) from Dincerler et al. (2005) and the sensitivity analysis (with changes on the normalized inventory data in the spline regression analysis) from Gorton et al. (2007) confirm a non-linear relationship between convenience yield and storage.

The backwardation theory is introduced by Keynes (1930) who gives a new idea on how and why it is possible for contract buyers to hold the futures contract. When the futures contract is discounted priced relative to its expected spot price, the convergence between futures price and spot price at expiration date provide risk premium to the contract’s holders. Normally, discounted contracts are offered by large producers or those people who want to hedge their physical products transaction risk in a future time point. When above scenario is on opposite side, the market is then referred to contango (the futures contract price is higher than the expected spot price). Taken together, these two scenarios describe the market equilibrium in which either futures contract seller or buyers will obtain (bear) corresponding premiums (risk).

Hedging pressure (net or long only option interest position for both speculators and hedgers)
is another explanation of risk premium (see Stoll [1979], Hirshleifer [1988] and Hirshleifer [1990]).

The risk premium identification procedure is built on market participators modelling: inter-temporal portfolio optimisation and informational barriers.

Specifically, markets are classified as non-marketable and marketable and both speculators and hedgers are included, which is different from previous partial-equilibrium model. In the sense of equilibrium, the futures risk premium is then decomposed into a systematic part (market beta which is marketable) and residual part (non-marketable).

In the spirit of Merton (1987), speculators are restricted to market as they need to pay certain set up cost to know the market. And the residual risk premium is positively marginal to the number of participators in the market. When the cost for speculators to enter the market is large, risk premium then goes up as few speculators can take the whole premium from hedgers. Back to the theory, when hedgers goes short (long), equilibrium implies backwardation (contango).

1.3.2 Skewness Preference in the CAPM

Within the case of the asset pricing model, the mean-variance method, used for describing the behaviour of investors in the whole financial market, is not accurate or to some extent not valid. The argument for this is due to the break of normality assumption of the distribution of the underlying asset return, which results in the imprecise modelling of investors’ decision based solely on the expected mean and volatility. In the meantime, narrowing problem by assuming homogeneous rational idea is also not generally approachable across all scenarios (e.g. "Lotto Investors") in the real world.

More recently, skewness preference, defined as a specific portfolio construction behaviour of selecting assets with a strong skewness (normally positive skewed assets), becomes one important topic in both equity and futures markets. The economics explanation behind is firstly attributed to the

---

1 see following section on behaviour finance discussion.
extension of the pricing model of the CAPM on the equity market.

Following the Sharpe-Lintner asset pricing model, demonstrated by (Sharpe, 1964, Mossin, 1966 and Lintner, 1965), Arditti and Levy (1975) extend their two-factor model into a three-factor model that takes into consideration of skewness effect on assets allocation. Given the theoretical argument, skewness is positively related to both mean return and volatility, which leads to a new efficient frontier on which portfolios are strictly dominating others (those portfolios who are below or above this new frontiers). One year later, Kraus and Litzenberger (1976) show that after considering the utility function with concave property and positive skewness preference, systematic skewness is a necessity for model pricing.

Empirical tests also support their proposal with the coefficient for systematic skewness being positive and significant. This indicates a new fact that previous mispricing in equity market is not attributed to borrowing constraints or lending rates. Instead, it is due to the lack of the consideration on the systematic skewness effect in asset pricing model. This is confirmed by Lim (1989) who employs the generalized moments method (GMM) to identify the importance of skewness.

Unlike the result in the equity market, there is no strong evidence to support existence of skewness effect in the future market, especially when the systematic risk is controlled (Junkus, 1991). The reason why they fail to test the significance may be attributed to the wrong selection of market portfolio: S&P 500 and BLS wholesale price index.

Followed by Junkus (1991), Christie-David and Chaudhry (2001) adopt nine market indices as proxy for the market portfolios and co-skewness and co-kurtosis as the risk factor. In terms of the testing method from Fama-Macbeth two-step analysis (Fama and MacBeth, 1973), systematic skewness has strong relation with future period return. $R^2$ is increased when adding co-skewness and co-kurtosis in the asset pricing model, which is also robust regardless of the market portfolio selec-
1.3.3 Behaviour Finance

Most recently, research studies on skewness preference explanation, both in equity and futures markets, are more concentrated on topics linked to behaviour finance. By allowing heterogeneous preference on skewness consideration in portfolio investing, the one-period agent utility maximisation problem is proposed, which predicts lower future returns for the positive skewness preference (Mitton and Vorkink, 2007). These agents who have strong propensity on positive skewness are marked as "Lotto Investors". In the sense of distribution property, they expect to have large compensation from investing in these positively skewed assets.

Another theory backing up this skewness preference investing is formed on the cumulative prospect theory by (Tversky and Kahneman, 1992 and Barberis and Huang, 2008). Regardless of the appearance of short-selling restrictions, the cumulative prospect theory explains that positively skewed assets (large positive return with small probability) are over weighted by investors. A special value function, which is convex in loss and concave in gain is applied to model the assets’ selection behaviour. Given more expected value on potential extreme positive returns in the future, the more they pay for this asset, the more likely it is for this to result in underlying product over-pricing. The consequent buying (selling) assets with positive (negative) skewness will generate fewer (more) positive returns in the subsequent period.

Their ideas are also consistent with the optimal belief framework from Brunnermeier and Parker (2005) and Brunnermeier et al. (2007) who state that there is a type of agents who maximise their current utilities by distorting their beliefs on future probability (endogenous-probabilities model). The results confirm that assets with higher idiosyncratic skewness yield less profits.
Up to now, the application of testing skewness preference in assets’ selection explicitly refer to
the realised skewness, an estimator based on the past historical data. This thesis is moving forward
to consider the expected skewness, a forward-looking measure obtained from the underlying asset
option market data.

Considering the market trading restrictions (e.g. short-selling) or potential downside risk, Gar-
leanu et al. (2009) state that option trading demand effect matters. The selective hedging idea (Stulz, 1996) is another aspect supporting the option-trading behaviour as buying out-of-money put option
can be regarded as a risk control method.

Regarding limitations in the real world, investors holding negative (positive) expectations about
the futures price movement will purchase more put (call) option contracts. This in return generates
more negative (positive) skewness measures. Over-pricing (under-pricing) is then consequently em-
bedded in the underlying asset price realisation, which leads to future negative (positive) returns.
It is worth mentioning that this proposal has the pre-requisite that heterogeneous beliefs about the
underlying future movement should exist (Han, 2008).

1.3.4 Technical Analysis

The technical concepts in this thesis are mainly attributed to the stochastic analysis, functional data
analysis and regression analysis. To be more specific, for the second chapter, the option implied mo-
ments estimation process relied on the stochastic analysis. Asset return is modelled as the integration
from both call and put option contract pay-off structures. The third moment is then easy to obtain by
taking the expectation of return with power 3.

In terms of the risk factor estimation procedure, several potential errors advertised from the litera-
ture are taken into account. For the implied volatility estimation, both the Bisection and the Newton-
Raphson method under the pricing framework of the European Black Model (Black and Scholes, 1973) are employed. The natural cubic spline method is used to fit the implied volatility within mon-eyness boundary and linear or flat extrapolation for outside boundary (Jiang and Tian, 2005, Jiang and Tian, 2007 and Carr and Wu, 2008). The hermite cubic spline is adopted to account for the potential calendar arbitrage issue suggested from (Leontsinis and Alexander, 2017) when interpolating the final estimator in a constant time-to-maturity way.

The risk-neutral third moment is also concerned under different measurements. Therefore this the-sis shows a more comprehensive comparison for the literature, from two different methods by Bakshi et al. (2003) and Kozhan et al. (2013). The first proposal is the traditional central moment method under the option pricing framework while the second one is under the price martingale assumption. This thesis is relied on the second measure while reserving the first one for the robustness check.

As for the second and third chapter, technical details on the functional data structure analysis are referred to (Ramsay, 2006 and Horváth and Kokoszka, 2012). Under their suggestion, discrete sam-ples can be converted into functional observations. In this sense, the commodity futures term structure data generating process is then modelled in a functional way (data observation across maturities).

Based on the Hilbert space property, infinite dimensional curves data can be projected onto certain pre-determined factors (Dynamic Nelson-Siegel three factors) to reduce the dimensionality (Bards-ley et al, 2017). Using the Kolmogorov-Smirnov functional form under the CUSUM (cumulative sum) method frame, a new detection statistics is formed for the curve data mean change detection procedure. Both asymptotic property and simulation analysis are considered and studied.

In the end, the functional autoregressive model is proposed to model and forecast futures market term structure dynamics. The main technique used in this part is the functional predictive factors model (Kargin and Onatski, 2008). For the completeness modelling and forecasting check, compar-
is offered under the frame of the statistical method namely, the Model Confidence Set (Hansen et al., 2011). A further trading strategy via backwardation and new proposed variance minimization method are discussed in the end for the sake of providing economic implications.

1.4 Empirical Studies Background

In this section, the structure is summarized into two parts: the asset pricing (discussion on previous risk factors pricing the global commodity futures return) and the term structure modelling (extension from single series to commodity futures term structure modelling). The first part is the foundation for chapter two and chapters three and four are based on the second part.

1.4.1 Asset Pricing

The first empirical study recording risk premium in the global futures market is from Dusak (1973), who adopts the value-weighted S&P 500 index as market portfolio and indicates that commodities such as wheat, corn and soy bean do not generate risk premium based on the CAPM asset pricing model (the same results followed by Kolb (1992) and Bessembinder (1992)). His conclusion is that the results are not related to the Keynesian theory and the principal reason might be the no correlation between those single assets and market index he uses (few information from commodity agriculture is embedded in the market index). If replacing the commodity by copper that is more related to the industrial process (or the underlying economy development), the beta could be significant to some extent.

Instead of applying the one-factor model, Ehrhardt et al. (1987) propose a two-factor model to justify the risk premium within the APT framework and fail to find the significance. However, strong evidence of normal backwardation has been documented by Carter et al. (1983), who use the same
model by Dusak (1973) but give two more extensions: allowing speculators to be net long or net short, and inclusion of the commodity index as the weighted market portfolio.

Different from previous studies on the commodity futures risk premium findings above, Baxter et al. (1985) argue that the market index chosen by both of them is overweighed as the S&P 500 index has already included a certain percentage of commodity products (similar argument inspired by Black (1976)). After constructing a new index using a weight construction method consistent with the theory suggested by Marcus (1984), they confirm that wheat, corn and soy bean do not expose to systematic risk and are devoid of risk premium.

Similarly, Bodie and Rosansky (1980) show that the beta coefficient roughly equals to one when research data coverage is on more commodity products. It is worth mentioning that their first trail is also consistent with the result from Dusak (1973) with no significance when focusing on certain specific products. What they suggest is that the results are more precise when more products as well as longer sample periods are employed.

Moreover, Fama and French (1987) document the instantiations of return premium in futures basis, denoted as the differential of future price and spot price. They find its time-varying property and statistically significance when regressing excess return on it in most commodities out of 21 commodities.

More recently, motivated by the risk factor model pricing principal (factor based long-short portfolio) in the equity market, the term structure factor, sorted by the futures basis, is constructed, demonstrating positive correlation with the individual assets return (Koijen et al., 2013, Erb and Harvey, 2005, Szymanowska et al., 2014 and Fuertes et al., 2015). They all statistically confirm the pricing ability of the term structure factor.

In terms of hedging and speculating behaviour, the hedging pressure factor is proposed with pric-
ing capability in future market, which is also portfolio strategy sorted on market participators’ option position. Positive correlation is found between the hedging pressure factor and future return, which explains the cross-sectional variation among most commodities (Bessembinder 1992, De Roon et al. 2000 and Basu and Miffre 2013), while De Roon et al. (2000) specify that hedging pressure can be treated as a non-systematic risk premium.

The momentum factor, calculated based on the moving average of the past historical return, also shows non-trivial effect in global commodity futures asset pricing. The intuition behind is that investors are more likely to hold financial assets with positive past performance as they believe this positive trend will continue in the next period. The long-short quantile portfolio is designed based on the past moving average return and found with pricing ability (positive correlation) in cross-sectional return for most commodity futures (Asness et al. 2013, Erb and Harvey 2005 and Miffre and Rallis 2007).

However, literature for these factors pricing ability conclusion are still mixed. Daskalaki et al. (2014) document no significance on the commodity-specified risk factors mentioned above. Although they nearly reject all factors in the commodity futures market, consideration on the risk-neutral high moments’ effect is still missing, which leads to the uniqueness of this thesis in this field.

More recently, the popular risk factor exploration direction is more related to the “idiosyncratic” property of return distribution, for example, volatility and skewness. The “idiosyncratic” here refers to a more generic idea that factors is out of the control of the systematic risk factor in traditional literature. In another world, there could be a factor whose pricing ability cannot be explained by the common traditional risk factors or is real idiosyncratic part obtained from regression residuals.

---

2 The calculation method will be different for speculators and hedgers, in this thesis, speculators’ positions are used for main conclusion while hedgers’ positions are tested for robustness, for formula, see methodology part.

3 Which can be referred to this thesis’s results given that product selection in portfolio construction can make a difference.
CHAPTER 1. INTRODUCTION

Earlier studies about individuals’ skewness are often found in the equity market. Conditional skewness is tested for the equity market among the stock pool with different sorting criteria and subsample analysis, they find that conditional skewness can price to some extent but not general for all assets (Harvey and Siddique (2000)). One reason for their failure might be the imperfection of skewness measurement as it is not the ex-ante measure of the skewness based approach.

As for the attractive attribute of the ex-ante measure, Boyer et al. (2010) find that idiosyncratic volatility can be a good proxy variable to linearly estimate the expected idiosyncratic skewness. Based on the linear regression approach, the expected idiosyncratic skewness plays an excellent role generating 1% abnormal return monthly. This is recorded with a negative relation between skewness and subsequent return after controlling the Fama-French three-factor model. Their result is also consistent with Amaya et al. (2011) who use a new estimation method with intra-day (high frequency) data to measure the realised skewness and document that it has a significant negative relation on subsequent returns. The success of their findings is more related to the employment of high frequency data due to benefit of the improved estimation accuracy compared with the usage of low frequency data. Different from previous researches, real ex ante skewness is obtained in two different ways, implemented by Bali and Murray (2013) and Conrad et al. (2013), both of them find the same relation for the Risk-Neutral Skewness and expected return. However, their empirical results are more focused on the stock market rather than on the future market, which point out how the gap is filled in the literature.

1.4.2 Term Structure Modelling

Different from the factor asset pricing modelling mentioned above (long-short portfolio via single time series return data), the global commodity futures are modelled on term structure dimension. By

---

4Coefficient obtained by regressing asset return on squared market portfolio return, the same as co-skewness, measuring the co-movement between market return variance and single asset return.
meaning of term structure modelling, it requires the model to be able to account for the maturity effect. Traditional asset modelling on the underlying process requires more assets characteristics consideration and factor dynamics estimation, which often includes the complex stochastic process modelling and large number of parameters’ estimation, see convenience yield and spot price two-factor model stochastic model from (Gibson and Schwartz, 1990 and Schwartz and Smith, 2000), stochastic model comparison and three-factor model in which the mean-reverting process of interest rate following Vasicek (1977) is extended by Schwartz (1997), also see Casassus and Collin D. (2005), asset seasonality modelling consideration (Sørensen, 2002).

However, the stochastic analysis framework tends to be complex in practical analysis for which a more flexible and easily estimated method is more welcome to capture the forward curve movement, in fitting and prediction across maturities’ range. The fundamental technique is following yield curve modelling, namely Nelson-Siegel and its dynamics extension, see (Nelson and Siegel, 1987, Diebold and Li, 2006 and Diebold and Rudebusch, 2013). The idea underneath this model is to find the factors which can accommodate term structure curve change on different exposures: level shift (parallel move of curve for all maturities point), slope shift (more weight (less) on short (long) maturity contract) and curvature shift (more weight on middle maturity contract).

Apart from the original dynamics modelling, new proposals to the DNS model on the yield curve modelling can be referred to the following studies. Modelling on the decaying factor lambda dynamics by Koopman et al. (2010), completely different factors namely intelligible factor introduced by Lengwiler and Lenz (2010), regime switching effect from Nieh et al. (2010) and Xiang and Zhu (2013) and regime switching based Marco-factor concern from Zhu and Rahman (2015).

However, not too many works focusing on the application and extension of the Dynamics Nelson-Siegel (DNS) model in the futures market. In the global commodity futures analysis, the DNS model
application and its extensions are mainly on futures products of the energy section, e.g. Grønborg
and Lunde (2016) use the DNS model within copula framework obtaining a better out-of-sample pre-
diction performance than the benchmark. Barunik and Malinska (2016) include the neural network
method in the DNS model, forward curve local dependence is discussed in Ohana (2010), regime de-
pendence is introduced in an error vector correction model on dynamics of level, slope and curvature

Karstanje et al. (2017), among others, who first propose a comprehensive DNS model study on
all global commodity futures products with both seasonality and sector effect considerations. In their
work, DNS three factors are selected to pass in the modelling for the sake of avoiding over-fitting
and a new factor (trigonometric functions) mimicking the seasonality effect is also tested at the same
time.

1.5 Research Motivations, Questions and Contributions

In this section, research motivation, question and contribution are formalised here. The second chapter
is on the Risk-Neutral Skewness pricing effect test on global futures market, which is motivated by its
pricing success in equity market (although pricing sign is mixed, see, Conrad et al., 2013; Stilger et al.,
2016; Kozhan et al., 2013 and Gkionis et al., 2017) and its realised counterpart, Pearson skewness
success pricing in the futures market (Fernandez-Perez et al., 2018).

The research question comes out whether the Risk-Neutral Skewness estimated from the futures
option market can price global futures return both from a time-series and cross-sectional perspective?
Furthermore, what are the superior points on the risk-neutral measure when comparing with the his-
torical calculated one, exactly the Pearson skewness coefficient in Fernandez-Perez et al. (2018)? By
answering these questions, this thesis is making the contribution to literature that the Risk-Neutral
CHAPTER 1. INTRODUCTION

Skewness can explain assets’ return variation and outperforms the Pearson skewness.

The third chapter is inspired by the global futures term structure modelling. Since the DNS model is originally proposed to deal with yield curve term structure modelling, applying the DNS model directly on the futures market might not necessarily able to accommodate the futures market pricing characters. In the meantime, following the recent model extension on DNS (e.g. regime switching), the research questions about whether the DNS model provides a good fitting in futures market or at least in some sample periods. For answering, is there a statistical evidence to prove that DNS fails to do its work?

The contribution to this relies on a new proposing statistical detection method on the mean change test. Given the DNS model estimated factors, no change will be found if DNS has a good fitting on the samples. Asymptotic property is confirmed with simulation outcomes showing decent testing power on this new statistic. In the line of testing, an economic analysis, via multivariate forecasting regression, is also conducted to further identify the situation before and after changes in the data.

The fourth chapter is motivated by both third chapter results and functional data analysis advantages. Following the results of the third chapter, the DNS model might fail to capture the term structure dynamics in some scenarios. In another way, term structure data is also treated as non-smoothed curve data in a discrete version, which can be naturally modelled in the view of functional data analysis. Different from the literature idea of modelling term structure on the forward price, the futures basis (log price difference between two maturities for the same underlying product) is studied.

The research questions try to explore whether the functional autoregressive model can offer a better out-of-sample prediction compared with the DNS model and other functional candidates. To address this question, the statistical measurement as well as the Model Confidence Set test on the forecasting error and trading strategy performance are experimented, indicating the new method’s
superiority. Far more than this, the term structure curve reservation property across different methods is also well studied with final results supporting the outperformance of the new functional model. A new variance reduction trading strategy is designed for practical application on how forecasted futures basis can be used in the real world.

Up to this writing moment, all research contributions in this thesis are new to the global commodity futures literature. Several robustness check concerns are conducted to ensure the solidarity of the results.
Chapter 2

Risk-Neutral Skewness on Commodity Pricing

In this chapter, the asset pricing test framework is deployed for the new Risk-Neutral Skewness (RNSK) factors estimated from weekly 10-year options and futures return data. The final results have significant validity from both time series and cross-sectional tests. A positive relation is recorded between the future asset return and the current RNSK. Risk control based option trading activities (supported by the Heterogeneous Belief and Selective Hedging concern of underlying assets’ performance) provides the mechanism of trading signal generation. Under-pricing (positive RNSK) and over-pricing (negative RNSK) based long-short portfolio outperform its counterpart (based on the realised skewness, e.g. the Pearson method) with an additional 14.6% annual return. The results are robust to several alternatives: on signal estimation techniques, regression control analysis and transaction cost analysis.
2.1 Introduction

The asset pricing ability of the third moment of asset return distribution has been studied a lot recently, mainly in the equity market and few of them are focusing on the commodity futures market. Equity studies on skewness, e.g. the conditional skewness\(^5\) and realised skewness\(^6\) by (Harvey and Siddique, 2000 and Amaya et al., 2011), the expected idiosyncratic skewness\(^7\) proposed by Boyer et al. (2010) and the Risk-Neutral Skewness discussed by (Conrad et al., 2013, Kozhan et al., 2013 and Dennis and Mayhew, 2002). The findings with respect to the pricing relation to futures return are mixed.

The commodity literature on this subject is much sparser. At this moment of writing, the closest study in the commodity market is conducted by Fernandez-Perez et al. (2018) on the realised third central moment. They show that commodity returns are strongly negatively related to the realised skewness (measured by the standard Pearson skewness coefficient estimated based on monthly observations with a past 12-month length window of daily excess return\(^8\)).

The rationale for the realised skewness pricing in the commodity future market relies primarily on investors’ "lottery-type” preference. The behaviour of pursuing positive skewed commodities pushes their prices to a higher level as investors are willing to pay more for having the potential opportunity of gambling potential extreme positive compensations. However, the over-pricing of these positive

---

5. Conditional measure differs from unconditional one (e.g. Pearson coefficient skewness) which is independent of other variables effect, requires conditional information to compute. In this chapter only, conditional skewness is specifically defined as the beta coefficient value by regression asset expected return on squared market return, measuring the co-movement between market return variance (volatility) and asset return, \(\beta = \frac{\text{cov}(r, r_m^2)}{\sqrt{\sigma_r} \sqrt{\sigma_{r_m^2}}}, \) where \(r\) is asset return, \(r_m^2\) is squared market return, \(\sigma\) is a standard volatility measure function. For example, if the beta coefficient value is large, then when market becomes more volatile, asset return will change dramatically depending on sign of beta.

6. All realised skewness is generally referred as using the past historical data to calculate, which represents investors’ skewness measure with all the historical information they have, skewness calculation formula can be different, but in this thesis, it is referred to Pearson Skewness coefficient, \(\text{skew} = \frac{E[(r-\mu)^3]}{(E[(r-\mu)^2])^{3/2}}, \mu\) is the mean of \(r, E\) is expectation operator

7. Ideas here are based on linear relation between volatility and skewness, showing an idiosyncratic view to accommodate the high order (3rd order, skewness) distributional information excluded from baseline model. Idiosyncratic volatility is calculated on residuals (obtained from regression of asset return on baseline factors) in a rolling window manner. Then expected idiosyncratic skewness is forecasted by idiosyncratic volatility in a linear regression.

8. Research on skewness explanation via CAPM and conditional skewness effect can be referred to Junkus (1991) and Christie-David and Chaudhry (2001) respectively.
CHAPTER 2. RISK-NEUTRAL SKEWNESS ON COMMODITY PRICING  

skewed commodities will be corrected as arbitrage is not allowed in the market, yielding negative returns on these positively skewed assets. In this chapter, skewness pricing ability has been tested using a superior measure of skewness, called Risk-Neutral Skewness. The empirical results indicate that the Risk-Neutral Skewness is positively related to the future return. This is not completely contradicting previous findings via the realised Pearson skewness studied by Fernandez-Perez et al. (2018). According to the bivariate sort, the average values sorted by the realised Pearson skewness shows increasing trend from the lowest quantile to the highest one. The Risk-Neutral Skewness quantiles sorting for traditional commodity characteristics are also consistent with their findings in the realised skewness measurement.

Since the realised skewness asset pricing ability has been well discussed by Fernandez-Perez et al. (2018), this chapter contributes the literature by stating that the Risk-Neutral Skewness can do better than its realised counterpart. Specifically, this chapter will contribute the existing literature by two main points. Firstly, standard realised estimation of high moments (central moments like the Pearson skewness employed in Fernandez-Perez et al. (2018)) has been argued with a strong estimation bias. The parameters, in the realised skewness calculation, such as the past window length selection and the data frequency usage can deteriorate the final estimated results remarkably, which in return leads to a different strategy performance.

This chapter tries to avoid these problems by using a model-free third moment estimation method (based on the risk-neutral probability measure and all investors are assumed to hold the same risk preference). Risk-neutral third central moment is initially introduced by Bakshi and Madan (2000) and further tested by Bakshi et al. (2003), Dennis and Mayhew, 2002, Bali and Murray, 2013, Conrad et al. 2013, Stilger et al. 2016 and Gkionis et al. 2017, showing a non-ignorable pricing effect on the equity market even though the current findings are mixed.
More recently, Neuberger (2012) propose a more general unbiased estimator of the realised moments with only one assumption that underlying price process is a martingale. Following their spirits, Kozhan et al. (2013) propose the implied skewness calculation method. Under the framework of the Risk-Neutral Skewness estimation, daily commodity options data are utilised to calculate daily option-implied skewness without any arbitrary selection of window length as well as frequency. This chapter embarks on various aspects, such as truncation error, discrete estimation error, interpolation and extrapolation error to provide empirical robustness.

Up to this writing moment, there is no similar works studying the Risk-Neutral Skewness pricing effect in the global commodity futures market. Secondly, a comparison exploring for the difference between the realised skewness (calculated in the way of Fernandez-Perez et al. (2018)) and the Risk-Neutral Skewness is conducted in the view of investment and trading. For investors in the commodity futures market, the Risk-Neutral Skewness is shown to be superior to its counterpart (the realised skewness) in terms of profitability, sharp ratio, maxdrawdown and etc.

This chapter is organised as follows, section 2 will go through the literature that is more related to this chapter, section 3 gives more explanations on why the Risk-Neutral Skewness is better and the pricing mechanism behind for the commodity futures market, data and methodology are discussed in section 4 and 5 respectively, section 6 shows the corresponding empirical results and section 7 summarises all findings.

2.2 Background Literature

Moving to the underlying distribution argument, the mean-variance model, used to describe the behaviour of investors in the whole financial market, is not accurate or to some extent not valid due to unsatisfactory assumptions. Rather than simply focusing on return chasing and risk avoiding, portfo-
lio construction with skewness inclusion should be considered. [Arditti and Levy (1975)] extend above two-factor model to a three-factor one and prove that the efficient frontier dominates under the later scenario. Similarly, necessity of skewness in asset pricing is identified when using utility function with concave property and positive skewness chasing attribute ([Kraus and Litzenberger (1976)]. This new inclusion of skewness clarifies the idea that previous mispricing is not caused by the agents’ borrowing constraints or lending rates. Empirical study by [Lim (1989)] who finds the positive relation between skewness and return. Continuing in this frame, [Junkus (1991)] states no significance for skewness to be systematic risk under CAPM, while [Christie-David and Chaudhry (2001)] find significant results by adopting nine market indices as the market portfolios instead of S&P 500 and BLS wholesale price index used by [Junkus (1991)]. Results about co-skewness\(^9\) in literature are still ambiguous, which direct the research interests to the idiosyncratic property of individual asset skewness estimation.

Earlier studies about individuals’ skewness are on equity market. Conditional skewness is tested for equity market among stock pool with different sorting criteria and subsample analysis, [Harvey and Siddique (2000)] find that conditional skewness can price to some extent but not general for all assets.

One important reason for this imperfect pricing may due to non-ex ante property of skewness. [Boyer et al. (2010)] states the excellent role of expected idiosyncratic skewness (linear regression forecasted values based on idiosyncratic volatility), with 1% abnormal return monthly generated from pricing test. Negative relation between skewness and subsequent return is recorded even after controlling the Fama-French three factors. Their results are also consistent with [Amaya et al. (2011)\(^{10}\)] who use new estimation method with intra-day (high frequency) data on realised skewness measurement.

---

\(^9\)In this chapter, this is the same definition as the conditional skewness footnote 5 although conditional measure is more general to conditional on other information, not only the market return.

\(^{10}\)Specifically, their calculation is based on Pearson skewness idea with return mean is set to be zero, \(\text{skew} = \frac{\sqrt{\sum_{i=1}^{N} r_{i,t}^3}}{\Sigma_{i=1}^{N} r_{i,t}^2}\), where \(N\) is the number of observation intra-day, \(r_{i,t}\) is the intra-day return for asset \(i\) at time \(t\) as they are use high frequency data. In more general case, data frequency can be adjusted.
Unlike previous research, real ex ante skewness is introduced and obtained in two different ways, Bali and Murray (2013) and Conrad et al. (2013), both of them find the negative relation between the Risk-Neutral Skewness and the subsequent asset return. However, this negative relation has been argued recently as some studies document opposite (positive) relationship (Stilger et al., 2016 and Gkionis et al., 2017). One argument comes to this difference may due to the fact that the Risk-Neutral Skewness is tagged by picking up short-term arbitrage while moving average manipulation from previous researches changes this pricing mechanism. In general, all these empirical results are more centred on stock market instead of on future market, which point out the gap in the literature which will be filled in this thesis.

Few studies are focusing on idiosyncratic factors (unexplained parts from traditional factors, term structure, momentum and hedging pressure) on global commodity futures return, Fuertes et al. (2015), among others, extracts residuals from regressing assets’ return on momentum and term structure and estimate the second moment called idiosyncratic volatility. Based on cross-sectional sorting approach, the triple-sorted portfolio (sorted by momentum, term structure and idiosyncratic volatility) does offer a more smoothing return and lower drawbacks. Cross-sectional regression with dummy included also confirms its pricing ability as coefficient is relatively larger than term structure and moment across all subsamples. Skewness, calculated no matter in the third central moment, in idiosyncratic or in expected way, shows non-trivial effect on asset return generating process (Fernandez-Perez et al., 2018). Their empirical results demonstrate that after controlling the traditional risk factors, extra 8% average annual return can be obtained from buying low and selling high skewed commodity futures assets. In conclusion, a strong negative relation is documented in their studies, which is consistent with the relevant studies applied in the equity market.

The closest research (in terms of only risk-neutral estimation idea) to this chapter is conducted

---

11 By saying high (low) skewness for a distribution in this thesis, this implies more positive (negative) skewness value.
by Triantafyllou et al. (2015) who use Bakshi et al. (2003) method to explore variance risk premium (differential of realised volatility and risk-neutral volatility). Therefore, in conclusion, studies on the Risk-Neutral Skewness pricing ability test for global futures market is new to the literature.

2.3 Theoretical Background

2.3.1 Out-Performance of Risk-Neutral Skewness

Compared with the realised skewness estimator that requires a set of historical data, the Risk-Neutral Skewness estimator studied in this chapter has several good features. Analogous to the implied volatility, the Risk-Neutral Skewness is retrieved from options data under risk-neutral probability measure assumption. Within this probability measure, all investors are assumed to hold the same preference on outcomes and therefore expected pay-off can be measured. Risk-neutral moments have been argued with valuable information embedded in, which reflects the market participates’ expectation about future assets’ characters movement (Bates, 1991, Jackwerth and Rubinstein, 1996 and Bakshi et al., 1997). This forward-looking property is more informative and matched with modern finance theory as investment decisions are based on maximisation of the future expected value, while realised moments reflect only the past information. Current popular studies on heterogeneous belief and market sentiment confirm that the Risk-Neutral Skewness tells more than its counterpart (Han, 2008).

The other advantage of risk-neutral estimation moments is mainly due to its unbiased property of representing true moments (Neuberger, 2012). Under the aggregation property theory, he shows that low frequency moments estimator can be calculated in an unbiased manner by using high frequency data (e.g. option data). Estimator based on the past sample history length selection can be easily plagued by outliers (Kim and White, 2004). According to the simulation comparison of different
moments estimation methods, they suggest that estimations, beyond the current standard approach, need more explorations to give more accurate and deep insight outcome. For data sample selected in this chapter, using realised moments cause bias as recent financial crisis period is included. In another way, short sample period variation can dominate the measure, while long-time sample estimation causes the loss of more historical information and freedom (Hansis et al., 2010).

In addition to the extra step of determining estimation window length, data frequency usage for realised moments calculation in terms of return series is also not reliable especially when return is in non i.i.d case (Neuberger, 2012). Bootstrap linear regression sampling shows that skewness calculated in daily (monthly) return is not proportional to the outcome in monthly (yearly). Following (Bakshi et al., 2003), they state that as long as the underlying process satisfying the martingale assumption, a model-free method with their aggregation property can provide an unbiased approximation no matter in realised measure or risk-neutral implied one.

### 2.3.2 Pricing Mechanism of Risk-Neutral Skewness

Before walking through the mechanism behind risk-neutral measure, the realised measure is first reviewed for later comparison. Those positively skewed assets preferred investors (known as lottery like behaviour) will push up these assets’ prices as they are willing to pay more for them (Mitton and Vorkink, 2007). As a result, over-pricing for these positively skewed assets will in return generate less profits and underpriced negative skewed assets offers higher returns. Although this over-pricing phenomenon is argued to be persistent due to short-selling restrictions, this does not fall into this chapter case as selling is allowed in commodity future markets.

Regardless of the appearance of short-selling restrictions, cumulative prospect theory proposed by (Tversky and Kahneman (1992) and Barberis and Huang (2008)) explain that positively skewed
assets are over weighted by investors holding special value function (convex in loss and concave in gain). By assigning more expected values on future potential extreme positive returns, investors are willing to pay more for those positive skewed assets, resulting in these assets’ over-pricing. Therefore, buying (selling) assets with positive (negative) skewness will generate lower (higher) return in the subsequent period once this arbitrage is corrected. Recent corresponding empirical study is conducted by (Fernandez-Perez et al., 2018) who confirm the negative relation between the realised skewness and subsequent futures asset return. However, it is worth mentioning that the market theory implied from (Deaton and Laroque, 1992) shows that positive relation is also reasonable to be expected, which is in the line with this thesis findings.

Different from the realised measure pricing mechanism, frameworks for the risk-neutral one does not directly reflect above statements as option trading possibilities have not been taken into account in those studies. Motivated by the demand-based option pricing theory from Garleanu et al. (2009) and empirical findings from Bollen and Whaley (2004), the first pricing framework is the net buying pressure idea borrowed from stock market. In the stock market, short-selling constraints leads to the impossibility of fully hedged position. In commodity futures market, higher margin requirement for extra short position, liquidity constraints and inventory level maintaining cost are some important aspects to be concerned. Investors with negative expectation about future return will buy more put OTM options, driving the Risk-Neutral Skewness to be more negative value. Negative skewness implies the over-pricing underlying and causes less return once miss-pricing is corrected.

For this argument to hold, heterogeneous belief about the underlying need to exist (Han, 2008). More recently, Friesen et al. (2012) states that the Risk-Neutral Skewness is strongly negatively related to several market sentiment proxies (e.g. idiosyncratic volatility). Following the empirical pricing test in the commodity futures market by Fuertes et al. (2015), idiosyncratic volatility has
been documented with significant pricing ability and negatively related to future return. To sum up, originating from the over-pricing perception of the underlying assets, demanding pressure effect with more OTM put options purchasing leads to the negative Risk-Neutral Skewness. When the arbitrage correction occurs, positive relation is observed between the current the Risk-Neutral Skewness and subsequent asset return.

Another framework explaining pricing intuition can be attributed to the selective hedging idea by (Stulz, 1996). It emphasizes that market participators tend to use selective hedging strategy rather than "full-cover" hedging given the consideration of future price change. In line with their arguments, trading in OTM put option can be regarded as a protection on the scenario of unexpected negative tail outcomes (distinct from the traditional mean-variance optimization frame).

In the commodity futures market, if there is a perception about the potential decrease (increase) of futures price due to the current assets overprice (under-pricing), continue trading in futures market is not attractive as hedgers may be exposed to a higher cumulative risk level and need to pay extra premium to their counterparts (speculators). Under the framework of selective hedging, buying OTM put option tends to be more satisfactory and in the end, leads to more negative Risk-Neutral Skewness. As a result, those over-pricing (under-pricing) assets accumulate a higher value of negative (positive) Risk-Neutral Skewness, which automatically yield lower (higher) return in the period when arbitrage correction happens.

However, this type of mispricing will disappear in a short time instead of being persistent. Compared with the equity market, short-selling allowance in commodity futures is fully flexible as shorting is equally treated, therefore, pricing correction process will be faster and shorter (Stilger et al., 2016 and Gkionis et al., 2017), which is consistent with the second chapter empirical evidence.
CHAPTER 2. RISK-NEUTRAL SKEWNESS ON COMMODITY PRICING

2.4 Data Description

Daily settlement futures price, trading volume and open interest data from 10/10/2007 to 01/03/2016 are collected from DataStream: agriculture sector (cocoa, coffee C, corn, cotton NO.2, frozen concentrated orange juice, oats, rough rice, soy bean meal, soy bean oil, soy beans, sugar NO.11, wheat), energy sector (WTI crude oil, RBOB gasoline, heating oil NO.2, light sweet crude oil, natural gas), livestock sector (feeder cattle, lean hogs, live cattle) and metal sector (gold, silver). Palladium and platinum are excluded for the consideration of estimation bias due to limited amount of OTM option data available. One-period return is calculated in the log price difference. Consistent with the literature, the nearest-to-maturity contract is used and rolled to the second nearest-to-maturity contract one month before the nearest-to-maturity contract expiring. Option data are obtained from DataStream with daily strike price, traded volume, contract market price (both call and put options) for each specific product. Hedgers’ future only aggregated long and short open interest data are downloaded from the Commodity Futures Trading Commission (CFTC) website in weekly frequency. In addition to commodity specific data, equity market related data (the Fama_French five factors) are downloaded from Fama_French data library website.

2.5 Methodology

Recently, model-free moments estimation is popular and widely cited method in the literature is by Bakshi et al. (2003) (denoted as BKM method in following content). Based on log return calculation and central moment idea, they show that asset’s moments can be approximated via using daily discrete

---

12 CFTC requires future trading participators to identify their types (hedgers, speculators, not reportable).
13 Data are regularly collected every Tuesday and released on the following Friday.
15 Estimation formula is listed in appendix A.
price data from option contracts. This idea has been argued more often recently as the central moment risk is not easy to hedge in reality somehow. To account for the jump risk, discrete error and downside risk bias from the BKM method, model-free estimator from Kozhan et al. (2013) is introduced and applied in this chapter analysis as the main estimation measurement of Risk-Neutral Skewness. One consideration for Kozhan et al. (2013) method to fit the commodity futures is to control the jumps possibilities caused by futures contract return roll-over. After estimating the skewness, time-series and cross-sectional ordinary least square (OLS) regression analysis are used to explore both time exposure and risk premium variation.

2.5.1 Risk-Neutral Skewness

Generally, Bakshi et al. (2003) show that volatility, skewness and kurtosis can be mimicked via a quadratic, cubic and quartic pay-off structure. The input to this structure uses daily observations cross over options data with different strike prices for the same underlying. However, it is worth mentioning that moments estimated under the framework of BKM are weighted with squared or cubed strike price of the underlying. This weighting scheme introduces potential estimation bias especially during the illiquid period in which call option part will be deteriorated and put option part will be overstated. Put option price increases rapidly when market exception falls in downside way, resulting in more negative value in estimation (Kozhan et al., 2013 and Leontsinis and Alexander, 2017). Following their arguments, within the framework of aggregation property theory (Neuberger, 2012), risk-neutral third moment estimation formula used in this chapter is following (Kozhan et al., 2013) and denoted

16Rationale for using OLS regression method is evidenced either by empirical study (Fama and MacBeth, 1973), Bakshi et al. (2013), Basu and Miffre (2013), Fuertes et al. (2015), Daskalaki et al. (2014), and Fernandez-Perez et al. (2018), by solving estimation bias by (Newey and West, 1986) and Hansen (1982), and by simulation comparisons with GMM, GLS by Shanken and Zhou (2007).
CHAPTER 2. RISK-NEUTRAL SKEWNESS ON COMMODITY PRICING

as RNSK hence in the following content specifically\footnote{For comparison, BKM method suggested factor property is also reported in the following tables and figures, with name referred to RNSK(B)}

\[ v_{t,T}^E = 2 \sum_{K_i \leq F_t,T} \frac{P_{t,T}(K_i)}{B_{t,T}K_i^2} \Delta I(K_i) + 2 \sum_{K_i > F_t,T} \frac{C_{t,T}(K_i)}{B_{t,T}K_i^2} \Delta I(K_i) \] (2.5.1)

\[ v_{t,T}^L = 2 \sum_{K_i \leq F_t,T} \frac{P_{t,T}(K_i)}{B_{t,T}K_iF_t,T} \Delta I(K_i) + 2 \sum_{K_i > F_t,T} \frac{C_{t,T}(K_i)}{B_{t,T}K_iF_t,T} \Delta I(K_i) \] (2.5.2)

where, \( B_{t,T} \) is the bond present value at time \( t \) with time-to-maturity is \( (T - t) \) given value at expiration date is unit, \( P_{t,T} \) and \( C_{t,T} \) are put and call option market price at time \( t \) with time-to-maturity is \( (T - t) \), \( K_i \) is the strike price level for underlying at difference value index \( i \).

\[ \Delta I(K_i) = \begin{cases} \frac{K_{i+1} - K_{i-1}}{2}, & \text{for } 0 \leq i \leq N \text{(with } K_{-1} = 2K_0 - K_1, K_{N+1} = 2K_N - K_{N-1}) \\ 0, & \text{Otherwise} \end{cases} \]

\[ RNSK_{t,T} \equiv \frac{3v_{t,T}^E - v_{t,T}^L}{\left(v_{t,T}^L\right)^{\frac{3}{2}}} \] (2.5.3)

where, \( RNSK_{t,T} \) is the Risk-Neutral Skewness at time \( t \) with the expiration time \( T \).

As for the specific \( RNSK \) calculation steps, the first filtration step deletes all in-the-money call (strike price lower than market price) and put (strike price lower than market price) options contracts and leave only out-of-money options. In order to make estimated results precisely, the minimum number of call and put OTM option price data required for calculation is at least 4 respectively. Meanwhile, the number of call and put options should be equal in order to estimate. Since the practical analysis is to deal with discrete data, trapezoidal approximation (\cite{Dennis and Mayhew 2002} and \cite{Conrad et al. 2013}) is implemented to calculate of the discrete integral equations (2.5.1)-(2.5.2). Finally, those options have only one week left to maturity will also be excluded as the trading behaviour.
on these options will distort the fair value of option value themselves to some extent.

After the filtration of data, discretion errors (e.g. the simple Riemann sum problem argued by Leontsinis and Alexander (2017)) and truncation errors are handled in the following ways. Given the option market quotes, a fine interval is constructed via the natural cubic spline interpolation within the strike price interval (Jiang and Tian, 2005, Jiang and Tian, 2007 and Carr and Wu, 2008). For data points beyond the current truncated extreme strike price range, a linear-interpolation method with the closest data is employed to account for the implied volatility smile effect or skew effect\(^{18}\) (Jiang and Tian, 2007). The interval is scaled by 2 standard deviation of underlying spot price to make sure the minimum effect from the truncation error (Jiang and Tian, 2005). The Bisection method\(^{19}\) is then applied to calculate the implied volatility via the Black model by Black and Scholes (1973). All fitted implied volatilities are then converted back to the call and put market price via the Black model\(^{20}\).

After the estimation procedure, following literature, 30 days constant maturity series Risk-Neutral Skewness for each underlying product is computed. In this computation process, if an exact time-to-maturity equal to 30\(^{21}\) does exist, the corresponding value is used directly, otherwise, the hermite cubic spline is employed to calculate this constant maturity value, which accounts for calendar arbitrage issue. It also shows non-linear trend property for long maturity data fitting as well as provide shape preserving merit (Leontsinis and Alexander, 2017). For robustness, linearly interpolation is also applied to estimate the constant time-to-maturity RNSK\(^{22}\).

---

\(^{18}\)Here also consider flat extrapolation (extreme value on two sides will be used for points outside strike price range without linear fitting) and no extrapolation (only consider data interpolation within strike price range), results are similar not reported here, can be requested from author  
\(^{19}\)Results are also robust to Newton-Raphson method.  
\(^{20}\)Black model here is simply treated as a bridge on pricing, which does not affect final results  
\(^{21}\)The reason why focusing on 30-day constant maturity is due to skewness estimation purpose limitation. By increasing maturity value to 60 and 120 days, the number of interpolated constant risk neutral skewness shows decreasing effect. Therefore, to avoid data inconsistency and bias in the further regression analysis, only 30-day maturity is considered  
\(^{22}\)Empirical results are similar to hermite cubic spline, so not reported, it can be requested from author
CHAPTER 2. RISK-NEUTRAL SKEWNESS ON COMMODITY PRICING

2.5.2 Commodity Market Variables

According to the commodity market characteristics, traditional risk factors are borrowed from the literature and calculated in the following ways.

The term structure factor portfolio uses the basis as a sorting signal. Basis here is defined as the log differential between the nearest-to-maturity contract price, $F_{T_1}^{T_i}$, and the second nearest-to-maturity contract price, $F_{T_2}^{T_i}$, of a commodity futures contract $i$ at time $t$ with $T_2 > T_1$. Following Koijen et al. (2013), in order to make the signal more informative to sort on a cross-sectional level, the scaled basis measure is employed

$$Basis_{i,t} = \log\left(\frac{F_{T_1}^{T_i}}{F_{T_2}^{T_i}}\right)$$ (2.5.4)

where $F_{T_1}^{T_i}$ acts as a proxy for the spot price, and $T_2 - T_1$ is the maturity differential in days. A positive basis signals a backwardated (contangoed) market and as such predicts that commodity futures prices will subsequently rise (fall).

The hedging pressure factor portfolio is based on participators’ open interest that signals the direction of trade of commodity trading participators (Fernandez-Perez et al., 2018). The hedging pressure for $i^{th}$ commodity futures contract at time $t$ is measured by the ratio of speculators’ long positions only to total position (also documented as large non-commercial traders in CFTC). The general formula is formatted as follows:

$$HP_{i,t} = \frac{\# long speculation positions_{i,t}}{total \# speculation positions_{i,t}}$$ (2.5.5)

where $HP_{i,t}$ is represented by large non-commercial traders (speculators) hedging pressure for partic-

\[23\] Long only hedger’s position (large commercial traders) is also computed and tested, showing similar results to speculators’ results, therefore not reported in this chapter.
The momentum is a portfolio sorted via signals that are the average commodity futures return over a past window. As in Asness et al. (2013), Szymanowska et al. (2014) and Miffre and Rallis (2007), a 12-month (52 weeks) window period is employed. Formally,

\[
MOM_{i,t} = \frac{\sum_{j=t-52}^{j=t-1} r_{i,j}}{52}
\]

where \( r_{i,t} = \ln F_{i,t}^{T_i} - \ln F_{i,t-1}^{T_i} \) is the log return of the nearest-to-maturity commodity futures contract \( i \) on week \( t \).

The realised skewness portfolios is sorted on signals that is the "Pearson’s moment coefficient of skewness of each commodity at month end \( t \) using the daily return history in the preceding 12-month window” proposed in Fernandez-Perez et al. (2018). The same daily return data over the past 12-month to estimate Pearson skewness coefficient is used. After daily measures are obtained, weekly results are selected for implementation. The only difference compared with their works is portfolio rebalancing frequency (weekly in this thesis, monthly in their study).

\[
SK_{i,t} = \left[ \frac{1}{D} \sum_{d=1}^{D} \left( r_{i,d,t} - \hat{\mu}_{i,t} \right) \right]^{3} \left( \hat{\sigma}_{i,t} \right)^{-3}
\]

where, \( r_{i,d,t} \) is the daily return for \( i^{th} \) commodity asset with \( D \) (the total number of observation) spanning from 1 to 252. \( \hat{\mu}_{i,t} \) is the standard mean estimation and \( \hat{\sigma}_{i,t} \) is standard error with scaling factor \( \sqrt{1/(D-1)} \).

---

\(^{24}\)Net position on hedging pressure measurement is awarded of in recent literature, see Szymanowska et al. (2014), De Roon et al. (2000), Basu and Miffre (2013) and Bessembinder (1992), robustness check on net position for both hedgers and speculators are conducted with similar results to long only one, no reported.

\(^{25}\)Another expected skewness, \( \text{Skew} = \frac{P_{99}-2P_{50}}{P_{99}-P_{1}} \), is also computed in a rolling manner with past one year daily sample distribution three quantiles (99%, 50% and 1%) and tested in the following, showing the same factor-return relation to Pearson skewness, pointing out that the option based skewness is distinct mainly due to risk-neutral property and no historical data inclusion Green and Hwang (2012).
2.5.3 Risk Factor Portfolio Construction

The weekly time-series of long-short portfolios excess returns (represented by TS, HP, MOM, SK and RNSK in this following content) are obtained by buying and selling quantile group assets simultaneously. At each time point, commodity assets are cross-sectional sorted via signals into quantile groups and corresponding groups equally weight returns are calculated. At the same time point, portfolio return is measured by the high quantile group mean minus the low quantile group mean. This procedure is then carried out when one new weekly observation become available and so forth.

Specific ranking period is identified for the sake of factor value calculation. The most recent 12-month window is used for MOM (52 weeks) and SK (252 days) signals according to Fernandez-Perez et al. (2017a), while TS, HP and RNSK employ the last week observation for comparison convenience as RNSK is argued to be less persistent (mispricing will be corrected in a short time, long time averaging weakens the signal effect) pricing factor in the literature (Stilger et al., 2016 and Gkionis et al., 2017). Specifically, denoting L and S the commodities included in the long and short portfolio, respectively. HP, TS, MOM and RNSK factors are constructed as high(L)-minus-low(S) portfolio, while only SK is constructed as low(L)-minus-high(S) portfolio. This follows from the wisdom that a high value of hedgers’ hedging pressure, term structure, momentum and Risk-Neutral Skewness predicts an increase in subsequent commodity futures prices whereas a high value of the realised skewness predicts instead a decrease in subsequent commodity futures prices (Bakshi et al., 2013, Bessembinder, 1992, Basu and Miffre, 2013, Miffre and Rallis, 2007, Stilger et al., 2016, Amaya et al., 2011 and Fernandez-Perez et al., 2018). Hereafter, the notation $HML$ denotes the corresponding long-short portfolio.
2.6 Empirical Results

2.6.1 Summary Statistics for Return and Commodity Risk Factors

Summary statistics of rolling continuous futures’ return is reported in the table 2.1. Mean is annualized based on weekly frequency, showing averagely negative performance in the sample period. Standard deviations are recorded with higher value, which is reasonably to be expected as recent global financial crisis period is included. Assets’ return distribution normality is rejected, which can be identified from high moments value column (skewness and kurtosis) as well as Jarque.Bera test from Jarque and Bera (1987).

Fully-collateralized long-short portfolio approach is used to construct time-series risk factors. For the sake of portfolio diversification idea, only two quartile groups are considered (bottom 25% and top 25%) due to non-negligible amount of missing values generated during RNSK estimation procedure. Commodity assets at the end of each week are grouped according to weekly risk factor value and held until the end of next week when new factor observations become available. Then, portfolio is rebalanced weekly and continues until the end of data sample. For the sake of practical strategy investing comparison, most related risk factors (their portfolios’ performance) proposed in the literature are reported in table 2.2. For further statistical analysis, time-series correlation matrix among risk factors is reported in table 2.3 with corresponding significance highlighted in bold value.

From table 2.2, results for traditional factors like TS, MOM, HP.C (measured as percentage of commercial traders’ short only positions) and HP are consistent with studies and findings in the most literature. The key interest point is about to what extent the Risk-Neutral Skewness performs different from the realised skewness and moreover, other well-established factors in commodity futures market. In general, portfolio performance suggested from RNSK factor is superior to all other factors

---

26 At least four assets in one quartile group is required for better portfolio risk diversification
Table 2.1: Summary Statistics for Commodity Futures Return

<table>
<thead>
<tr>
<th>Panel A: Agriculture Sector</th>
<th>N</th>
<th>Mean</th>
<th>SD</th>
<th>Median</th>
<th>Min</th>
<th>Max</th>
<th>Skew</th>
<th>Kurtosis</th>
<th>Jarque.Bera</th>
<th>AR(1)</th>
<th>T test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cocoa</td>
<td>433</td>
<td>0.057</td>
<td>0.289</td>
<td>0.000</td>
<td>-0.167</td>
<td>0.143</td>
<td>0.051</td>
<td>1.295</td>
<td>0.000</td>
<td>0.014</td>
<td>0.571</td>
</tr>
<tr>
<td>Coffee</td>
<td>433</td>
<td>-0.075</td>
<td>0.329</td>
<td>-0.002</td>
<td>-0.145</td>
<td>0.177</td>
<td>0.133</td>
<td>0.854</td>
<td>0.001</td>
<td>0.013</td>
<td>-0.662</td>
</tr>
<tr>
<td>Corn</td>
<td>433</td>
<td>-0.047</td>
<td>0.319</td>
<td>0.001</td>
<td>-0.165</td>
<td>0.184</td>
<td>0.006</td>
<td>1.811</td>
<td>0.000</td>
<td>-0.322</td>
<td>-0.421</td>
</tr>
<tr>
<td>Cotton</td>
<td>433</td>
<td>-0.043</td>
<td>0.320</td>
<td>-0.001</td>
<td>-0.168</td>
<td>0.162</td>
<td>-0.028</td>
<td>1.320</td>
<td>0.000</td>
<td>0.025</td>
<td>-0.385</td>
</tr>
<tr>
<td>Live Cattle</td>
<td>433</td>
<td>-0.015</td>
<td>0.151</td>
<td>0.001</td>
<td>-0.089</td>
<td>0.071</td>
<td>-0.364</td>
<td>1.154</td>
<td>0.000</td>
<td>-0.118</td>
<td>-0.278</td>
</tr>
<tr>
<td>Oat</td>
<td>433</td>
<td>-0.031</td>
<td>0.376</td>
<td>0.000</td>
<td>-0.233</td>
<td>0.334</td>
<td>0.300</td>
<td>4.959</td>
<td>0.000</td>
<td>-0.062</td>
<td>-0.240</td>
</tr>
<tr>
<td>Orange</td>
<td>433</td>
<td>-0.009</td>
<td>0.369</td>
<td>0.000</td>
<td>-0.247</td>
<td>0.184</td>
<td>0.069</td>
<td>2.279</td>
<td>0.000</td>
<td>-0.025</td>
<td>-0.067</td>
</tr>
<tr>
<td>Rough Rice</td>
<td>433</td>
<td>-0.087</td>
<td>0.243</td>
<td>0.000</td>
<td>-0.142</td>
<td>0.099</td>
<td>-0.195</td>
<td>0.990</td>
<td>0.000</td>
<td>0.070</td>
<td>-1.036</td>
</tr>
<tr>
<td>Soybean Meal</td>
<td>433</td>
<td>0.142</td>
<td>0.301</td>
<td>0.003</td>
<td>-0.152</td>
<td>0.132</td>
<td>-0.103</td>
<td>0.390</td>
<td>0.155</td>
<td>-0.068</td>
<td>1.358</td>
</tr>
<tr>
<td>Soybean</td>
<td>433</td>
<td>0.068</td>
<td>0.264</td>
<td>0.003</td>
<td>-0.127</td>
<td>0.113</td>
<td>-0.183</td>
<td>0.631</td>
<td>0.007</td>
<td>-0.035</td>
<td>0.746</td>
</tr>
<tr>
<td>Soybean Oil</td>
<td>433</td>
<td>-0.081</td>
<td>0.264</td>
<td>0.000</td>
<td>-0.116</td>
<td>0.140</td>
<td>0.051</td>
<td>0.894</td>
<td>0.000</td>
<td>0.002</td>
<td>-0.891</td>
</tr>
<tr>
<td>Sugar</td>
<td>433</td>
<td>-0.046</td>
<td>0.356</td>
<td>-0.001</td>
<td>-0.230</td>
<td>0.145</td>
<td>-0.266</td>
<td>1.248</td>
<td>0.000</td>
<td>-0.075</td>
<td>-0.373</td>
</tr>
<tr>
<td>Wheat</td>
<td>433</td>
<td>-0.182</td>
<td>0.329</td>
<td>-0.001</td>
<td>-0.176</td>
<td>0.147</td>
<td>0.050</td>
<td>0.974</td>
<td>0.000</td>
<td>-0.022</td>
<td>-1.599</td>
</tr>
<tr>
<td>Lean Hogs</td>
<td>433</td>
<td>-0.058</td>
<td>0.240</td>
<td>0.002</td>
<td>-0.122</td>
<td>0.105</td>
<td>-0.324</td>
<td>0.464</td>
<td>0.003</td>
<td>0.057</td>
<td>-0.693</td>
</tr>
<tr>
<td>Feeder Cattle</td>
<td>433</td>
<td>-0.023</td>
<td>0.157</td>
<td>0.002</td>
<td>-0.120</td>
<td>0.080</td>
<td>-0.577</td>
<td>2.416</td>
<td>0.000</td>
<td>0.013</td>
<td>-0.423</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel B: Energy Sector</th>
<th>N</th>
<th>Mean</th>
<th>SD</th>
<th>Median</th>
<th>Min</th>
<th>Max</th>
<th>Skew</th>
<th>Kurtosis</th>
<th>Jarque.Bera</th>
<th>AR(1)</th>
<th>T test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brent</td>
<td>433</td>
<td>-0.168</td>
<td>0.348</td>
<td>0.000</td>
<td>-0.176</td>
<td>0.216</td>
<td>-0.196</td>
<td>1.894</td>
<td>0.000</td>
<td>0.026</td>
<td>-1.388</td>
</tr>
<tr>
<td>Heating Oil</td>
<td>433</td>
<td>-0.143</td>
<td>0.319</td>
<td>-0.002</td>
<td>-0.144</td>
<td>0.219</td>
<td>0.183</td>
<td>1.992</td>
<td>0.000</td>
<td>0.061</td>
<td>-1.292</td>
</tr>
<tr>
<td>Light Crude</td>
<td>433</td>
<td>-0.233</td>
<td>0.369</td>
<td>-0.001</td>
<td>-0.181</td>
<td>0.219</td>
<td>-0.174</td>
<td>1.630</td>
<td>0.000</td>
<td>-0.038</td>
<td>-1.824</td>
</tr>
<tr>
<td>WTI</td>
<td>433</td>
<td>-0.233</td>
<td>0.369</td>
<td>-0.001</td>
<td>-0.181</td>
<td>0.217</td>
<td>-0.193</td>
<td>1.618</td>
<td>0.000</td>
<td>-0.037</td>
<td>-1.826</td>
</tr>
<tr>
<td>Natural Gas</td>
<td>433</td>
<td>-0.450</td>
<td>0.409</td>
<td>-0.004</td>
<td>-0.162</td>
<td>0.208</td>
<td>0.117</td>
<td>0.520</td>
<td>0.045</td>
<td>-0.020</td>
<td>-3.171</td>
</tr>
<tr>
<td>RBOB</td>
<td>433</td>
<td>-0.075</td>
<td>0.356</td>
<td>0.000</td>
<td>-0.183</td>
<td>0.243</td>
<td>-0.224</td>
<td>2.349</td>
<td>0.000</td>
<td>0.003</td>
<td>-0.608</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel C: Metal Sector</th>
<th>N</th>
<th>Mean</th>
<th>SD</th>
<th>Median</th>
<th>Min</th>
<th>Max</th>
<th>Skew</th>
<th>Kurtosis</th>
<th>Jarque.Bera</th>
<th>AR(1)</th>
<th>T test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gold</td>
<td>433</td>
<td>0.055</td>
<td>0.196</td>
<td>0.002</td>
<td>-0.134</td>
<td>0.131</td>
<td>-0.085</td>
<td>3.235</td>
<td>0.000</td>
<td>0.002</td>
<td>0.819</td>
</tr>
<tr>
<td>Silver</td>
<td>433</td>
<td>0.008</td>
<td>0.349</td>
<td>-0.001</td>
<td>-0.241</td>
<td>0.226</td>
<td>-0.198</td>
<td>2.550</td>
<td>0.000</td>
<td>-0.005</td>
<td>0.063</td>
</tr>
</tbody>
</table>

Notes: the table reports summary statistics of weekly commodity futures returns from 10/10/2007 to 01/03/2016. First row of this table is for descriptive statistics and first column is for specific assets. Results are organized by sectors based on commodities’ attributes: Panel A: Agriculture sector, Panel B: Energy sector and Panel C: Metal sector. From the second column, standard first four central moments are reported, labelled Mean, SD, Skew and Kurtosis; median, minimum and maximum value of return series: Median, Min and Max; Jarque.Bera test results for return distribution normality test are shown in the 9th column; autocorrelation with one week lag coefficient results are in the 10th column with name AR(1); the last column is unconditional asset return mean zero T test statistics.

Based portfolios. In terms of the portfolio return realization direction prediction, RNSK based portfolio from either Bakshi et al. (2003) or Kozhan et al. (2013) is over 50%, while others fall into the group of less than 50%. Meanwhile, Sharpe ratios for these two portfolios are also over 1, implying strong risk adjusted compensation given one unit of risk bearing. Considering the threshold or target portfolio return, omega and sortino ratio listed in table also point out the superior performance on two RNSKs. For portfolio performance via Pearson skewness coefficient, stated by Fernandez-Perez et al. (2018), annual return is only 2.5%. Compared with this chapter interest, this number increase
Table 2.2: Portfolio Performance Statistics for Commodity Risk Factors

<table>
<thead>
<tr>
<th></th>
<th>EW</th>
<th>MOM</th>
<th>TS</th>
<th>HP.C</th>
<th>HP</th>
<th>RNSK(B)</th>
<th>RNSK</th>
<th>SK</th>
<th>LIQUID</th>
<th>CV</th>
<th>IDIOSK</th>
<th>∆OP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>-0.073</td>
<td>0.024</td>
<td>0.062</td>
<td>0.052</td>
<td>0.033</td>
<td>0.115</td>
<td>0.133</td>
<td>0.018</td>
<td>-0.007</td>
<td>-0.028</td>
<td>-0.004</td>
<td>0.002</td>
</tr>
<tr>
<td>StdDev</td>
<td>0.173</td>
<td>0.126</td>
<td>0.117</td>
<td>0.095</td>
<td>0.087</td>
<td>0.095</td>
<td>0.095</td>
<td>0.113</td>
<td>0.093</td>
<td>0.085</td>
<td>0.097</td>
<td>0.102</td>
</tr>
<tr>
<td>Sharp</td>
<td>-0.420</td>
<td>0.192</td>
<td>0.526</td>
<td>0.549</td>
<td>0.374</td>
<td>1.220</td>
<td>1.393</td>
<td>0.165</td>
<td>-0.075</td>
<td>-0.332</td>
<td>-0.041</td>
<td>0.015</td>
</tr>
<tr>
<td>SortinoRatio (0%)</td>
<td>-0.078</td>
<td>0.037</td>
<td>0.107</td>
<td>0.112</td>
<td>0.077</td>
<td>0.286</td>
<td>0.346</td>
<td>0.033</td>
<td>-0.015</td>
<td>-0.062</td>
<td>-0.008</td>
<td>0.003</td>
</tr>
<tr>
<td>OmegaSharpeRatio (0%)</td>
<td>-0.146</td>
<td>0.081</td>
<td>0.231</td>
<td>0.243</td>
<td>0.159</td>
<td>0.582</td>
<td>0.691</td>
<td>0.064</td>
<td>-0.028</td>
<td>-0.114</td>
<td>-0.015</td>
<td>0.011</td>
</tr>
<tr>
<td>Skewness</td>
<td>-0.145</td>
<td>-0.570</td>
<td>-0.183</td>
<td>-0.092</td>
<td>0.102</td>
<td>0.206</td>
<td>0.404</td>
<td>0.192</td>
<td>-0.004</td>
<td>-0.077</td>
<td>-0.005</td>
<td>-0.908</td>
</tr>
<tr>
<td>99% VAR(Cornish-Fisher)</td>
<td>-0.087</td>
<td>-0.060</td>
<td>-0.052</td>
<td>-0.046</td>
<td>-0.046</td>
<td>-0.036</td>
<td>-0.028</td>
<td>-0.061</td>
<td>-0.049</td>
<td>-0.035</td>
<td>-0.039</td>
<td>-0.093</td>
</tr>
<tr>
<td>MaxDrawdown</td>
<td>0.633</td>
<td>0.164</td>
<td>0.213</td>
<td>0.187</td>
<td>0.137</td>
<td>0.096</td>
<td>0.083</td>
<td>0.388</td>
<td>0.235</td>
<td>0.416</td>
<td>0.311</td>
<td>0.316</td>
</tr>
<tr>
<td>% of positive months</td>
<td>0.471</td>
<td>0.443</td>
<td>0.476</td>
<td>0.457</td>
<td>0.446</td>
<td>0.520</td>
<td>0.515</td>
<td>0.491</td>
<td>0.457</td>
<td>0.464</td>
<td>0.448</td>
<td>0.182</td>
</tr>
<tr>
<td>ADF.Test</td>
<td>0.010</td>
<td>0.010</td>
<td>0.010</td>
<td>0.010</td>
<td>0.010</td>
<td>0.010</td>
<td>0.010</td>
<td>0.010</td>
<td>0.010</td>
<td>0.010</td>
<td>0.010</td>
<td>0.010</td>
</tr>
</tbody>
</table>

Notes: most commodity related risk factors based long-short portfolios’ performance are reported in this table. The first column reports all portfolios related statistics and time-series stationary test. The first row with bold label stands for long-short portfolios based risk factors from the third column to the right end: momentum (MOM), term structure (TS), hedges’ short open interest over total open interest (HP.C), speculators’ long open interest over total open interest (HP), BKM Risk-Neutral Skewness (RNSK(B)), Kozhan Risk-Neutral Skewness (RNSK), Pearson skewness coefficient over past 12 months (SK), dollar volume over absolute return in past 2 months (LIQUID), variance-over-mean over past 36 months (CV), skewness of the residuals in time-series regressions of weekly commodity futures returns on weekly observations for the EW, TS, MOM and HP factors (IDIOSK), change of entire term structure open interest (∆OP). EW, on the second column, is an exception with equally weighted all available assets long only portfolio.

remarkable to 11.5% and 13.3% on RNSK(B) and RNSK sorted portfolios separately. Regarding the risk management idea, trading via RNSK is less risky in terms of some common measures: maximum drawdown (maxDrawdown) and Value-at-Risk (VaR). Overall, compared with all other factors, RNSK factors present better return and less risk.

### 2.6.2 Trading Strategy Performance

The practical idea of how these factors contribute can be referred to their portfolio cumulative returns. This can be treated as a measurement on how stable this trading signal implies. The long-short portfolios based on risk factor signal are plotted in figure 2.1,
Table 2.3: Pair-Wise Correlation Matrix of Risk Factors

<table>
<thead>
<tr>
<th></th>
<th>EW</th>
<th>MOM</th>
<th>TS</th>
<th>HP.C</th>
<th>HP</th>
<th>SMB</th>
<th>HML</th>
<th>RMW</th>
<th>CMA</th>
<th>RNSK(B)</th>
<th>RNSK</th>
<th>SK</th>
<th>LIQUID</th>
<th>CV</th>
<th>IDIOSK</th>
<th>∆OP</th>
</tr>
</thead>
<tbody>
<tr>
<td>EW</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MOM</td>
<td>-0.263</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TS</td>
<td>-0.107</td>
<td>0.593</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HP.C</td>
<td>0.018</td>
<td>0.381</td>
<td>0.289</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HP</td>
<td>0.056</td>
<td>-0.068</td>
<td>-0.049</td>
<td>-0.045</td>
<td>-0.046</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SMB</td>
<td>0.062</td>
<td>0.032</td>
<td>0.046</td>
<td>0.064</td>
<td>0.071</td>
<td>0.140</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HML</td>
<td>-0.017</td>
<td>-0.069</td>
<td>-0.038</td>
<td>-0.065</td>
<td>-0.094</td>
<td>0.502</td>
<td>0.099</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RMW</td>
<td>-0.030</td>
<td>0.030</td>
<td>0.016</td>
<td>0.002</td>
<td>-0.039</td>
<td>-0.446</td>
<td>-0.302</td>
<td>-0.498</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CMA</td>
<td>-0.058</td>
<td>0.000</td>
<td>-0.042</td>
<td>-0.006</td>
<td>-0.023</td>
<td>-0.151</td>
<td>0.094</td>
<td>0.144</td>
<td>0.049</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RNSK(B)</td>
<td>-0.065</td>
<td>0.079</td>
<td>-0.053</td>
<td>-0.063</td>
<td>-0.096</td>
<td>-0.003</td>
<td>-0.007</td>
<td>-0.007</td>
<td>0.026</td>
<td>-0.032</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RNSK</td>
<td>0.006</td>
<td>0.011</td>
<td>0.005</td>
<td>-0.139</td>
<td>-0.121</td>
<td>-0.010</td>
<td>0.038</td>
<td>0.015</td>
<td>0.048</td>
<td>-0.019</td>
<td>0.721</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SK</td>
<td>0.122</td>
<td>0.131</td>
<td>-0.141</td>
<td>0.060</td>
<td>0.081</td>
<td>-0.0004</td>
<td>-0.044</td>
<td>0.027</td>
<td>0.002</td>
<td>0.011</td>
<td>-0.213</td>
<td>-0.203</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LIQUID</td>
<td>0.038</td>
<td>-0.328</td>
<td>-0.185</td>
<td>-0.059</td>
<td>-0.090</td>
<td>-0.026</td>
<td>0.071</td>
<td>0.014</td>
<td>-0.022</td>
<td>-0.027</td>
<td>-0.083</td>
<td>0.067</td>
<td>-0.0001</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CV</td>
<td>-0.253</td>
<td>-0.077</td>
<td>-0.044</td>
<td>-0.163</td>
<td>-0.138</td>
<td>-0.040</td>
<td>0.004</td>
<td>-0.059</td>
<td>0.036</td>
<td>0.025</td>
<td>-0.050</td>
<td>-0.066</td>
<td>-0.153</td>
<td>0.049</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IDIOSK</td>
<td>0.114</td>
<td>-0.207</td>
<td>0.119</td>
<td>-0.194</td>
<td>-0.189</td>
<td>0.041</td>
<td>-0.036</td>
<td>0.031</td>
<td>0.004</td>
<td>-0.036</td>
<td>-0.142</td>
<td>-0.093</td>
<td>0.303</td>
<td>-0.012</td>
<td>0.122</td>
<td></td>
</tr>
<tr>
<td>∆OP</td>
<td>0.200</td>
<td>-0.101</td>
<td>-0.111</td>
<td>0.012</td>
<td>0.026</td>
<td>-0.124</td>
<td>0.052</td>
<td>0.007</td>
<td>-0.088</td>
<td>-0.060</td>
<td>0.050</td>
<td>-0.022</td>
<td>-0.009</td>
<td>0.065</td>
<td>0.021</td>
<td>-0.037</td>
</tr>
</tbody>
</table>

Notes: following the table 2.2, this table shows pair-wise correlation matrix between two risk factors (long-short portfolio) with more extensions on the number of factors. All factor names are listed on both the first row and column. In addition to factors specified in table 2.2, more factors motivated from stock market (Fama-French five factors) are include: Mk (Long only market portfolio), SMB (long-short portfolio sorted by company market capitalization), HML (long-short portfolio sorted by book-to-market ratio), RMW (long-short portfolio sorted by firms’ operating profitability), CMA (long-short portfolio sorted by investing style). Bold value in this table means at least 90% significant.

HP, TS and RNSK are in general performing better than other trading strategies, among which SK is not good as suggested from Fernandez-Perez et al. (2018) and the equally weighted long only portfolio (also known as commodity market portfolio) are in the worst group. The reason for Pearson skewness coefficient failing to delivery good performance may due to the time period selection, observation frequency as well as underlying commodity assets difference. The recent financial crisis could be another big impactor for momentum and commodity market portfolio as they mainly reply on the global market trend. It is worth mentioning that momentum factor provides nearly no benefits (starts with 50% cumulative return immediately after financial crisis and ends at slightly over 50%
cumulative return) to invest, which can be also evidenced by its lowest annual mean return among traditional factors from Table 2.1. Overall, RNSK is superior to TS, HP.C and HP and beats other trading strategies within this sample period.

The quartile portfolio cumulative performance is plotted in Figure 2.2 to show the dynamics of each quartile portfolio. Since RNSK is computed in a rolling window manner, the first one year sample is not available to plot. Two extreme quartiles, P1 and P4, show opposite cumulative return path with strong asymmetric property on the lowest RNSK group performing much sever than the highest RNSK group.

In addition to its superior performance, RNSK is also more flexible and has less parametrization.
problem compared with the realised skewness (e.g. Pearson skewness coefficient) when constructing a long-short portfolio. The merits of using RNSK can be summarized as parametrization reduction in two aspects: window length and data frequency. Two different data frequencies (daily and weekly) and five rolling window lengths (1 month, 3 months, 6 months, 9 months, 12 months and 24 months) for the realised skewness calculation are taken into account. Dynamics of the skewness and corresponding sorted long-short portfolios are constructed in figure 2.3.

From figure 2.3, it is easy to see that estimation results are not proportional to each other when using the same rolling window but different data frequency (return is not normally distributed). Results are more likely to be impacted by some outliers in the data distribution, which deteriorates the calculation precision (Kim and White 2004, Neuberger 2012 and Hansis et al. 2010). In terms of
Figure 2.3: Skewness Comparison Analysis - Parametrization Problem

These four figures show consideration on estimation window and data frequency usage when signal is estimated via realised skewness (Pearson Skewness coefficient used here, same method used in Fernandez-Perez et al. [2018]). From the top plots to the bottom one, data are formatted in daily, weekly, daily and weekly frequency. Therefore, for the first and the third plot, daily observation starts from 10/10/2007 to 01/03/2016, with T = 1, . . . , 2119; for the second and the forth figures, data is on the same time span but T = 1, . . . , 433. Rolling window is scheduled as (30, 90, 125, 252 and 504) days and (5, 15, 26, 52 and 104) weeks for (daily) and (weekly) signal and portfolio generation purpose separately. Regarding the value of plots, the upper two plots display the dynamics of averaged cross-sectional Pearson skewness coefficient estimators based on different rolling windows (depends on different frequencies). By saying averaged signal, individual signal on rolling manner is first obtained and cross-sectional commodities’ values mean are calculated for dynamic plots. Risk-neutral estimated dynamics plot is not reported here as there is no problem on window selection issue. The bottom two plots are the top two plots corresponding rolling window estimators based long-short trading strategy cumulative return. Portfolio rebalance frequency is daily and weekly repetitively. In these bottom two plots, long-short portfolio return is reported for the Risk-Neutral Skewness as well.
the return performance plot in bottom two panels, the realised skewness varies within different data frequencies and rolling window lengths while RNSK works stable throughout two frequencies. Overall, RNSK can provide stable long-short portfolio return and free of parameters selection, which is more general and flexible for practitioner usage purpose.

2.6.3 Risk-Neutral Skewness Characteristics

This section explores pricing behaviour of RNSK by taking into account the traditional suggested factors. Results can be found at table 2.4 where mean of sorted factors at each quartile group and high-quartile minus low-quartile (HML) t-test value are calculated and reported.

Although RNSK differs from the "lottery-like" behaviour and positive skewness chasing idea under the cumulative prospect theory, the underlying intuition behind is still on asset’s mispricing: over-pricing and under-pricing. Overall, RNSK shows consistent group mean value sorted by SK factor in panel A, increasing from lowest group -0.24897 to highest group 0.16366. This relation is also supported by the fact that the HML quartile performance has a significant t-test value.

In panel B of table 2.4, all factors are sorted by the RNSK to identify their relations to the RNSK pricing characteristics. In general, the SK factor mean values for all quartile groups show a decreasing trend from P1 to P4. However, the SK factor value signs are always negative, which differs significantly from the opposite sorting procedure results from the panel A. Meanwhile, group mean value for the RNSK in panel B implies magnitude difference compared panel A results when sorting based on SK. The main conclusion here is, the RNSK is well matched from the SK sorting but in meantime shows more distinct character on sorting behaviour that is evidenced by non-trivial magnitude first row (RNSK) and all negative values in second row (SK).

Consistent with Fernandez-Perez et al. (2018), P1 (P4) quartiles in panel B shows strong backwar-
### Table 2.4: Risk-Neutral Skewness Quartile Characteristics

<table>
<thead>
<tr>
<th></th>
<th>P1</th>
<th>P2</th>
<th>P3</th>
<th>P4</th>
<th>HML</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel A: Post-Ranking via Pearson Skewness</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RNSK</td>
<td>-0.24897</td>
<td>-0.02201</td>
<td>0.07194</td>
<td>0.16366</td>
<td>0.21916***</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(3.289081)</td>
</tr>
<tr>
<td><strong>Panel B: Post-Ranking via Risk-Neutral Skewness</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RNSK</td>
<td>-1.48903</td>
<td>-0.16700</td>
<td>0.50670</td>
<td>1.67060</td>
<td>1.54302***</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(41.17034)</td>
</tr>
<tr>
<td>SK</td>
<td>-0.16827</td>
<td>-0.0878</td>
<td>-0.06521</td>
<td>-0.06364</td>
<td>0.0523***</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(7.981162)</td>
</tr>
<tr>
<td>Basis</td>
<td>-0.00062</td>
<td>-0.00019</td>
<td>-0.00099</td>
<td>-0.00119</td>
<td>-0.00029***</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(-2.162495)</td>
</tr>
<tr>
<td>Momentum</td>
<td>0.00028</td>
<td>-0.00021</td>
<td>-0.00109</td>
<td>-0.00156</td>
<td>-0.00092***</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(-12.13342)</td>
</tr>
<tr>
<td>Hedging pressure (Hedgers)</td>
<td>0.39108</td>
<td>0.40661</td>
<td>0.43370</td>
<td>0.43060</td>
<td>0.01976***</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(12.31988)</td>
</tr>
<tr>
<td>Hedging pressure (Speculators)</td>
<td>0.59947</td>
<td>0.58455</td>
<td>0.55984</td>
<td>0.56232</td>
<td>-0.01858***</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(-7.26529)</td>
</tr>
</tbody>
</table>

Notes: this table shows characteristics of the Risk-Neutral Skewness via quartile ranking and sorting. Specifically, in panel A, RNSK is calculated by taking the average according to the cross-sectional rank of Pearson skewness coefficient in increasing order within each group from P1 (lowest group) to P4 (highest group). HML (High minus Low) is difference between P4 and P1 and its corresponding t statistics test is reported in bracket. In panel B, all things keep same as panel A except for group average calculation is sorted according to the cross-sectional rank of RNSK in increasing order from P1 to P4.

dation (contango) ideas, which is accompanied by high (low) values of Basis and HP from speculators and by low (high) values of HP from hedgers. From another perspective, HP from either hedgers or
speculators are consistent with the literature finding on RNSK pricing resource \([\text{Han}] 2008\). When HP value (long position) is small for hedgers, a negative expectation about future price movement is formed, implying that hedgers are more likely to short assets to avoid risks. Speculators are standing on the opposite side, therefore HP value will be different from hedgers. To describe to what extent their characteristics correlated to assets selection, at the stage of portfolio formation, it is quantified by exploring how many same assets overlap at each extreme quartile group in a time-series dimension.

**Figure 2.4: Extreme Quartile Commodity Assets Overlapping Analysis**

Notes: this figure characterize the overlapping information during the quantile ranking procedure between RNSK and other risk factors \((\text{realised skewness, term structure, momentum and hedging pressure})\). The number of overlapping is calculated based on trading signal on each quantile. For \text{realised skewness and hedging pressure}, top (bottom) quantile of RNSK is matched with bottom (top) quantile of it. For \text{term structure, and momentum}, bottom (top) quantile of RNSK is matched with bottom (top) quantile of it. Red line and dot stand for number of overlapping for top quantile and blue line and dot stand for that for bottom quantile.

Specifically, a time-series overlapping number at both highest and lowest quartile will be recorded. The purpose of this procedure is to test long-short portfolio assets selection similarities between
RNSK and the traditional factors. For term structure, momentum and hedging pressure, overlapping with RNSK use Top-to-Top (long assets in momentum, term structure and hedging pressure with long assets in RNSK) and Bottom-to-Bottom (short assets in momentum, term structure and hedging pressure with short assets in RNSK) quartile, while Top-to-Bottom (Bottom-to-Top) is only for realised skewness. From figure 2.4, all of them are less than 4 (full overlapping number at each quartile). It is still possible to observe that 3 same assets overlapped at one balance time for all factors at certain time point, indicating strong similar signal generation idea but not consistent for most periods.

RNSK overlaps more with SK factor during the 2008 financial crisis, which implies that the external driving force, market trend, has caused the same pricing mechanism between these two factors. Term structure and momentum has more clear observation from Top-to-Top quartile number than Bottom-to-Bottom, implying backwardation and winner group are more exposed to high implied skewness. Overall, except for certain time point, the overlapping information indicates that RNSK has more information that are not exposed to the current known baseline model and the realised skewness.

2.6.4 Time Series Analysis

In this section, time series analysis is conducted for each quartile sorted groups. By doing so, portfolio performance analysis is extended to the quartile level. Secondly, OLS regression method is applied for each quartile to quantify new factors’ exposure with respect to traditional risk factors. Moreover, the extra unexplained information generated by new risk factor, called alpha (well known as the abnormal return which excludes the known information), is also studied on both value and significance.

Long-short portfolio (RNSK) is regressed on equally weighted portfolio (standing for sample overall long only performance), term structure (market backwardation and contango information), hedging pressure (market expectation from trading participators) and momentum portfolio (market
where, $PR_{i,t}$ is the quartile RNSK sorted portfolio return at time $t$, for instance $PR_{1,t}$ is matched for P1 at time $t$. $\alpha_i$ is the abnormal return in $i^{th}$ portfolio regression analysis. $\epsilon_{i,t}$ is the $i^{th}$ portfolio regressed error term at time $t$. $i$ is the index correlated to which portfolio specify from P1 to P4. $TS$, $HP$ and $MOM$ are buying and selling portfolio returns, which are identical for each asset time series quartile regression.

From table 2.5, panel A, for the annualized mean, it is obvious that there is a general upward-sloping trend for quartile portfolio return from P1 to P4, from -19.7% to 6.8%, showing asymmetric property documented in the literature. After including selling availability, HML portfolio delivers annual 13% return with Sharpe ratio 1.39, which is significant higher than each quartile long only strategy (P1 to P4) and other HML strategies (based on traditional risk factors) mentioned in the literature (considering the Sharpe ratio at global futures market so far, 0.47 term structure strategy in Erb and Harvey (2005), 0.75 hedging pressure strategy in Basu and Miffre (2013), 0.67 Carry strategy in Koijen et al. (2013), value and momentum factors less than 1 in normal case in Asness et al. (2013) and 1.1 time series momentum strategy with all available futures products in global market in Moskowitz et al. (2012)). Moreover, RNSK sorted HML portfolio is distributional close to the normal with less asymmetry and flat tail. Compared with single buying portfolio, it is more reliable when applying in the real market with less risk to some extent. It also shows the smallest VaR value among all portfolios.

From the panel B, alpha (abnormal return) is significant for both two extreme quartiles at 99% significant level, demonstrating the fact that the baseline model from the traditional commodity risk
Table 2.5: Time Series Analysis – Portfolio Property based on Risk-Neutral Skewness

<table>
<thead>
<tr>
<th></th>
<th>P1</th>
<th>P2</th>
<th>P3</th>
<th>P4</th>
<th>HML</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>-0.1971&lt;sup&gt;***&lt;/sup&gt;</td>
<td>-0.0534</td>
<td>-0.0168</td>
<td>0.0686</td>
<td>0.1328&lt;sup&gt;***&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td>(-2.6922)</td>
<td>(-0.7587)</td>
<td>(-0.2059)</td>
<td>(0.9350)</td>
<td>(4.0211)</td>
</tr>
<tr>
<td>StDev</td>
<td>0.2113</td>
<td>0.2029</td>
<td>0.2359</td>
<td>0.0668</td>
<td>0.3459</td>
</tr>
<tr>
<td>Sharpe</td>
<td>-0.1615</td>
<td>-0.0501</td>
<td>-0.0137</td>
<td>0.0668</td>
<td>0.3459</td>
</tr>
<tr>
<td>Skewness</td>
<td>-0.3125</td>
<td>-0.0741</td>
<td>-0.2197</td>
<td>0.0324</td>
<td>0.4035</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>4.9285</td>
<td>4.7057</td>
<td>4.4848</td>
<td>5.4418</td>
<td>3.8076</td>
</tr>
<tr>
<td>99% VAR(Cornish-Fisher)</td>
<td>-0.1060</td>
<td>-0.1035</td>
<td>-0.1157</td>
<td>-0.1123</td>
<td>-0.0281</td>
</tr>
<tr>
<td>MaxDrawdown</td>
<td>0.8547</td>
<td>0.6003</td>
<td>0.5239</td>
<td>0.4111</td>
<td>0.0834</td>
</tr>
<tr>
<td>% of positive months</td>
<td>0.4018</td>
<td>0.4434</td>
<td>0.4688</td>
<td>0.4642</td>
<td>0.5150</td>
</tr>
</tbody>
</table>

Panel A: Quantile Portfolio Performance

Panel B: Time-Series Regression

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
<th>(8)</th>
<th>(9)</th>
<th>(10)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alpha</td>
<td>-0.1132&lt;sup&gt;***&lt;/sup&gt;</td>
<td>-0.1180&lt;sup&gt;***&lt;/sup&gt;</td>
<td>0.0243</td>
<td>0.0227</td>
<td>0.0788</td>
<td>0.0804</td>
<td>0.1789&lt;sup&gt;***&lt;/sup&gt;</td>
<td>0.1814&lt;sup&gt;***&lt;/sup&gt;</td>
<td>0.1460&lt;sup&gt;***&lt;/sup&gt;</td>
<td>0.1500&lt;sup&gt;***&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td>(0.0009)</td>
<td>(0.0009)</td>
<td>(0.0009)</td>
<td>(0.0009)</td>
<td>(0.0012)</td>
<td>(0.0012)</td>
<td>(0.0009)</td>
<td>(0.0009)</td>
<td>(0.0009)</td>
<td>(0.0007)</td>
</tr>
<tr>
<td>EW</td>
<td>1.0388&lt;sup&gt;***&lt;/sup&gt;</td>
<td>1.0244&lt;sup&gt;***&lt;/sup&gt;</td>
<td>1.0699&lt;sup&gt;***&lt;/sup&gt;</td>
<td>1.0586&lt;sup&gt;***&lt;/sup&gt;</td>
<td>1.0326&lt;sup&gt;***&lt;/sup&gt;</td>
<td>1.0369&lt;sup&gt;***&lt;/sup&gt;</td>
<td>1.0511&lt;sup&gt;***&lt;/sup&gt;</td>
<td>1.0620&lt;sup&gt;***&lt;/sup&gt;</td>
<td>0.0082</td>
<td>0.0187</td>
</tr>
<tr>
<td></td>
<td>(0.0375)</td>
<td>(0.0375)</td>
<td>(0.0375)</td>
<td>(0.0375)</td>
<td>(0.0511)</td>
<td>(0.0511)</td>
<td>(0.0369)</td>
<td>(0.0368)</td>
<td>(0.0295)</td>
<td>(0.0294)</td>
</tr>
<tr>
<td>MOM</td>
<td>-0.0164</td>
<td>0.0136&lt;sup&gt;***&lt;/sup&gt;</td>
<td>0.1954&lt;sup&gt;***&lt;/sup&gt;</td>
<td>0.1992&lt;sup&gt;***&lt;/sup&gt;</td>
<td>0.0995</td>
<td>0.0887</td>
<td>0.0059</td>
<td>-0.0110</td>
<td>0.0111</td>
<td>-0.0148</td>
</tr>
<tr>
<td></td>
<td>(0.0473)</td>
<td>(0.0474)</td>
<td>(0.0474)</td>
<td>(0.0475)</td>
<td>(0.0645)</td>
<td>(0.0646)</td>
<td>(0.0466)</td>
<td>(0.0466)</td>
<td>(0.0372)</td>
<td>(0.0372)</td>
</tr>
<tr>
<td>HP</td>
<td>0.1467&lt;sup&gt;***&lt;/sup&gt;</td>
<td>0.1158</td>
<td>0.1676&lt;sup&gt;***&lt;/sup&gt;</td>
<td>0.1642&lt;sup&gt;***&lt;/sup&gt;</td>
<td>-0.2809&lt;sup&gt;***&lt;/sup&gt;</td>
<td>-0.2813&lt;sup&gt;***&lt;/sup&gt;</td>
<td>-0.1526&lt;sup&gt;***&lt;/sup&gt;</td>
<td>-0.1496&lt;sup&gt;***&lt;/sup&gt;</td>
<td>-0.1267&lt;sup&gt;***&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0734)</td>
<td>(0.0737)</td>
<td>(0.0735)</td>
<td>(0.0738)</td>
<td>(0.1001)</td>
<td>(0.1005)</td>
<td>(0.0722)</td>
<td>(0.0724)</td>
<td>(0.0577)</td>
<td>(0.0579)</td>
</tr>
<tr>
<td>SK</td>
<td>0.0673&lt;sup&gt;***&lt;/sup&gt;</td>
<td>0.2296</td>
<td>-0.0707</td>
<td>-0.1114&lt;sup&gt;**&lt;/sup&gt;</td>
<td>-0.175&lt;sup&gt;**&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adj.R.square</td>
<td>0.6736</td>
<td>0.687</td>
<td>0.6456</td>
<td>0.6488</td>
<td>0.5144</td>
<td>0.5145</td>
<td>0.686</td>
<td>0.6886</td>
<td>0.0069</td>
<td>0.044</td>
</tr>
</tbody>
</table>

Adjusted R-square: 0.6736, 0.687, 0.6456, 0.6488, 0.5144, 0.5145, 0.686, 0.6886, 0.0069, 0.044

Notes: weekly Quantile RNSK sorted equally weighted portfolio performance and regression analysis on baseline models results from 10/10/2007 to 01/03/2016 are reported in this table. In panel A, portfolio performance statistics are reported for long only portfolio sorted by RNSK from P1 (lowest RNSK group) to P4 (highest RNSK group) and HML (high RNSK minus low RNSK group). In panel B, each quantile time-series portfolio is regressed on the baseline model (EW, TS, MOM, and HP) and baseline model plus SK factor for robustness check. The first row in panel B reports the annualized mean with coefficient multiplied by 52. Standard errors are reported under the estimated coefficients, with standard error and Newey-West corrected standard error (12 weeks lags setting) in the round and squared bracket respectively.

Factors cannot explain under-performance (over-performance) in lowest (highest) quartile. Regarding to the beta significance for each quartile group, term structure and momentum betas are not significant overall. Hedging pressure, on the contrary, presents strong statistical significances in most quartiles analysis. However, the signs of hedging pressure in the lower quartile is opposite from the higher quartile. HML group implies that hedging pressure based trading strategy has opposite trading signal from RNSK based on trading strategy<sup>27</sup>. On the level of the long-short portfolio, alpha is also signifi-

<sup>27</sup>This is consistent with Speculators’ perspective, which echo the RNSK pricing mechanism from the perspective of
cant with annual return up to 14.6%. $R^2$, measuring regression fitting degree, with low value (4.4% in HML column) indicates that baseline more is not a good candidate for explaining HML RNSK sorted portfolio. This evidence is also robust when SK factor is controlled on each quartile group following the baseline model. This implies that performance from RNSK cannot be explained by both baseline model and the realised skewness sorted portfolio.

In order to give clearer picture of alpha dynamic path in time-series analysis, rolling regression with window length 1 year on weekly observation is used. Specifically, the RNSK long-short portfolio return is regressed on the baseline traditional commodity risk factors model using the first 52 observations and then repeat this step when one new observation comes in (adding it into the end of 52 sample and dropping the first observation, making total number of observation in regression consistent and fixed, always 52). Alpha dynamics and its corresponding 95% bootstrap confidence interval are then plotted.

From figure 2.5, the mean for rolling alpha is 0.00289, or 15% annually. The worst performance is around the year 2011 and then average around 0.00189 (9.84% annually) afterwards. Generally, abnormal return is non-trivial in value across whole testing period as most of it is above zero value line despite the fact that zero value line is fall into bootstrap bound making this abnormal return insignificant somehow.

### 2.6.5 Robustness Check

This section employs other risk factors that have been discussed in the recent literature for controlling effect check. The reason to select following factors is due to their extra pricing effects in addition to the baseline model. Purpose here is to clarify whether alpha generated from RNSK is still significant and robust when controlling these extra pricing effects mentioned before. Moreover, transaction costs hedgers.
Notes: the RNSK is regressed on commodity baseline model: EW, TS, MOM and HP with 52 weeks rolling window. Starting from 1 to 52 observations, rolling forward, dropping the sample first observation and adding one new observation at the sample end. At each subsample regression, 1000 bootstrap sampling is applied to estimate the confidence interval and lower 5% and upper 95% bounds are extracted. The figure has dropped the first 52 observations as they are used for initializing this regression procedure. Red line is estimated rolling alpha, blue line is upper 95% bound and green line is lower 5% bound.

To be specific, the following six factors are utilized with five of them from the commodity futures market and one of them from stock market: (1) SK, Pearson skewness coefficient, long-short portfolio sorted by third central moment of daily futures return, (2) IDIOSK, long-short portfolio sorted by skewness calculated on the residual (obtained from regression of asset return on baseline model), (3) CV, long-short portfolio sorted by variance-over-mean of daily futures returns over prior 36 months, (4) LIQUID, long-short portfolio sorted by prior 2-month dollar volume over absolute return, (5) ∆OP, long-short portfolio sorted by the change of entire open interest of commodity futures,
(6) Fama_French five factors, motivated by stock market: Mk, SMB, HML, RMW (Robust Minus Weak)\(^{28}\) and CMA (Conservative Minus Aggressive)\(^{29}\). (7) Transaction cost analysis is conducted by deducting 6.6 bps per trade from Risk-Neutral Skewness portfolio (Fernandez-Perez et al. (2018), Hong and Yogo (2012), Erb and Harvey (2005), Amihud (2002) and Locke and Venkatesh (1997)).

From table 2.6 and table 2.7, in general, alpha is significant through all scenarios despite the fact that \(R^2\) changes a bit with different factors. Alpha is annualized and generated with 15.6\% percentage average return for generally all cases at 99\% significance level. One exception on alpha value is the column 13 where alpha reduces to 13.2\% due to the fact that all explanatory variables are added in regression model, increasing the total explaining power, see R square at 0.049.

Another exception is for transaction cost analysis in the last two columns, with around 5.6\% annually and 95\% significance level. One finding to be well expected is the sign and significance of idiosyncratic skewness (IDIOSK)\(^{30}\). Following the test result from (Fernandez-Perez et al. 2018), idiosyncratic skewness has the same characteristics as Pearson skewness coefficient, and therefore, it is reasonable to expect a negative correlation between IDIOSK and RNSK. Moreover, hedging pressure measurements from speculators effect is also tested with row name labelled as HP. This result is not consistent from Han (2008) as his calculation is on net position rather than long position only. Research focus in his study is on the S&P 500 index future, which aggregates the individual stock level effect\(^{31}\). Table 2.7 suggests that portfolio sorted by hedging pressure and by RNSK are performing in opposite way. The difference could come up with dispersion effect among individual commodity futures.

\(^{28}\)the average return on the two robust operating profitability portfolios minus the average return on the two weak operating profitability portfolios

\(^{29}\)the average return on the two conservative investment portfolios minus the average return on the two aggressive investment portfolios

\(^{30}\)Because of its property as a sentiment proxy variable argued in pricing mechanism.

\(^{31}\)Analysis for individual commodity Risk-Neutral Skewness is repeated on hedging pressure, finding mixed results for individual level.
Table 2.6: Risk-Neutral Skewness Time-Series Portfolio Performance Test

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
<th>(8)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alpha</td>
<td>0.156***</td>
<td>0.156***</td>
<td>0.156***</td>
<td>0.156***</td>
<td>0.156***</td>
<td>0.156***</td>
<td>0.156***</td>
<td>0.156***</td>
</tr>
<tr>
<td></td>
<td>[0.001]</td>
<td>[0.001]</td>
<td>[0.001]</td>
<td>[0.001]</td>
<td>[0.001]</td>
<td>[0.001]</td>
<td>[0.001]</td>
<td>[0.001]</td>
</tr>
<tr>
<td>CV</td>
<td>−0.074</td>
<td>−0.112*</td>
<td>−0.112*</td>
<td>−0.112*</td>
<td>−0.112*</td>
<td>−0.112*</td>
<td>−0.112*</td>
<td>−0.112*</td>
</tr>
<tr>
<td></td>
<td>(0.064)</td>
<td>(0.058)</td>
<td>(0.058)</td>
<td>(0.058)</td>
<td>(0.058)</td>
<td>(0.058)</td>
<td>(0.058)</td>
<td>(0.058)</td>
</tr>
<tr>
<td>SK</td>
<td>−0.182***</td>
<td>−0.169***</td>
<td>−0.169***</td>
<td>−0.169***</td>
<td>−0.169***</td>
<td>−0.169***</td>
<td>−0.169***</td>
<td>−0.169***</td>
</tr>
<tr>
<td></td>
<td>[0.051]</td>
<td>[0.050]</td>
<td>[0.050]</td>
<td>[0.050]</td>
<td>[0.050]</td>
<td>[0.050]</td>
<td>[0.050]</td>
<td>[0.050]</td>
</tr>
<tr>
<td>IDIOSK</td>
<td>−0.091</td>
<td>−0.034</td>
<td>−0.034</td>
<td>−0.034</td>
<td>−0.034</td>
<td>−0.034</td>
<td>−0.034</td>
<td>−0.034</td>
</tr>
<tr>
<td></td>
<td>[0.055]</td>
<td>[0.046]</td>
<td>[0.046]</td>
<td>[0.046]</td>
<td>[0.046]</td>
<td>[0.046]</td>
<td>[0.046]</td>
<td>[0.046]</td>
</tr>
<tr>
<td>LIQUID</td>
<td></td>
<td>0.068</td>
<td></td>
<td>0.068</td>
<td></td>
<td>0.068</td>
<td></td>
<td>0.068</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[0.076]</td>
<td></td>
<td>[0.076]</td>
<td></td>
<td>[0.078]</td>
<td></td>
<td>[0.078]</td>
</tr>
<tr>
<td>∆OP</td>
<td></td>
<td></td>
<td>−0.020</td>
<td></td>
<td></td>
<td>−0.020</td>
<td></td>
<td>−0.011</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>[0.042]</td>
<td></td>
<td></td>
<td>[0.038]</td>
<td></td>
<td>[0.038]</td>
</tr>
<tr>
<td>EW</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MOM</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HP</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mk</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SMB</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HML</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RMW</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CMA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Adjusted $R^2$          0.002  0.008  0.006  0.010  0.002  0.006 −0.002  0.003

F Statistic  1.893 (df = 1; 431)  2.847* (df = 2; 430)  3.744* (df = 1; 431)  3.261** (df = 2; 430)  1.927 (df = 1; 431)  2.349* (df = 2; 430)  0.200 (df = 1; 431)  1.675 (df = 2; 430)

Notes: weekly time-series regression analysis is repeated for 14 groups (different models) from (1) to (15) shown in the first row to explore the significance of alpha (at the top of each model). From (1) to (12), every odd column, RNSK portfolio is regressed on one of following commodity risk factors (CV, IDIOSK, LIQUID and ∆OP), baseline model and stock market factors sequentially in weekly frequency, while in each even column, extra SK factors is added for robustness check. In column (13), all explanatory variables are included to for full regression test purpose. In column (14), transaction cost is added by deducting trading cost per trade 6.6 bps from RNSK portfolio return. Column (15) is corresponding transaction cost robustness check when realised skewness (SK) factor is controlled. Newey-West corrected standard error is reported in the square bracket with 12 weeks lags setting.

In addition to the alpha coefficient and its significance, the negative significant relation between RNSK and SK in table 2.6 is consistent with the factors pairwise correlation matrix result from table
Table 2.7: Risk-Neutral Skewness Time-Series Portfolio Performance Test Continues

<table>
<thead>
<tr>
<th></th>
<th>(9)</th>
<th>(10)</th>
<th>(11)</th>
<th>(12)</th>
<th>(13)</th>
<th>(14)</th>
<th>(15)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alpha</td>
<td>0.156**</td>
<td>0.156**</td>
<td>0.156**</td>
<td>0.156**</td>
<td>0.132***</td>
<td>0.0561**</td>
<td>0.0598***</td>
</tr>
<tr>
<td></td>
<td>[0.001]</td>
<td>[0.001]</td>
<td>[0.001]</td>
<td>[0.001]</td>
<td>[0.001]</td>
<td>[0.001]</td>
<td>[0.001]</td>
</tr>
<tr>
<td>CV</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-0.126*</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>[0.044]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SK</td>
<td>-0.167***</td>
<td>-0.167***</td>
<td>-0.167***</td>
<td>-0.164***</td>
<td>-0.166***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[0.052]</td>
<td>[0.054]</td>
<td>[0.051]</td>
<td>[0.001]</td>
<td>[0.054]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IDIOSK</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-0.061</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>[0.175]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LIQUID</td>
<td>0.062</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[0.423]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ΔOP</td>
<td></td>
<td>-0.004</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>[0.928]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EW</td>
<td>0.002</td>
<td>0.013</td>
<td>-0.002</td>
<td>0.007</td>
<td>0.017</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[0.031]</td>
<td>[0.029]</td>
<td>[0.995]</td>
<td>[0.030]</td>
<td>[0.027]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TS</td>
<td>0.027</td>
<td>0.048</td>
<td>0.067</td>
<td>0.036</td>
<td>0.055</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[0.053]</td>
<td>[0.053]</td>
<td>[0.223]</td>
<td>[0.053]</td>
<td>[0.054]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MOM</td>
<td>0.010</td>
<td>-0.015</td>
<td>-0.022</td>
<td>0.009</td>
<td>-0.015</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[0.051]</td>
<td>[0.045]</td>
<td>[0.649]</td>
<td>[0.049]</td>
<td>[0.045]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HP</td>
<td>-0.150**</td>
<td>-0.126*</td>
<td>-0.156*</td>
<td>-0.149**</td>
<td>-0.126*</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[0.073]</td>
<td>[0.072]</td>
<td>[0.029]</td>
<td>[0.074]</td>
<td>[0.073]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mk</td>
<td>0.0004</td>
<td>-0.002</td>
<td>-0.003</td>
<td>-0.002</td>
<td>0.004</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[0.041]</td>
<td>[0.040]</td>
<td>[0.039]</td>
<td>[0.038]</td>
<td>[0.908]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SMB</td>
<td>0.130**</td>
<td>0.111</td>
<td>0.143**</td>
<td>0.112</td>
<td>0.103</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[0.066]</td>
<td>[0.072]</td>
<td>[0.067]</td>
<td>[0.073]</td>
<td>[0.152]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HML</td>
<td>0.038</td>
<td>0.050</td>
<td>0.013</td>
<td>0.022</td>
<td>0.0001</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[0.088]</td>
<td>[0.086]</td>
<td>[0.090]</td>
<td>[0.090]</td>
<td>[0.999]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RWA</td>
<td>0.277</td>
<td>0.277</td>
<td>0.229</td>
<td>0.232</td>
<td>0.223</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[0.175]</td>
<td>[0.178]</td>
<td>[0.178]</td>
<td>[0.182]</td>
<td>[0.189]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CMA</td>
<td>-0.143</td>
<td>-0.135</td>
<td>-0.142</td>
<td>-0.116</td>
<td>-0.096</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[0.195]</td>
<td>[0.191]</td>
<td>[0.195]</td>
<td>[0.190]</td>
<td>[0.606]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adjusted $R^2$</td>
<td>-0.005</td>
<td>-0.005</td>
<td>0.002</td>
<td>0.005</td>
<td>0.049</td>
<td>0.007</td>
<td>0.0425</td>
</tr>
</tbody>
</table>

| F Statistic | 0.556 (df = 5; 427) | 0.966 (df = 6; 426) | 1.099 (df = 9; 423) | 1.234 (df = 10; 422) | 2.595 (df = 14; 418) | 1.801 (df = 4; 428) | 4.843 (df = 5; 427) |

$*** \ p < 0.01, ** \ p < 0.05, * \ p < 0.1$

2.3 in which RNSK and SK are negatively correlated at -0.203.

### 2.6.6 Cross Sectional Analysis

Cross-sectional regression approach is applied following the two-step method by Fama and MacBeth [1973]. In the first step, time series OLS regression is used to estimate the betas coefficient for each
asset. Then, assuming all betas are constant for the whole sample period, the second step is to run cross-sectional regression at each time period to get the risk premium. Specifically, cross-sectional asset returns are regressed against all constant betas, yielding the risk premium called gamma. Finally, gammas are averaged out to get the averaged estimated risk premiums and their t-statistics are calculated as well. Intuitively, the idea held here is to explore the average risk premium that can be obtained by one unit change risk factor exposure beta. In model set-up, the first step follows as:

$$R_{i,t} = \alpha_i + \beta_{1,i} \cdot EW_t + \beta_{2,i} \cdot TS_t + \beta_{3,i} \cdot HP_t + \beta_{4,i} \cdot MOM_t + \beta_{5,i} \cdot RNSK_t + \epsilon_{i,t} \quad (2.6.2)$$

where, $i$ is standing for each asset instead of each portfolio in time series regression. $R_i$ is the $i^{th}$ single asset return in the time series dimension. $\beta_i$ is the $i^{th}$ asset risk exposure coefficients set to risk factors.

In the second step, the estimated beta from the first step is used for explanatory variable:

$$R_{i,t} = \alpha_{i,t} + \gamma_{1,t} \cdot \hat{\beta}_{1,i} + \gamma_{2,t} \cdot \hat{\beta}_{2,i} + \gamma_{3,t} \cdot \hat{\beta}_{3,i} + \gamma_{4,t} \cdot \hat{\beta}_{4,i} + \gamma_{5,t} \cdot \hat{\beta}_{5,i} + \mu_t \quad (2.6.3)$$

To account for "Error-in-Variable" problem (second-step explanatory variables are estimated value from the first step), corrected standard error in the spirit of Shanken (1992) is employed. Results also report cross-sectional correlation adjusted standard error via the idea from Newey and West (1987).

RNSK is passed to test with the inclusion of baseline model as well as other commodity risk factors and stock market Fama_French five factors for robustness check.

Table 2.8 and 2.9 show the Fama_Macbeth two-step regression results with consideration on: (1) the baseline model test, (2) case (1) + RNSK factor test and (3) case (1) + case (2) + more importantly Pearson skewness (SK) factor test. Generally speaking, starting from the baseline model (TS, HP
and MOM), one or more variables is or are sequentially added in regression test procedure. At the same time, for extra variable(s) adding, statistical property test is conducted for RNSK solely or plus Pearson skewness control effect.

In table 2.8 and 2.9, generally speaking, RNSK shows strong pricing ability across nearly all models (even SK factor is controlled). This can be observed via RNSK coefficients and t statistics values on the first row cross-sectional from the column (3) to (22). In another perspective, RNSK shows non-negligible contribution on regression fittings due to the fact that $R^2$ is improved in all cases when RNSK is added on. Lastly, RNSK is overall significantly negatively priced with averaged at least annualized premium 15.6% ($\lambda_5$ in Equation 2.6.3 is at least 0.3% each week, so annually $0.003 \times 52 = 15.6\%$).

Although this may seemingly counterintuitive, it can be: (1) explained by the negative correlation coefficient among estimated RNSK betas and SK betas at the first step regression results, indicating opposite pricing behaviour, see the second last row, $\rho(\hat{\beta}_{RNSK}, \hat{\beta}_{SK})$, of table 2.8 and 2.9 and (2) consistent with the negative significant pairwise correlation coefficient value from table 2.3.

Besides this, this RNSK negative pricing coefficient is linked to the RNSK pricing mechanism section where trading on options is a selective hedging tool to fight against potential downside risk. Following this idea, RNSK should reflect the same trading direction as HP.C (hedgers) does since both of these two variables refers to large hedgers’ market participation direction and the degree of their activities. Correspondingly, speculators’ market open interest should be opposite to RNSK. This explanation is confirmed by table 2.4 where post-ranking implied hedgers (speculators) are positively (negatively) correlated with RNSK. The negative signs for HP (speculators) in results from time series (table 2.5) and cross-sectional (table 2.8 and table 2.9) regressions are also consistent with this pricing mechanism.
Table 2.8: Risk-Neutral Skewness Cross-Sectional Fama-Macbeth Regression

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
<th>(8)</th>
<th>(9)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alpha</td>
<td>-0.0037</td>
<td>0.00059</td>
<td>0.00080</td>
<td>0.00011</td>
<td>0.00048</td>
<td>0.00078</td>
<td>-0.0028</td>
<td>0.00059</td>
<td>0.00080</td>
</tr>
<tr>
<td></td>
<td>(0.00102)</td>
<td>(0.00091)</td>
<td>(0.00099)</td>
<td>(0.00088)</td>
<td>(0.00085)</td>
<td>(0.00092)</td>
<td>(0.00100)</td>
<td>(0.00091)</td>
<td>(0.00099)</td>
</tr>
<tr>
<td>RNSK</td>
<td>-0.00380*</td>
<td>-0.00321*</td>
<td>-0.00390**</td>
<td>-0.00323*</td>
<td>-0.00380**</td>
<td>-0.00321*</td>
<td>-0.00380**</td>
<td>-0.00321*</td>
<td>-0.00380**</td>
</tr>
<tr>
<td></td>
<td>(0.00169)</td>
<td>(0.00173)</td>
<td>(0.00174)</td>
<td>(0.00174)</td>
<td>(0.00170)</td>
<td>(0.00174)</td>
<td>(0.00177)</td>
<td>(0.00175)</td>
<td>(0.00175)</td>
</tr>
<tr>
<td>TS</td>
<td>0.00289*</td>
<td>0.00264</td>
<td>0.00209</td>
<td>0.00298*</td>
<td>0.00260</td>
<td>0.00209</td>
<td>0.00293*</td>
<td>0.00264</td>
<td>0.00209</td>
</tr>
<tr>
<td></td>
<td>(0.00175)</td>
<td>(0.00176)</td>
<td>(0.00190)</td>
<td>(0.00174)</td>
<td>(0.00175)</td>
<td>(0.00190)</td>
<td>(0.00175)</td>
<td>(0.00177)</td>
<td>(0.00191)</td>
</tr>
<tr>
<td>MOM</td>
<td>0.00424</td>
<td>0.00787***</td>
<td>0.00864***</td>
<td>0.00830</td>
<td>0.00800***</td>
<td>0.00865***</td>
<td>0.00530</td>
<td>0.00786***</td>
<td>0.00863***</td>
</tr>
<tr>
<td></td>
<td>(0.00285)</td>
<td>(0.00308)</td>
<td>(0.00319)</td>
<td>(0.00287)</td>
<td>(0.00294)</td>
<td>(0.00312)</td>
<td>(0.00291)</td>
<td>(0.00309)</td>
<td>(0.00323)</td>
</tr>
<tr>
<td>HP</td>
<td>0.00333**</td>
<td>0.00441***</td>
<td>0.00428***</td>
<td>0.00362***</td>
<td>0.00436***</td>
<td>0.00428***</td>
<td>0.00395***</td>
<td>0.00440***</td>
<td>0.00427***</td>
</tr>
<tr>
<td></td>
<td>(0.00133)</td>
<td>(0.00138)</td>
<td>(0.00137)</td>
<td>(0.00144)</td>
<td>(0.00144)</td>
<td>(0.00142)</td>
<td>(0.00147)</td>
<td>(0.00148)</td>
<td>(0.00146)</td>
</tr>
<tr>
<td>CV</td>
<td>0.00030</td>
<td>-0.00076</td>
<td>-0.00107</td>
<td>-0.00140</td>
<td>-0.00140</td>
<td>-0.00150</td>
<td>-0.00140</td>
<td>-0.00150</td>
<td>-0.00149</td>
</tr>
<tr>
<td></td>
<td>(0.00131)</td>
<td>(0.00139)</td>
<td>(0.00149)</td>
<td>(0.00132)</td>
<td>(0.00140)</td>
<td>(0.00150)</td>
<td>(0.00132)</td>
<td>(0.00140)</td>
<td>(0.00149)</td>
</tr>
<tr>
<td>LIQUID</td>
<td>-0.00307</td>
<td>-0.00177</td>
<td>-0.00195</td>
<td>-0.00140</td>
<td>-0.00140</td>
<td>-0.00150</td>
<td>-0.00140</td>
<td>-0.00150</td>
<td>-0.00149</td>
</tr>
<tr>
<td></td>
<td>(0.00282)</td>
<td>(0.00287)</td>
<td>(0.00290)</td>
<td>(0.00288)</td>
<td>(0.00293)</td>
<td>(0.00297)</td>
<td>(0.00288)</td>
<td>(0.00293)</td>
<td>(0.00290)</td>
</tr>
<tr>
<td>SK</td>
<td>0.00534</td>
<td>0.00529</td>
<td>0.00534</td>
<td>0.00558</td>
<td>0.00541</td>
<td>0.00560</td>
<td>0.00558</td>
<td>0.00541</td>
<td>0.00560</td>
</tr>
<tr>
<td></td>
<td>(0.00558)</td>
<td>(0.00595)</td>
<td>(0.00575)</td>
<td>(0.00558)</td>
<td>(0.00541)</td>
<td>(0.00560)</td>
<td>(0.00558)</td>
<td>(0.00541)</td>
<td>(0.00560)</td>
</tr>
<tr>
<td>FAMA_FRENCH</td>
<td>(p(\hat{\beta}<em>{RNSK}, \hat{\beta}</em>{SK}))</td>
<td>-0.61053</td>
<td>-0.70244</td>
<td>-0.61053</td>
<td>-0.61053</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adj. R²</td>
<td>0.17210</td>
<td>0.21580</td>
<td>0.23848</td>
<td>0.22060</td>
<td>0.26645</td>
<td>0.28666</td>
<td>0.19944</td>
<td>0.24602</td>
<td>0.27193</td>
</tr>
</tbody>
</table>

*** p < 0.01, ** p < 0.05, * p < 0.1

Fama-Macbeth two step regression are applied for 20 models from left to right marked as (1), ..., (22) based on weekly observation. Risk factors that included in this table are listed on the first column. Similar to time-series analysis, except for the first column listing all independent variable names, from second column afterwards, odd column is for target model while even column is for robustness check with extra consideration on SK factor. All standard errors are listed underneath estimated coefficient with round bracket for normal standard error, squared bracket for adjusted error by autocorrelation and heterogeneity with lag 12 based on [Newey and West, 1987] and curly bracket for adjusted error by error in variable (EIV) problem proposed by [Shanken, 1992]. FAMA_FRENCH row stands for whether Fama-French five factors are taken into controlled in cross-sectional regression \(Y = \text{Yes, Blank = No}\). \(p(\hat{\beta}_{RNSK}, \hat{\beta}_{SK})\) row reports correlation coefficient between estimated betas: RNSK and SK in the first time series step of Fama-Macbeth. The last row in table reports adjusted \(R^2\) for each regression.
Regarding the results from other factors, HP and MOM, in general, are significantly positively priced across all scenarios. A positive significant evidence is witnessed on TS factor, while this effect disappears when RNSK is well controlled. For MOM to work, this might due to increasing market trading trend following the recent financial crisis as assets are marked as high correlation in this period. Meanwhile, strong heterogeneous beliefs about oil is well formed when market is
under extreme uncertainty, which supports HP pricing effect somehow.

As for explaining the TS factor pricing effect, after financial crisis, there might be a sluggish reaction on the physical production side of commodity market, yielding an increasing and stacked inventory. Argued by Fama and French (1988) who state that spot and future price will vary similarly when underlying inventory is high, deviating from what theory of storage suggests. This might be one potential explanation for TS factor as the signal implied from spot and future price relation may suggest the loss of the cost of carry and convenience yield premium when inventory is high.

Other factors from the commodity literature do not show significant effect in cross-sectional regression even though regression fittings $R^2$ are somehow increased. Whether controlling the stock market implied factors does not make a difference from table 2.9. To sum up, RNSK has strong conditional pricing ability across all scenarios and its effect can be not ignored.

### 2.7 Conclusion

This chapter investigates the relation between the Risk-Neutral Skewness and commodity futures return in both time-series and cross-sectional approaches. Comparison between the Risk-Neutral Skewness and the realised Pearson skewness suggested by Fernandez-Perez et al. (2018) is conducted along all analysis parts. Evidence shows that Risk-Neutral Skewness is positively related to the subsequent futures contract return. A portfolio formed by longing commodities with the high Risk-Neutral Skewness and shorting commodities with the low Risk-Neutral Skewness will generate significant positive return, which can be not explained by the traditional factors. Strategy abnormal return is robust in all cases where the baseline model and other commodity risk factors suggested from the commodity literature are controlled.

On practical view, the Risk-Neutral Skewness is superior to the realised skewness in terms of
nearly all portfolio performance statistics. Moreover, this new factor is easier to calculate without
special care on the past window length selection and is well controlled without bias, which are two
aspects mostly argued on the realised skewness calculation method. From its characteristics, this new
risk factor is partly linked to backwardation theory suggested from commodity futures Basis and HP.
Far more than, the pricing mechanism behind this new factor is more centred at over-pricing (under-
pricing) idea within framework of net buying pressure (high margins requirement), heterogeneous
belief and selective hedging ideas.
Chapter 3

Term Structure Mean Change Detection

Seeing the recent commodity futures term structure modelling popularity and motivated by the recent yield curve modelling change detection research findings, a new functional change detection procedure (via Kolmogorov-Smirnov functional form) is proposed and investigated under the alternative. Monte Carlo simulation shows decent testing power and converging rejection probability when the change size is large enough. Consistent with the literature, change detection performance is subjected to the inclusion of break points (controlling the heterogeneous effect outside of model) in the error curve volatility process. The empirical finite sample analysis tested on the gold and oil futures market suggest different change locations which are mainly correspondent to financial crisis. Multivariate forecasting regression shows that trading behaviour in the gold market after change points is captured and controlled by traditional Producers while oil market is attributed to Hedgers and Speculators’ trading activity.
CHAPTER 3. TERM STRUCTURE MEAN CHANGE DETECTION

3.1 Introduction

Motivated by the recent increasing popularity of commodity futures price term structure modelling and forecasting studies via the factor model, specifically the Nelson-Siegel model (NS) (Nelson and Siegel, 1987) and the Dynamic Nelson-Siegel model (DNS) (Diebold and Rudebusch, 2013), a natural question comes to mind is that whether this specific factor model can explain term structure variation in the commodity futures market. One argument for this is on model validation as these two models mentioned above are initialised for dealing with yield curve data in rates products rather than futures market.

The yield curve and futures prices are somehow comparable in the view of similar data structure, namely term structure (for each daily observation, different time-to-maturity contracts can be recorded). However, futures market, especially commodity futures market has been exposed to more heterogeneous characters’ effects (supply and demand shocks from Routledge et al. (2000), seasonality impact from Sørensen (2002), speculating and hedging activities from De Roon et al. (2000)). Mapping NS/DNS model directly to futures data modelling needs more considerate concerns and tests.

Furthermore, with the reference to the yield curve studies, recent results have been focusing on data change detection, modelling and forecasting with extensions on the pure DNS factor model (e.g. adaptive method with including structural change for better forecasting by Chen and Niu (2014), better and robust forecasting recorded under the regime-switching model with reversible jump Markov Chain Monte Carlo method by Xiang and Zhu (2013) and its extension with macro variables by Zhu and Rahman (2015), Hidden Markov Chain estimation method with parameter change by Nieh et al. (2010)).

Summarising the above research outcomes indicates that model stability becomes a sever concern
when applying DNS model on yield curve data. It is smooth to link this concern back to the DNS application on futures market and then argue that model validation and further extensions on a pure DNS model are necessary.

Some relevant futures market DNS modelling studies are listed in the following: Grønborg and Lunde (2016) obtain better forecasting results by including the GARCH model with Normal Inverse Gaussian innovation under the copula framework. Barunik and Malinska (2016) apply the artificial neural networks (ANN) on modelling DNS suggested beta coefficients and document its outperformance to the benchmark model in terms of different time-to-maturity contracts’ price forecasting. Karstanje et al. (2017) find the best fitting via generalising DNS model with seasonality, market and sector impact consideration.

However, none of them pay attention to the DNS model stability, which leads to the contribution of this chapter that is going to fill this gap by investigating the potential change point under the DNS model fitting of commodity futures data. For the sake of the testing model stability, this chapter proposes to use a new test statistics and procedure under the framework of functional data analysis.

Before introducing on the functional method, at this writing moment, we do aware that there is a large literature group discussing time series structure change detection method. Hypothesis tests have been set up for detecting a unknown break location, for some study about it, readers are referred to some theories and finance applications by Bai and Perron (1998), Chu et al. (1996), Davis et al. (2006), Csörgő and Horváth (1997), Aue and Horváth (2013) as well as other literatures cited inside and referred on it.). But, all of them studies are limited on one single time series detection, which is not suitable for this chapter as what this chapter care about is to explore detection procedure on a collection of time series (per maturity) data representing an entire term structure.

Presenting and analysing data in a functional way has been experiencing an increasing trend over
the decade. This is attributed to the special property of functional modelling of handling data that is out of the explanation framework of either scalar or vector observation. Its natural and parsimonious way of describing data has been widely used in sparse data smoothing and modelling, statistical inference and forecasting (Ramsay, 2006; Horváth and Kokoszka, 2012; Berkes et al., 2009; Kargin and Onatski, 2008; and Kokoszka and Reimherr, 2013).

The specific statistical testing procedure used in this chapter is inspired by the recent new mean change detection procedure by Bardsley et al. (2017) in which functional statistics is introduced with a careful investigation on large sample asymptotic properties. Standing on the same framework, this chapter extends their current findings by exploring a new functional statistics’ (Kolmogorov-Smirnov functional) large sample property under an alternative hypothesis and discussing its testing power further. This chapter further includes an empirical data analysis to uncover the economic intuition behind the mean changes. In a straightforward way, the regression analysis is employed to analyse the commodity futures subsample market data divided by change point location.

With respect to the change point detection, the simple scalar, single time series as well as multivariate series have been studied for the past several decades starting from the inspection control CUSUM test by Page (1954). More recent extensions on linear, panel data and functional data can be referred to survey papers by Aue and Horváth (2013) and Horváth and Rice (2014). Although changes can happen in both the first order and the second order, this chapter focuses on the first order (mean) change only (see (Csörgő and Horváth, 1997)) as involving the second order property cause complex asymptotic theory even in the simple independent normal case (Horváth, 1993).

In the line of functional change point detection, the sample difference change point test is proposed (see, change point in the functional linear operators in functional linear model by Horváth et al. (2009), functional principal component analysis based change detection via bootstrap by Benko et al.
CHAPTER 3. TERM STRUCTURE MEAN CHANGE DETECTION

(2009), general functional observation mean change detection by Horváth et al. (2013). Functional projection method is another branch of this testing procedure. A general functional data generating process change detection proposed by Berkes et al. (2009) and Aue, Gabrys, Horváth and Kokoszka (2009) under independent innovations assumption, under weekly dependent innovations assumption by Hörmann et al. (2010). However, all of these statistics forms are in weighted form with penalisation on the corresponding eigenvalues.

Following Bardsley et al. (2017), assumptions for the innovation terms dependence structure is also adopted in this chapter, which shows a more general fitting to real data (Aue, Hörmann, Horváth, Reimherr et al., 2009; Hörmann et al., 2013; Kokoszka and Reimherr, 2013 and Aue et al., 2012). For comprehensive summary on the functional data analysis, relevant survey paper by Horváth and Rice (2015) is well recommended.

In terms of the selection of pre-determined factors (used to project infinite dimensional functional curve data onto finite vectors), well-defined orthonormal basis function factors (e.g. Fourier and Spline basis) are normally used due to their well-known function properties. However, this automatically leads to one question that how to interpret the meaning behind the change point with respect to those selected functional factors. Recent studies from Chib and Kang (2012) who adopt macroeconomic factors and Bardsley et al. (2017) who use the Dynamic Nelson-Siegel model (Diebold and Rudebusch, 2013) suggested factors give a more interpretable idea behind this change. Implied by data structure, a new functional detection statistic formed in this chapter is based on DNS implied factors.

Regarding the strength of these factors, the DNS three factors (“Level”, “Slope” and “Curvature”) have been argued with respect to its validity (e.g. Lengwiler and Lenz (2010) propose new factors called intelligible factors, stating better interpretation and Hays et al. (2012) model DNS beta
coefficients in a functional dynamic way). Despite the potential disadvantages of DNS model factors mentioned above, the fundamental reasons for using DNS in this chapter is mainly due to its wider documentation in the literature (from the initial fixed factor approach see [Nelson and Siegel (1987)], to time-varying approach by [Diebold and Li (2006) and Diebold and Rudebusch (2013)] and also other cited studies in their reference and based on them). Another two concerns of testing DNS is (1) for the sake of comparison with reference studies by [Bardsley et al. (2017)], (2) easy interpretation and implementation from the real data perspective.

On the page of the DNS, lambda ($\lambda$) a parameter in model, defined as to measuring the maturity decaying effect, is introduced with time variation property rather than being constant by [Hautsch and Ou (2012) and Koopman et al. (2010)]. However, the fitting and forecasting results suggest that no significance improvement is observed when taking time varying lambda into consideration. This chapter will ignore this time-varying lambda idea and keep lambda fixed.

In addition to this, it is worth mentioning that, the possibilities of change point in lambda may cause functional detection procedure bias in the time dimension. If this is the case, the DNS model factors suggested from this scenario (where lambda value is not stable) will not be proper to use in the following testing procedure. To account for this problem, robust check is conducted with projecting data onto basis function, denoted as fully functional method in [Bardsley et al. (2017)].

Regarding the main contributions to the commodity futures literature, up to now, there is no any functional change point detection method discussion formed on Kolmogorov-Smirnov functional, any empirical functional change detection application on global commodity futures market data and further regression analysis on subsamples to quantify structure change driving force. To sum up, this chapter is going to achieve the following points in the end. Firstly, exploring new functional change point detection statistics with clear document on both large sample property and testing power
performance. Secondly, empirically providing evidence on data structure mean change of DNS model fitting via real time commodity futures data. Thirdly, the market driving force behind the mean change of DNS method is evidenced and explained.

The structure of this chapter is organized as follows: section 2 is the methodology description on how the Dynamic Nelson-Siegel model is formed, section 3 is functional models and projected method specification, section 4 shows simulation implementation details; section 5 refers to the empirical data description and analysis, section 6 concludes all findings on the simulation and empirical data analysis. Theoretical proofs in section 3 are presented in the Appendix B.

### 3.2 Dynamic Nelson-Siegel Model Estimation

Capturing the term structure price movement on different maturities bases, the DNS model shows different variations given different parameters and factor forms. Apart from the time-varying effect, a decaying factor, $\lambda$, allocates DNS modelling emphasise by placing different weights on different maturities’ contracts. In practical way, estimated optimal lambda is chosen to fix the shape of three factors in DNS. As this estimation process, grid search and optimization are adopted following studies in government bond yield curve application by (Diebold and Li, 2006). DNS is formulated as follows:

$$X_t(\tau) = \beta_{t,1} + \beta_{t,2} \left( \frac{1-e^{-\lambda \tau}}{\lambda \tau} \right) + \beta_{t,3} \left( \frac{1-e^{-\lambda \tau}}{\lambda \tau} - e^{-\lambda \tau} \right) + \epsilon_t(\tau) \quad (3.2.1)$$

where, subscript $t$ is for time dimension, $\tau$ is for maturity, $X_t(\tau)$ is data observation with respect to time-to-maturity $\tau$ observed at time $t$, $\beta_{t,1}$ is the first factor with a constant corresponding factor loading which is independent of time and $\lambda$ value, measuring the long-time variation of term structure. $\beta_{t,2}$ is the second factor with factor loading $\left( \frac{1-e^{-\lambda \tau}}{\lambda \tau} \right)$ in DNS model gauging the fitting of short time
contract data variation, $\beta_{t,3}$ is the third factor with factor loading $\left(\frac{1-e^{-\lambda t}}{\lambda t} - e^{-\lambda t}\right)$ measuring the middle term of contract data variation.

In the literature of the commodity futures market, the DNS three factors are not only the main components explaining the dynamics of term structure variation (e.g. PCA forecasting method Chantziara and Skiadopoulos (2008)) but also delegating economical meaning, denoted as "Level" (the average price level shift across all time-to-maturity contracts), "Slope" (presenting the relative strength between the short and long term price change, indicating market situation on backwardation or contango and forming an important factor called basis term structure Fernandez-Perez et al. (2017b)) and "Curvature" (controlling middle time-to-maturity contract, showing expectation shift speed in commodity futures market) respectively.

More importantly, $\lambda$ in this formula adjusts the importance within last two factors. In the yield curve literature Diebold and Li (2006) and Diebold and Rudebusch (2013) optimize the $\lambda$ estimation based on maximizing middle term fitting. While this differs in futures market GrØnborg and Lunde (2016) and Barunik and Malinska (2016), in which lambda is carried out by standard error square minimization. For the sake of demonstrating how sensitive and responsive of factor shape to different values of lambda, figure 3.1 shows four examples about sensitivities of 'Slope' and "Curvature" factors when decaying parameters are changed while "Level" factor stay constant for all cases.

It is clear to state that the larger the $\lambda$ value, the higher the factor will be placed on the short-term maturity contract. When the $\lambda$ value is equal to 0.1, factor value over 0.5 are put on the contract with maturity less than around 12\footnote{which can be scaled in reality, here the magnitude of 12 is for example} in x-axis figure 3.1 with second factor declining quickly and third factor peak at this point.

In this chapter, grid search method is applied to find out the best lambda value (in terms of minimizing the sum of squared error over time) for fixing the shape of the DNS three factors. The first step
Figure 3.1: Sensitivity of Dynamic Nelson-Siegel Model to Lambda Values

Note: this figure shows how shapes of three factors change in response to the change of lambda values. Factor loadings are plotted with red solid line for “Level” factor, green dot line for “Slope” factor and blue dash line for “Curvature” line from top to bottom, left to right, in terms of lambda value of 0.01, 0.05, 0.1 and 0.3.

is composed of a fine grid search. During each search, one $\lambda$ is selected and assigned to the cross-sectional linear regression method. Given the estimation results from the first step, the best lambda will be extracted out in terms of the minimum value of squared error summation in the previous step and its corresponding model coefficients (three latent time-series factors) are obtained in a time-series manner (OLS method).
3.3 Functional Models and Projected Method

Following the functional change point test framework in [Bardsley et al. (2017)], the functional observation $X_t(i)$ is modelled by the curves

$$X_t(i) = \sum_{k=1}^{K} B_{t,k} f_k(i) + \epsilon_t(i)$$  \hspace{1cm} (3.3.1)

where, $X_t(i)$ is $t^{th}$ functional observation, in this chapter, referred to time dimension, $f_k(i)$ is known $k$ dimensional factors. $K$ is the total number of functions used for the linear representation of the mean of $X_t(i)$. On the framework of DNS model, $K = 3$ and the function $f_k(i)$ can be then referred to "Level" ($k=1$), "Slope" ($k=2$) and "Curvature" ($k=3$). The functions $f_k(i)$ are assumed to be linearly independent, but the asymptotic results are not affected if this assumption does not hold. The random coefficient $B_{t,k}$ can be written as:

$$B_{t,k} = \mu_{t,k} + b_{t,k}, \quad E[b_{t,k}] = 0$$  \hspace{1cm} (3.3.2)

The hypothesis of a constant functional mean is stated as:

$$H_0 : \quad \mu_1 = \mu_2 = \cdots = \mu_N$$  \hspace{1cm} (3.3.3)

where

$$\mu_t = [\mu_{t,1}, \mu_{t,2}, \cdots, \mu_{t,K}]^T$$

Under the alternative hypothesis, let $r_1 < r_2 < \cdots < r_R$ denote the time of change in the functional mean, in which $r_0 = 0$ and $r_{R+1} = N$, therefore the number of change point in the functional mean is
R and the means after changes are $\mu^*_1, \mu^*_2, \ldots, \mu^*_{R+1}$, see

$$H_A: \quad \mu_t = \mu^*_m, \quad \text{if} \quad r_{m-1} < t \leq r_m, \quad 1 \leq t \leq N$$

(3.3.4)

Under the null hypothesis, the mean is assumed to be unchanged while error structure of $(\sum_{k=1}^{K} b_{t,k} f_{k}(i) + \varepsilon_t(i))$ is allowed to change at certain time point, called break point, say $t_m$:

$$1 = t_0 < t_1 < t_2 < t_3 < \cdots < t_M < t_{M+1} = N,$$

The data are second order stationary on the interval $(t_\ell, t_{\ell+1}]$ with $1 \leq \ell \leq M$ and the purpose of above statement is to detect the mean change regardless of the second order (variance and covariance) behaviour of error curve. Intuitively, break point is an idea stating possible exogenous information inclusion in the data generating process, which could be government intervention impact on data like yield curve, spot market demand and supply temporary shock from market or OPEC announcement for commodity futures market and oil related future products.

For example, it is natural to choose the point as the day when financial market reflects a strong signal like Lehman Brother Bankruptcy in 2008. Statistical procedure of testing the constant mean change normally requires estimation of long run covariance matrix. However, due to possible breaks in the covariance structure, estimation procedure without this consideration might result in misleading outcome.

Following the framework in Bardsley et al. (2017), the test statistics are obtained as functions of the CUSUM process of projections:

$$\alpha_N(x) = N^{-\frac{1}{2}} \left( \frac{[N\alpha]}{N} \sum_{t=1}^{N} z_t - \frac{[N\alpha]}{N} \sum_{t=1}^{N} z_t \right), \quad 0 \leq x \leq 1,$$

(3.3.5)
where, $\lfloor \cdot \rfloor$ denotes the integer value and $\langle \cdot , \cdot \rangle$ is the inner product,

$$z_t = \langle X_t, f_1 \rangle, \ldots, \langle X_t, f_K \rangle, \quad (3.3.6)$$

Rather than using the Cramer-von-Mises functional of $\alpha_N(x)$, Kolmogorov-Smirnov statistics is employed.

$$K_N = \sup_{0 \leq x \leq 1} \| a_N(x) \|, \quad (3.3.7)$$

where, $\| \cdot \|$ stands for the norm in the $\mathbb{R}^K$. According to Theorem 3.1 in Bardsley et al. (2017), under $H_0$,

$$\alpha_N(x) \rightarrow G^0(x), \quad \text{in} \quad D^K([0,1])$$

where the process $G^0$ is defined by

$$G^0(x) = G(x) - xG(1), \quad 0 \leq x \leq 1, \quad (3.3.8)$$

and $G(x)$, $x \in [0,1]$, is a mean zero $\mathbb{R}^K$-valued Gaussian process with covariance:

$$E \left( G(x) G(y)^T \right) = \sum_{j=1}^m (\theta_j - \theta_{j-1}) V_j + (x - \theta_m) V_{m+1}, \quad \theta_m \leq x \leq \theta_{m+1}, \quad y \geq x,$$

$V$ is a long-run matrix given by:

$$V_m = \sum_{l=-\infty}^{\infty} \text{Cov} \left( w(a_i^m), w(a^m_{i+1}) \right),$$

where, $w$ is a known function, $a_i^m$ is defined in Assumption 1 and 2 of Bardsley et al. (2017).

It is followed immediately from the equation (3.3.8) that,

$$K_N \xrightarrow{D} \sup_{0 \leq x \leq 1} \| G^0(x) \|, \quad (3.3.9)$$
and based on the equation (3.3.9), the functional statistic can be approximated with the supreme of norm of $G^0$. The Gaussian process $G^0(x)$ can be represented as an infinite sum using the Karhunen-Loève expansion, namely,

$$G^0(x) = \sum_{j=1}^{\infty} \lambda_j^{1/2} Z_j \phi_j(x),$$

therefore,

$$\sup_{0 \leq x \leq 1} \|G^0(x)\| = \sup_{0 \leq x \leq 1} \| \sum_{j=1}^{\infty} \lambda_j^{1/2} Z_j \phi_j(x) \|,$$

(3.3.10)

where, $j$ denotes $j^{th}$ observation, $\lambda_1 \geq \lambda_2 \geq \ldots$, are the eigenvalues of covariance kernel $R(x,y) = E(G^0(x)G^0(y)^T)$, $\phi_j(x)$ is eigenfunctions of $R$ and $Z_1, Z_2, \ldots$ are independent standard normal random variables. The eigenfunctions are orthonormal and satisfy:

$$\int_0^1 R(x,y) \phi_j(y) dy = \lambda_j \phi_j(x), \quad j = 1, 2, \ldots$$

(3.3.11)

It is shown in the [Bardsley et al. (2017)] that $R(x,y)$ is unknown and must be estimated from the sample. They suggest that $\hat{R}(x,y)$, a kernel estimated from the sample satisfying the condition,

$$\int_0^1 \int_0^1 \| \hat{R}(x,y) - R(x,y) \| \overset{P}{\to} 0$$

Defining the empirical eigenvalues $\hat{\lambda}_i$ and eigenfunction $\hat{\phi}_i(x)$ by,

$$\int_0^1 \hat{R}(x,y) \hat{\phi}_j(y) dy = \hat{\lambda}_j \hat{\phi}_j(x), \quad j = 1, 2, \ldots,$$

and approximate $G^0(x)$ by,

$$G^0(x) = \sum_{j=1}^{\infty} \hat{\lambda}_j^{1/2} Z_j \hat{\phi}_j(x),$$

where, $Z_1, Z_2, \ldots$ are independent standard normal random variables.
In the empirical and simulation, calculation of \( R(x,y) \) is replaced by computing covariance of \( G^0(x) \), which is as follows,

\[
R(x,y) = E \left( G^0(x)G^0(y)^T \right) \\
= (1 - y) \left[ \sum_{j=1}^{m} (\theta_j - \theta_{j-1})V_j + (x - \theta_{m})V_{m+1} \right] \\
- x \left[ \sum_{j=1}^{m'} (\theta_j - \theta_{j-1})V_j + (x - \theta_{m'})V_{m'+1} \right] \\
+ xy \sum_{j=1}^{M+1} (\theta_j - \theta_{j-1})V_j, \quad (3.3.12)
\]

where \( 0 \leq x \leq y \leq 1, \theta_m \leq x \leq \theta_{m+1}, \) and \( \theta_{m'} \leq y \leq \theta_{m'+1}. \)

To state the results under the alternative, a matrix \( C \) is then introduced,

\[
C = \{ < f_i, f_j >, \quad 1 \leq i, j \leq k \}, \quad (3.3.13)
\]

since \( f_k \) for \( k = 1, 2, \ldots, K \) are assumed to be linearly independent, stating that \( C^{-1} \) exists, \( z_t \) is now written as,

\[
z_t = C\mu_t + \gamma_t, \quad (3.3.14)
\]

where, \( \gamma_t \) is the error term.

Using the definition of \( z_t \) and get,

\[
E(\ z_t \ ) = c\mu^*_t, \quad \text{if} \quad r_{m-1} < t \leq r_m, \quad 1 \leq t \leq N \quad (3.3.15)
\]
and then the assumption of change points in the mean occurring at \( r_1 < r_2 < r_3 < \cdots < r_K \) means that,

\[
\lim_{N \to \infty} N^{-1/2} \delta_N (t, m) = \infty, \quad t < m, \quad 1 \leq t < m < R
\]  \hspace{1cm} (3.3.16)

where,

\[
\delta_N (t, m) = \max_{t \leq k \leq m} \left| \sum_{\ell=t}^{k} \mu_{t}^{\ast} (r_{\ell} - r_{\ell-1}) - \frac{r_k - r_{t-1}}{r_m - r_{t-1}} \sum_{\ell=t}^{m} \mu_{t}^{\ast} (r_{\ell} - r_{\ell-1}) \right|
\]  \hspace{1cm} (3.3.17)

the equation (3.3.17) indicates that there is at least one detectable mean change in the subset \( \{ X_{r_{t-1}} (i), \ldots, X_{r_m} (i) \} \).

It is easy to see that \( \delta_N (t, m) \) compares the means of \( z_t, z_{t+1}, \ldots, z_k, t \leq k \leq m \) to the fraction of the total sum of \( z_t, z_{t+1}, \ldots, z_m \). If assuming that locations of these changes are proportional to the sample size \( N \),

\[
r_t = \lfloor N \tau_t \rfloor, \quad 1 \leq t \leq R,
\]

\[
0 = \tau_0 < \tau_1 < \tau_2 < \cdots < \tau_R < \tau_{R+1} = 1,
\]

then equation (3.3.16) and (3.3.17) is rewritten,

\[
\lim_{N \to \infty} N^{1/2} \delta_N (t, m) = \infty, \quad t < m, \quad 1 \leq t < m < R
\]  \hspace{1cm} (3.3.18)

where,

\[
\delta_N (t, m) = \max_{t \leq k \leq m} \left| \sum_{\ell=t}^{k} \mu_{t}^{\ast} (\tau_{\ell} - \tau_{\ell-1}) - \frac{\tau_k - \tau_{t-1}}{\tau_m - \tau_{t-1}} \sum_{\ell=t}^{m} \mu_{t}^{\ast} (\tau_{\ell} - \tau_{\ell-1}) \right|
\]  \hspace{1cm} (3.3.19)

Let

\[
\alpha_N^{\ast} (x) = N^{-1/2} \left( \sum_{l=1}^{[N]} (z_l - E (z)) \frac{[N]}{N} \sum_{l=1}^{N} (z_l - E (z)) \right),
\]  \hspace{1cm} (3.3.20)

under the null hypothesis, \( \alpha_N^{\ast} (x) = \alpha_N (x) \) and assumptions of Theorem 3.1 of Bardsley et al. (2017).
are satisfied, then will have,

\[ \sup_{0 \leq x \leq 1} \| \alpha_N(x) \| = O_p(1) \]  

(3.3.21)

**THEOREM 3.1** If conditions from the equations (3.3.18), (3.3.19) and (3.3.20) are satisfied, then,

\[ \sup_{0 \leq x \leq 1} \| \alpha_N(x) \| \xrightarrow{p} \infty, \]  

(3.3.22)

This implies the consistency of the testing procedure for the validity of \( H_0 \). \( H_0 \) is rejected if \( K_N \geq \hat{k}_N(\alpha) \), where \( 1 - \alpha \) is the significance and \( \hat{k}_N(\alpha) \) is the empirical significance level computed from the data. If \( H_0 \) holds, then

\[ \lim_{N \to \infty} P\{ K_N \geq \hat{k}_N(\alpha) \} = \alpha, \]  

(3.3.23)

while under the change point alternative,

\[ \lim_{N \to \infty} P\{ K_N \geq \hat{k}_N(\alpha) \} = 1 \]  

(3.3.24)

In the testing procedure, binary segmentation method is used to divide the data into mean stationary subsets. Let,

\[ K_N(t, m) = (m - t)^{-1/2} \sup_{0 \leq t \leq 1} \left\| \sum_{\ell=t+1}^{[m-t]x+t} z_\ell - \frac{(m-t)x}{(m-t)} \sum_{\ell=t+1}^{i} z_\ell \right\| \]  

(3.3.25)

The statistics \( K_N(t, m) \) is computed from the observations \( X_{t+1}(i), \ldots, X_m(i) \) only. The estimator of a change in the mean of the observations \( X_{t+1}(i), \ldots, X_m(i) \) is defined by,

\[ \hat{r}(t, m) = \min \left\{ k \in \{ t + 1, \ldots, m \} : (m-t)^{-1/2} K_N(t, m) = (m-t)^{-1/2} \left\| \sum_{\ell=t+1}^{k-1} z_\ell - \frac{k-i}{m-t} \sum_{\ell=t+1}^{m} z_\ell \right\| \right\} \]  

(3.3.26)

According to the binary segmentation method, \( H_0 \) is rejected and find \( \hat{r}_1 = \hat{r}(0, N) \) if \( K_N(0, N) \geq \hat{k}_N(\alpha) \)
\( \hat{k}_N(0,N,\alpha) \) \(^{33}\) Then it tests whether \( K_N(0,\hat{r}_1) \) and \( K_N(\hat{r}_1,N) \) are above their corresponding critical values. If \( K_N(0,\hat{r}_1) \) is above its critical value, \( \hat{r}_2 = \hat{r}(0,\hat{r}_1) \) is computed and sample \( X_1, \ldots, X_{\hat{r}_1} \) is divide into subsets \( X_1, \ldots, X_{\hat{r}_2} \) and \( X_{\hat{r}_2}, \ldots, X_{\hat{r}_1} \), otherwise \( X_1, \ldots, X_{\hat{r}_1} \) is a mean stationary subset. This same procedure is applied for \( K_N(\hat{r}_1,N) \). By the binary algorithm, sets are separated into \( \hat{R} + 1 \) mean stationary subsequence and corresponding change points found is \( \hat{R} \).

**THEOREM 3.2** If conditions (3.3.16), (3.3.17) and (3.3.21) are satisfied, then

\[
\lim_{N \to \infty} P\{\hat{R} \geq R\} = 1 \quad (3.3.27)
\]

and for all \( k \geq 1 \),

\[
P\{\hat{R} = R + k\} \leq \alpha^k, \quad k = 1, 2, \ldots \quad (3.3.28)
\]

The algorithm finds that the change points \( \hat{r}_1^*, \hat{r}_2^*, \ldots, \hat{r}_R^* \), the size of each change is denoted as,

\[
\hat{\delta}_{N,t} = \left\| \frac{\hat{r}_t}{N} - \frac{\hat{r}_t}{N} \sum_{\ell=1}^{N} z_\ell \right\| \quad (3.3.29)
\]

**THEOREM 3.3** If conditions of the equations (3.3.16), (3.3.17) and (3.3.21) are satisfied, then for any \( \varepsilon > 0 \),

\[
\lim_{N \to \infty} P\left\{ \frac{\hat{r}_t^*}{N} \in \bigcup_{\ell=1}^{M} [\tau_\ell - \varepsilon, \tau_\ell + \varepsilon] \right\} = 1, \quad for \ all \ 1 \leq t \leq M \quad (3.3.30)
\]

According to the **THEOREM 3.3** it is easier to locate the sample where the changes in the mean occur. By the equation (3.3.16) or (3.3.18), the largest change size \( \hat{\delta}_{N,t} \) must be much larger than the remaining ones, if there is any, ”artificial change points” that are occurring with probability less than

\(^{33}\)Empirical critical threshold value given sample starting from 0 to N with significance \( 1 - \alpha \)
CHAPTER 3. TERM STRUCTURE MEAN CHANGE DETECTION

α and corresponding size $\delta_{N,t}$'s are much smaller.

3.4 Simulation Study

In this section, Kolmogorov-Smirnov functional statistics proposed in above section is tested under both the null and alternative hypothesis. Functional coefficient processes are generated with functional mean observation including change and excluding change. Monte Carlo method with replication 1000 times is deployed. In the interest of performance comparison from functional statistics (Cramer-von-Mises functional from Bardsley et al. (2017)), simulation settings are the same as they did. Then, sensitivity analysis for Kolmogorov-Smirnov functional statistics with respect to the change magnitude is also conducted.

Real change point location and observed mean values before and after change location are settled in the following rules,

Mean under the Null $H_0$:

$$\mu = (4.54, -2.82, -3.03)$$

Mean under Change $H_{A(1)}$:

$$\mu_t = \begin{cases} 
4.54 \\
-2.82 \\
-3.03 
\end{cases}, \quad t \leq \frac{N}{2}, \quad \mu_t = \begin{cases} 
4.2 \\
-3.0 \\
-3.2 
\end{cases}, \quad t > \frac{N}{2}$$

Mean under Change $H_{A(2)}$:
Mean under Change $H_A(3)$:

$$
\mu_t = \begin{pmatrix} 4.54 \\ -2.82 \\ -3.03 \end{pmatrix}, \quad t \leq \frac{N}{2}; \quad \mu_t = \begin{pmatrix} 3.89 \\ -3.32 \end{pmatrix}, \quad t > \frac{N}{2}
$$

Under the framework of DNS model, functional coefficients $B_{t,k}$ are generated following AR(1) process:

$$B_{t,k} = \mu_k(1 - \phi_k) + \phi_k B_{t-1,k} + u_{t,k}, \quad t = 1, \ldots, N, \quad k = 1, 2, 3,$$

and the error curve $\varepsilon_t(i)$ is simulated from

$$\varepsilon_t(i) = \frac{2}{25} \zeta_i + \frac{1}{25} \zeta_i \sin(2\pi i), \quad i \in [0, 1],$$

where, $\zeta_{t,k} = 0.9 \zeta_{t-1,m} + Z_{t,m}$ and $Z_{t,m} \sim N(0, 1), m = 1, 2, t = 1, \ldots, N$, AR(1) parameters are fixed at: $\phi_1 = 0.9, \phi_2 = 0.9, \phi_3 = 0.9$.

Once the simulated functional coefficient $B_{t,k}$ is obtained, functional curve observation is constructed by multiplying it with the DNS factors pre-determined before. Therefore, vector of projection $z_t$ in $\alpha_N(x)$ is obtained and related CUSUM test is formed. In order to approximate the right-hand side of the equation (3.3.10), one supreme value is computed given 10000 sampling random variables from the standard normal distribution. Following the same procedure, repeating this 1000 times to obtain its distribution.
Heterogeneous information is also taken into account and represented by the change of error curve variance at certain time point \( t \), denoted as Break Point (BP). By doing this, the intuition is to allow functional curve observation dynamics containing information outside system that may be responded to the real change or may simply data process variation. In this simulation, BP is chosen in the functional error curve variance at either the point which is exactly the same as the real functional mean change (BP = \( \frac{N}{2} \)) or the one which is not (BP \( \neq \frac{N}{2} \)).

It is worth mentioning that break point (BP) is the consideration on variance change while the change point (CP) is consideration on mean change. In the simulation, we insert the BP into data generation process by allowing the error structure variance to change at a certain time. However, in real data testing procedure, BP is assumed to be included in its error process (variance becomes a new value after certain time point), therefore no need to specify it.

Simulated AR(1) error heterogeneous variance:

\[
\begin{align*}
\Var[u_t, 1] &= 0.003, \quad t \leq t_1 \text{ and } \Var[u_t, 1] = 0.012, \quad t > t_1, \\
\Var[u_t, 2] &= 0.006, \quad t \leq t_1 \text{ and } \Var[u_t, 2] = 0.026, \quad t > t_1, \\
\Var[u_t, 3] &= 0.053, \quad t \leq t_1 \text{ and } \Var[u_t, 3] = 0.095, \quad t > t_1
\end{align*}
\]

Break points are set to be equal to \( N/2, 2N/3 \) and 1, corresponding to scenarios with respect to real change points (CP): BP = CP, BP \( \neq \) CP and BP at last observation of sample (no consideration on the heterogeneous information effect). As for the simulation sample size, 250 and 500 are used to test all scenarios respectively. DNS projected method simulation power test results for sample size equal to 250 and 500 is reported in table 3.1.

From table 3.1, whether break point is included (BP = 1, standing for no break point in testing sample) does make a significant difference on the projection method testing power. For all test
## Chapter 3. Term Structure Mean Change Detection

### Table 3.1: Simulation Size and Power for DNS Projection Method

<table>
<thead>
<tr>
<th>Break Point</th>
<th>Significance Level</th>
<th>Sample Size</th>
<th>Change Point</th>
<th>( H_0 )</th>
<th>( H_1 )</th>
<th>( H_2 )</th>
<th>( H_3 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5</td>
<td>0.99</td>
<td>250</td>
<td>0.5</td>
<td>0.072</td>
<td>0.634</td>
<td>0.979</td>
<td>0.797</td>
</tr>
<tr>
<td>0.667</td>
<td>0.99</td>
<td>250</td>
<td>0.5</td>
<td>0.1</td>
<td>0.426</td>
<td>0.505</td>
<td>0.657</td>
</tr>
<tr>
<td>1</td>
<td>0.99</td>
<td>250</td>
<td>0.5</td>
<td>0.016</td>
<td>0.096</td>
<td>0.012</td>
<td>0.06</td>
</tr>
<tr>
<td>0.5</td>
<td>0.95</td>
<td>250</td>
<td>0.5</td>
<td>0.118</td>
<td>0.728</td>
<td>0.987</td>
<td>0.866</td>
</tr>
<tr>
<td>0.667</td>
<td>0.95</td>
<td>250</td>
<td>0.5</td>
<td>0.135</td>
<td>0.56</td>
<td>0.637</td>
<td>0.775</td>
</tr>
<tr>
<td>1</td>
<td>0.95</td>
<td>250</td>
<td>0.5</td>
<td>0.046</td>
<td>0.285</td>
<td>0.074</td>
<td>0.18</td>
</tr>
<tr>
<td>0.5</td>
<td>0.90</td>
<td>250</td>
<td>0.5</td>
<td>0.147</td>
<td>0.772</td>
<td>0.992</td>
<td>0.897</td>
</tr>
<tr>
<td>0.667</td>
<td>0.90</td>
<td>250</td>
<td>0.5</td>
<td>0.164</td>
<td>0.619</td>
<td>0.719</td>
<td>0.815</td>
</tr>
<tr>
<td>1</td>
<td>0.90</td>
<td>250</td>
<td>0.5</td>
<td>0.074</td>
<td>0.415</td>
<td>0.142</td>
<td>0.291</td>
</tr>
</tbody>
</table>

### Panel A: Sample Size = 250

### Panel B: Sample Size = 500

<table>
<thead>
<tr>
<th>Break Point</th>
<th>Significance Level</th>
<th>Sample Size</th>
<th>Change Point</th>
<th>( H_0 )</th>
<th>( H_1 )</th>
<th>( H_2 )</th>
<th>( H_3 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5</td>
<td>0.99</td>
<td>500</td>
<td>0.5</td>
<td>0.056</td>
<td>0.895</td>
<td>1</td>
<td>0.98</td>
</tr>
<tr>
<td>0.667</td>
<td>0.99</td>
<td>500</td>
<td>0.5</td>
<td>0.065</td>
<td>0.809</td>
<td>0.934</td>
<td>0.94</td>
</tr>
<tr>
<td>1</td>
<td>0.99</td>
<td>500</td>
<td>0.5</td>
<td>0.029</td>
<td>0.764</td>
<td>0.367</td>
<td>0.673</td>
</tr>
<tr>
<td>0.5</td>
<td>0.95</td>
<td>500</td>
<td>0.5</td>
<td>0.112</td>
<td>0.945</td>
<td>1</td>
<td>0.99</td>
</tr>
<tr>
<td>0.667</td>
<td>0.95</td>
<td>500</td>
<td>0.5</td>
<td>0.114</td>
<td>0.882</td>
<td>0.978</td>
<td>0.964</td>
</tr>
<tr>
<td>1</td>
<td>0.95</td>
<td>500</td>
<td>0.5</td>
<td>0.075</td>
<td>0.931</td>
<td>0.738</td>
<td>0.903</td>
</tr>
<tr>
<td>0.5</td>
<td>0.90</td>
<td>500</td>
<td>0.5</td>
<td>0.154</td>
<td>0.955</td>
<td>1</td>
<td>0.992</td>
</tr>
<tr>
<td>0.667</td>
<td>0.90</td>
<td>500</td>
<td>0.5</td>
<td>0.137</td>
<td>0.912</td>
<td>0.988</td>
<td>0.977</td>
</tr>
<tr>
<td>1</td>
<td>0.90</td>
<td>500</td>
<td>0.5</td>
<td>0.106</td>
<td>0.967</td>
<td>0.879</td>
<td>0.958</td>
</tr>
</tbody>
</table>

Note: this table reports the simulation test power results based on Kolmogorov-Smirnov Functional projection method with Dynamic Nelson-Siegel model factors as projected factors Sample. All scenarios are reported for all three significance levels (99%, 95% and 90%) under one null hypothesis (\( H_0 \)) and three alternative hypothesis (\( H_1, H_2 \) and \( H_3 \)) with details specified in the simulation section. In panel A, break points are settled as 0.5 (1/2 of sample size), 0.667 (2/3 of sample size) and 1 (break is at the last observation of sample, implying no variance change in data generation simulation process) given the sample size is 250. In panel B, break points are settled as 0.5 (1/2 of sample size), 0.667 (2/3 of sample size) and 1 (no break point in test sample) given the sample size is 500.

Scenarios with sample size 500, as long as break point is introduced, simulation testing power after thousands trail shows reasonable results with acceptable errors. Compared with the result from Bard-sley et al. (2017) who use Cramer-von-Mises functional statistics, Kolmogorov-Smirnov based test has relatively upward testing powers under all scenarios. Similar to their findings, there is no consistent evidence for testing power in this simulation in terms of the different alternative hypotheses.

Specifically, testing power under \( H_2 \) alternative is higher than the one under \( H_1 \) given the change size
### Table 3.2: Simulation Size and Power for Full Functional Projection Method

<table>
<thead>
<tr>
<th>Break Point</th>
<th>Significance Level</th>
<th>Sample Size</th>
<th>Change Point</th>
<th>$H_0$</th>
<th>$H_1$</th>
<th>$H_2$</th>
<th>$H_3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5</td>
<td>0.99</td>
<td>250</td>
<td>0.5</td>
<td>0.08</td>
<td>0.63</td>
<td>0.988</td>
<td>0.814</td>
</tr>
<tr>
<td>0.667</td>
<td>0.99</td>
<td>250</td>
<td>0.5</td>
<td>0.087</td>
<td>0.366</td>
<td>0.353</td>
<td>0.657</td>
</tr>
<tr>
<td>1</td>
<td>0.99</td>
<td>250</td>
<td>0.5</td>
<td>0.007</td>
<td>0.028</td>
<td>0</td>
<td>0.004</td>
</tr>
<tr>
<td>0.5</td>
<td>0.95</td>
<td>250</td>
<td>0.5</td>
<td>0.121</td>
<td>0.726</td>
<td>0.995</td>
<td>0.873</td>
</tr>
<tr>
<td>0.667</td>
<td>0.95</td>
<td>250</td>
<td>0.5</td>
<td>0.129</td>
<td>0.488</td>
<td>0.484</td>
<td>0.759</td>
</tr>
<tr>
<td>1</td>
<td>0.95</td>
<td>250</td>
<td>0.5</td>
<td>0.033</td>
<td>0.115</td>
<td>0.003</td>
<td>0.018</td>
</tr>
<tr>
<td>0.5</td>
<td>0.90</td>
<td>250</td>
<td>0.5</td>
<td>0.157</td>
<td>0.772</td>
<td>0.997</td>
<td>0.902</td>
</tr>
<tr>
<td>0.667</td>
<td>0.90</td>
<td>250</td>
<td>0.5</td>
<td>0.164</td>
<td>0.544</td>
<td>0.56</td>
<td>0.804</td>
</tr>
<tr>
<td>1</td>
<td>0.90</td>
<td>250</td>
<td>0.5</td>
<td>0.058</td>
<td>0.201</td>
<td>0.01</td>
<td>0.039</td>
</tr>
</tbody>
</table>

Panel A: Sample Size = 250

<table>
<thead>
<tr>
<th>Break Point</th>
<th>Significance Level</th>
<th>Sample Size</th>
<th>Change Point</th>
<th>$H_0$</th>
<th>$H_1$</th>
<th>$H_2$</th>
<th>$H_3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5</td>
<td>0.99</td>
<td>500</td>
<td>0.5</td>
<td>0.056</td>
<td>0.887</td>
<td>1</td>
<td>0.981</td>
</tr>
<tr>
<td>0.667</td>
<td>0.99</td>
<td>500</td>
<td>0.5</td>
<td>0.073</td>
<td>0.714</td>
<td>0.685</td>
<td>0.935</td>
</tr>
<tr>
<td>1</td>
<td>0.99</td>
<td>500</td>
<td>0.5</td>
<td>0.027</td>
<td>0.406</td>
<td>0.005</td>
<td>0.083</td>
</tr>
<tr>
<td>0.5</td>
<td>0.95</td>
<td>500</td>
<td>0.5</td>
<td>0.11</td>
<td>0.93</td>
<td>1</td>
<td>0.995</td>
</tr>
<tr>
<td>0.667</td>
<td>0.95</td>
<td>500</td>
<td>0.5</td>
<td>0.13</td>
<td>0.825</td>
<td>0.837</td>
<td>0.973</td>
</tr>
<tr>
<td>1</td>
<td>0.95</td>
<td>500</td>
<td>0.5</td>
<td>0.068</td>
<td>0.76</td>
<td>0.067</td>
<td>0.337</td>
</tr>
<tr>
<td>0.5</td>
<td>0.90</td>
<td>500</td>
<td>0.5</td>
<td>0.154</td>
<td>0.945</td>
<td>1</td>
<td>0.996</td>
</tr>
<tr>
<td>0.667</td>
<td>0.90</td>
<td>500</td>
<td>0.5</td>
<td>0.164</td>
<td>0.86</td>
<td>0.894</td>
<td>0.977</td>
</tr>
<tr>
<td>1</td>
<td>0.90</td>
<td>500</td>
<td>0.5</td>
<td>0.108</td>
<td>0.875</td>
<td>0.178</td>
<td>0.517</td>
</tr>
</tbody>
</table>

Panel B: Sample Size = 500

Note: this table reports the simulation test power results based on Kolmogorov-Smirnov Functional projection method with full functional factors as projected factors. All scenarios are reported for three significance levels (99%, 95% and 90%) under one null hypothesis ($H_0$) and three alternative hypothesis ($H_1$, $H_2$ and $H_3$) with details specified in the simulation section. In panel A, break points are settled as 0.5 (1/2 of sample size), 0.667 (2/3 of sample size) and 1 (no break point in testing sample) for sample size equal to 250. In panel B, break points are settled as 0.5 (1/2 of sample size), 0.667 (2/3 of sample size) and 1 (no break point in testing sample) for sample size equal to 500.

departure from $H_0$ is larger for $H_2$ case. Sample size 250 does generally underperform in all scenarios compared with sample size 500.

In order to control the DNS factors selection bias, fully functional method, projecting functional curve observations on orthogonal basis curve (e.g. Fourier series), is conducted with the same settings under all scenarios. Results for simulation power test can be found in table 3.2 for sample size 250 and 500 respectively.
CHAPTER 3. TERM STRUCTURE MEAN CHANGE DETECTION

From table 3.2, fully functional factors projected method has slightly lower testing power for most scenarios compared with the Dynamic Nelson-Siegel projected factors tests. Similar to table 3.1, in the absence of break points setting, functional detection power has been distorted. Sample size under 500 does also perform stable than sample size under 250, consistent with table 3.1 outcome.

To explore the statistics detection power in terms of different change point size, sensitivity on size of change is tested based on size range \([0, 1]\) with steps 0.02 under DNS model framework.

Mean under Change \(H_A\):

\[
\mu_i = \begin{cases} 
4.54 \\ -2.82 \\ -3.03 
\end{cases}, \quad i \leq \frac{N}{2}; \quad \mu_i = \begin{cases} 
4.54 + (0.02r) \\ -2.82 + (0.02r) \\ -3.03 + (0.02r) 
\end{cases}, \quad i > \frac{N}{2}
\]

where, \(r\) is the index to control the size of change (\(r = 0, 1, 2, \ldots, 50\)).

Simulations are based on the middle real change point \((N/2)\), sample size 250 and 500, break point \((N/2\) and \(3N/4)\). Specifically, simulation is implemented with adding new vector values on the functional mean, starting from \([0, 0, 0]\) up to \([1, 1, 1]\) with an increment vector \([0.02, 0.02, 0.02]\). It is worth mentioning that under the initial case, the change size equal to 0 is the same as testing rejection error under the null hypothesis. Results are presented in plot with detection power against the size of change for both 250 and 500 sample sizes and different significance levels.

Change point magnitude sensitivity analysis is conducted with the sample size 250 and 500 in figure 3.2 and 3.3 respectively. Given x-axis standing for change point size, figure 3.2 and 3.3 show change point test power on y-axis for three significance level (90%, 95% and 99%) under two scenarios: \(BP = CP\) and \(BP \neq CP\). Similar conclusions echoing the previous part, location of break point does not alter the testing power significantly as long as the break point has been settled in the simulated data generating process. Moreover, monotonic momentum on testing power with respect to
change magnitude is confirmed under case of BP = CP only.

Figure 3.2: Sensitivity of Testing Power on Change Point Magnitude with Sample Size 250

Note: this figure shows sensitivity of functional testing power (Y-axis) with respect to change point magnitude (X-axis). Y-axis is the rejection rate ranging from 0 to 1 and X-axis is change point magnitude ranging from 0 to 1 with step 0.2. Simulation sample size is 250 with top panel plot under scenario break point equal to change point (BP = CP) and bottom panel plot under scenario break point is not equal to change point (BP ≠ CP). Red solid, green dot and blue dash lines represent testing power value under significance level 90%, 95% and 99% respectively.
Figure 3.3: Sensitivity of Testing Power on Change Point Magnitude with Sample Size 500

Note: this figure shows sensitivity of functional testing power (Y-axis) with respect to change point magnitude (X-axis). Y-axis is the rejection rate ranging from 0 to 1 and X-axis is change point magnitude ranging from 0 to 1 with step 0.2. Simulation sample size is 500 with top panel plot under scenario break point equal to change point (BP = CP) and bottom panel plot under scenario break point is not equal to change point (BP ≠ CP). Red solid, green dot and blue dash lines represent testing power value under significance level 90%, 95% and 99% respectively.

By comparison, sample size does make a strong difference regardless of change point magnitude on size interval. Roughly speaking, sample size 500 does give a much better performance in terms of the rejection rate on all change size cases. Convergence speed is relatively faster in the large sample size 500. In both case when break point is consistent with change point, testing power will converge
to probability 1 as long as the change size is relatively large enough.

### 3.5 Empirical Finite Sample Analysis

Real time data application based on Kolmogorov-Smirnov functional statistics is conducted. The purpose is to (1) provide the real life financial market application on this new statistic (2) conduct post-change analysis on subsample obtained from whole sample partition via change point location. Once the change point is documented, empirical sample data is separated according to change points location. Then further regression analysis is run for the purpose of explaining underlying reasons behind this change. Binary segmentation method is implemented until no changes detected since method in this chapter allows multiple finite changes in the data generating process. Specifically, whole testing sample is divided by changes location and continue the same detection procedure for divided samples until no rejection is recorded.

The initial attempt is via the Dynamic Nelson-Siegel (DNS) model in which "Level", "Slope" and "Curvature" factors will be used to project infinite dimensional functional data onto finite space. Furthermore, for change point detection robustness check, fully functional method is applied to avoiding the DNS model dependent issue (bias) caused by model misspecification problem (e.g. change point possibilities embedded in the lambda value).

Driving forces behind the change point detected on projected functional observation is normally not easy. However, given the advantages of the DNS model, research interest can be well transferred on analysing the latent factors dynamics for simplicity and clear interpretation view. By analysing factors especially on the "Slope" and "Curvature" latent factors, regime switching (market expectation shift) can be statistically identified, matching backwardation and contango theory in commodity futures literature (see discussions from Gorton et al. (2007), Gorton and Rouwenhorst (2004) and
In this functional setting, structure break point on the term structure identification is more solid compared with similar studies on the single time series analysis, often on either spot market data or front future contract data. Formally speaking, functional test considers term structure exposure across all maturities which is superior to the single maturity series test.

3.5.1 Data Description

Commodity futures market data are selected as the testing sample with specified products: metal product Gold from the Commodity Mercantile Exchange (COMEX) and energy product Light Crude Oil from the New York Mercantile Exchange (NYMEX). Considerations on these two commodity futures market data structure selection in the functional change point test are listed in following reasons: (1) consistent data structure (on each observed point, contracts prices are spanning across an increasing ordered maturities’ interval without missing points), (2) literature supported structure data good fitting from DNS (Grønborg and Lunde, 2016 Barunik and Malinska, 2016 and Karstanje et al., 2017), (3) futures market specification research interests (detecting changes on then DNS model in this market can provide both statistical argument on the current futures term structure modelling imperfection (inclusion of regime and error correction in DNS model (Nomikos and Pouliasis, 2015) and adaptive change detection in DNS modelling (Chen and Niu, 2014)) and reflect market change on demand-supply relation as well as future expectation (Karstanje et al., 2017)), (4) these two products have the best term structure in terms of numbers of maturities and time dimension(5) gold and light crude oil are closely related to financial crisis change but reacts differently (e.g. in financial crisis 2008, gold gains popularity as markets’ risk level goes up and investors would prefer less-risky asset, while light crude oil priced in USD and linked to fundamental industry will react more obvious than gold somehow, other products, for example, cotton does not vary too much as driving forces behind

34 Other products from data source have discrete maturity interval varying across years.
this category are different).

Data is from DataStream with time spanning from 04/05/2007 to 31/12/2009 in weekly frequency, 139 observations. Futures contract has pre-determined expiration date which details the last trading date for physical products’ transaction. Consistent with the standard literature, contract rolling is first applied and the linear interpolation is then deployed for constant time-to-maturity calculation (Diebold and Li [2006]). Commodity futures price term structure is organized as a matrix with columns standing for time-to-maturity and rows standing for continuous time dimension.

In this section, functional change point detection procedure is implemented with number of maturity contract equal to 18 (months) and 11 (months) separately. The reason for short time-to-maturity inclusion is aimed at controlling the illiquid effect caused by longer maturity contracts as they are less traded in reality (Heidorn et al., 2015). The empirical experiments are consistent across both gold and light crude oil products data. In conclusion, generally, testing procedure is applied to both matrices with 18 and 11 columns respectively. The maximum number of column is limited by the data itself (there is no extrapolation to manipulate the data structure in this case).

Compared with simulation test, empirical analysis is somehow referred as a finite sample test. Small sample size testing is selected to avoid mainly due to change detection property under the alternative when sample size goes to infinity. That is, statistics under large sample size scenario will suggest change point occurrence with probability equal to one.

For regression analysis, inspired by large scale database re-factoring method (principal component analysis) from Stock and Watson (2012) and closely related studies on commodity futures market by Karstanje et al. (2017), following data are collected (1) interest rate (Federal Funds, 3-month treasure bill and 6-month treasury bill), exchange rate (traded weighted U.S. dollar index, major currency, index) and TED spread (spread between 3-month treasury bill and 3-month interbank Libor rate)
are downloaded from FRED St. Louis database, (2) Equity data (Dow Jones Industry Index and S&P 500 Index), BDI (Baltic Dry Index) and CRBSPOT (Commodity Research Bureau spot market price index) from DataStream (3) ADS (Aruoba Diebold Scotti financial conditions index) is collected from Federal Reserve Bank of Philadelphia.

To explore the contribution or deriving force from trading specific behaviour, both short and long open interest for different trader types (Managed Money, Swap Dealer, Producers, Hedgers and Speculators) are collected from Commodity Futures Trading Commission following studies in the literature (Büyükşahin and Harris, 2011, Irwin and Sanders, 2012 and Heidorn et al., 2015). For measurement calculation of different types of traders, both long position effect and net long position effect are considered in this chapter for comprehensive comparison.

### 3.5.2 Dynamic Nelson-Siegel model Estimation

Following the literature of estimating latent factors in the DNS model, fixing lambda is a prerequisite. The two-step process is implemented by fixing lambda first and then doing cross-sectional regression for term structure data on each week observation (Diebold and Li, 2006). Grid search optimization (minimization of sum of squared errors under cross-sectional fitting) with interval range for lambda value (0,1] and step 0.001 is implemented practically. The optimal lambda value selection criterion can be represented as the following logical:

\[
\lambda^* = \inf \left\{ \lambda_k : \min \left[ \sum_{j=1}^{N} \sum_{k=1}^{K} \varepsilon_{ij}(\lambda_k)^T \varepsilon_{ij}(\lambda_k) \right] \right\}, \quad \lambda_k = 0.001, 0.002, 0.003, \ldots, 1, \quad k = 1, 2, 3, \ldots, K = 1000
\]

where, \( N \) is the sample length, \( \varepsilon \) is the cross-sectional estimation error. With the optimal lambda value, cross-sectional regression will automatically yield the estimated coefficients for latent factors.

Figure 3.4 shows the estimation of latent factors dynamics across time dimension with optimal
Figure 3.4: Projected Factor Dynamics for Assets Cross Financial Crisis Period

Gold Projected Factor Dynamics

Light Crude Oil Projected Factor Dynamics

Note: this figure reports the Dynamic Nelson-Siegel model fitting results on weekly price term structure of crude oil, Gold and Light crude oil from 04/05/2007 to 31/12/2009 with daily frequency data (T = 1, 2, …, 139). Lambda value is first optimized within sample data and then regression method is applied to estimate three factors: “Level”, “Slope” and “Curvature”.
CHAPTER 3. TERM STRUCTURE MEAN CHANGE DETECTION

lambda value 0.1 for both gold and light crude oil. Factors estimation is conducted without scaling, therefore “Level” factor (red solid line) represents the average price level across all time-to-maturity. It is also obvious to see that this time-to-maturity averaged effect has a big drop responded to the financial crisis in 2008, from around 1700 to nearly 1000 for gold and from 125 to below 50 for light crude oil. ”Slope” and ”Curvature” factors also document this change with an opposite movement.

It is worth mentioning that light crude oil market generally shows more fluctuation within full sample compared with gold. This can be evidenced from [Kesicki, 2010 and Singleton, 2013] who identify that oil market shift from backwardation to contango after 16/05/2008. On the contrary, in the whole sample period, gold market is always on backwardation, witnessed by the negative slope factors across financial crisis period. Given the increasing risk fact, a safer invest asset is expected and gold in somehow is a good candidate, which then push up the short-term contract price. This is also consistent with the findings by [Andreasson et al., 2016] who document that gold price jumps without experiencing trough between 2007 - 2009.

3.5.3 Empirical Functional Change Point Detection

In this part, empirical functional change detection is completed by testing selected commodity futures market data during the recent financial crisis period. Summary statistics of commodity futures price term structure in terms of different time-to-maturity are reported for both crude oil and gold in table 3.3 and 3.4 respectively.

The sample size is fixed for 139 trading weeks (last Friday daily observation) for two assets. Averaged contango effect can be observed across tables for two products. The longer the time-to-maturity, the higher the price level, which shows an upward sloping curve. Volatilities with respect to time-to-maturity across two assets show a decreasing trend, which is referred to Balassa-Samuelson
effect in the literature (Samuelson [1965]). This is mainly because the short-maturity contracts are more actively traded compared with the long-maturity ones. Therefore, it is more obvious to observe that the short-maturity contracts’ price variance is larger as this reflect the trading open interests and volumes contributed from market participants. As for long-maturity contracts, lack of liquidity (less traded) in another way makes price stable.

Accompanying with the summary statistics, dynamics of futures price term structure is presented via three-dimensional surface plots in figure 3.5. Different from the averaged effect, light crude oil structure shows a clear shift from backwardation to contango spanning financial crisis period while gold structure presents a change on slope from this figure.
CHAPTER 3. TERM STRUCTURE MEAN CHANGE DETECTION

Figure 3.5: Price Term Structure 3-Dimensional Surface Plot

Note: this figure shows dynamics of price term structure for two different assets. Top left panel is for Gold and bottom panel is for Light crude oil. X-axis is time-to-maturity starting from 1 to 18 standing for (1 month and 18-month maturity contract). Y-axis is the time dimension weekly data starting T = 0 to T = 139 (04/05/2007 to 31/12/2009)
CHAPTER 3. TERM STRUCTURE MEAN CHANGE DETECTION

Table 3.3: Summary Statistics for Gold Price Term Structure

<table>
<thead>
<tr>
<th>Maturity</th>
<th>N</th>
<th>Mean</th>
<th>SD</th>
<th>Median</th>
<th>Min</th>
<th>Max</th>
<th>Skew</th>
<th>Kurtosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maturity.1</td>
<td>139</td>
<td>871.188</td>
<td>123.288</td>
<td>895.400</td>
<td>645.500</td>
<td>1174.200</td>
<td>-0.021</td>
<td>-0.371</td>
</tr>
<tr>
<td>Maturity.2</td>
<td>139</td>
<td>872.581</td>
<td>122.811</td>
<td>896.500</td>
<td>647.500</td>
<td>1175.500</td>
<td>-0.018</td>
<td>-0.365</td>
</tr>
<tr>
<td>Maturity.3</td>
<td>139</td>
<td>874.657</td>
<td>122.144</td>
<td>897.300</td>
<td>650.300</td>
<td>1176.800</td>
<td>-0.014</td>
<td>-0.358</td>
</tr>
<tr>
<td>Maturity.4</td>
<td>139</td>
<td>877.632</td>
<td>121.045</td>
<td>899.000</td>
<td>656.400</td>
<td>1177.900</td>
<td>-0.012</td>
<td>-0.352</td>
</tr>
<tr>
<td>Maturity.5</td>
<td>139</td>
<td>881.047</td>
<td>119.737</td>
<td>901.100</td>
<td>662.500</td>
<td>1179.100</td>
<td>-0.007</td>
<td>-0.347</td>
</tr>
<tr>
<td>Maturity.6</td>
<td>139</td>
<td>884.422</td>
<td>118.509</td>
<td>904.500</td>
<td>668.500</td>
<td>1180.300</td>
<td>-0.001</td>
<td>-0.342</td>
</tr>
<tr>
<td>Maturity.7</td>
<td>139</td>
<td>887.831</td>
<td>117.393</td>
<td>908.900</td>
<td>674.500</td>
<td>1182.000</td>
<td>0.007</td>
<td>-0.336</td>
</tr>
<tr>
<td>Maturity.8</td>
<td>139</td>
<td>891.356</td>
<td>116.412</td>
<td>913.700</td>
<td>680.700</td>
<td>1184.300</td>
<td>0.017</td>
<td>-0.329</td>
</tr>
<tr>
<td>Maturity.9</td>
<td>139</td>
<td>895.037</td>
<td>115.553</td>
<td>918.800</td>
<td>686.800</td>
<td>1186.800</td>
<td>0.027</td>
<td>-0.323</td>
</tr>
<tr>
<td>Maturity.10</td>
<td>139</td>
<td>898.904</td>
<td>114.814</td>
<td>922.000</td>
<td>693.000</td>
<td>1189.600</td>
<td>0.038</td>
<td>-0.318</td>
</tr>
<tr>
<td>Maturity.11</td>
<td>139</td>
<td>902.927</td>
<td>114.205</td>
<td>924.600</td>
<td>699.300</td>
<td>1192.900</td>
<td>0.049</td>
<td>-0.315</td>
</tr>
<tr>
<td>Maturity.12</td>
<td>139</td>
<td>907.196</td>
<td>113.876</td>
<td>929.200</td>
<td>705.600</td>
<td>1200.900</td>
<td>0.070</td>
<td>-0.292</td>
</tr>
<tr>
<td>Maturity.13</td>
<td>139</td>
<td>913.024</td>
<td>113.948</td>
<td>928.700</td>
<td>711.700</td>
<td>1215.200</td>
<td>0.101</td>
<td>-0.256</td>
</tr>
<tr>
<td>Maturity.14</td>
<td>139</td>
<td>924.340</td>
<td>114.704</td>
<td>938.700</td>
<td>717.800</td>
<td>1232.700</td>
<td>0.144</td>
<td>-0.222</td>
</tr>
<tr>
<td>Maturity.15</td>
<td>139</td>
<td>940.069</td>
<td>114.911</td>
<td>955.800</td>
<td>736.800</td>
<td>1252.500</td>
<td>0.186</td>
<td>-0.199</td>
</tr>
<tr>
<td>Maturity.16</td>
<td>139</td>
<td>957.006</td>
<td>115.654</td>
<td>972.900</td>
<td>755.800</td>
<td>1273.700</td>
<td>0.228</td>
<td>-0.178</td>
</tr>
<tr>
<td>Maturity.17</td>
<td>139</td>
<td>974.891</td>
<td>116.836</td>
<td>988.100</td>
<td>775.200</td>
<td>1296.600</td>
<td>0.269</td>
<td>-0.160</td>
</tr>
<tr>
<td>Maturity.18</td>
<td>139</td>
<td>992.737</td>
<td>116.126</td>
<td>1002.600</td>
<td>794.900</td>
<td>1302.500</td>
<td>0.208</td>
<td>-0.382</td>
</tr>
</tbody>
</table>

Note: the table reports summary statistics of daily commodity future price data for different time-to-maturity from 01/01/2007 to 31/12/2009. The table first row is organized as the label for each column meaning: N (number of observations), Mean (mean value), SD (standard deviation), Median (median of price level), Max (maximum price value), Min (minimum price value), Skew (skewness of price) and Kurtosis (kurtosis of price).

Given the optimal lambda and corresponding DNS three factors, projected vector is easier to be obtained once functional observation is projected onto "Level", "Slope" and "Curvature". Functional change detection is then experimented on testing the mean change of projected vectors.

This whole procedure is repeated for both gold and light crude oil estimated projected factors with price term structure maturities equalling 18 and 11 months. In the fully functional change detection process, factors are selected from orthonormal basis functions (Fourier basis in this chapter) and implemented for gold and light crude oil with the same maturity level as well. Functional empirical change detection results are shown in table 3.5.
Table 3.4: Summary Statistics for Light Crude Oil Price Term Structure

<table>
<thead>
<tr>
<th>Maturity</th>
<th>N</th>
<th>Mean</th>
<th>SD</th>
<th>Median</th>
<th>Min</th>
<th>Max</th>
<th>Skew</th>
<th>Kurtosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maturity.1</td>
<td>139</td>
<td>80.690</td>
<td>25.806</td>
<td>75.470</td>
<td>37.710</td>
<td>145.290</td>
<td>0.561</td>
<td>-0.337</td>
</tr>
<tr>
<td>Maturity.2</td>
<td>139</td>
<td>81.315</td>
<td>25.037</td>
<td>75.370</td>
<td>40.530</td>
<td>145.860</td>
<td>0.660</td>
<td>-0.222</td>
</tr>
<tr>
<td>Maturity.3</td>
<td>139</td>
<td>81.734</td>
<td>24.500</td>
<td>75.260</td>
<td>42.280</td>
<td>146.120</td>
<td>0.724</td>
<td>-0.122</td>
</tr>
<tr>
<td>Maturity.4</td>
<td>139</td>
<td>82.051</td>
<td>24.078</td>
<td>75.060</td>
<td>43.750</td>
<td>146.310</td>
<td>0.774</td>
<td>-0.036</td>
</tr>
<tr>
<td>Maturity.5</td>
<td>139</td>
<td>82.322</td>
<td>23.719</td>
<td>75.470</td>
<td>45.080</td>
<td>146.500</td>
<td>0.816</td>
<td>0.040</td>
</tr>
<tr>
<td>Maturity.6</td>
<td>139</td>
<td>82.565</td>
<td>23.395</td>
<td>75.950</td>
<td>46.330</td>
<td>146.680</td>
<td>0.852</td>
<td>0.110</td>
</tr>
<tr>
<td>Maturity.7</td>
<td>139</td>
<td>82.774</td>
<td>23.097</td>
<td>76.450</td>
<td>47.340</td>
<td>146.790</td>
<td>0.884</td>
<td>0.174</td>
</tr>
<tr>
<td>Maturity.8</td>
<td>139</td>
<td>82.961</td>
<td>22.810</td>
<td>76.910</td>
<td>48.120</td>
<td>146.810</td>
<td>0.914</td>
<td>0.235</td>
</tr>
<tr>
<td>Maturity.9</td>
<td>139</td>
<td>83.129</td>
<td>22.526</td>
<td>76.950</td>
<td>48.900</td>
<td>146.680</td>
<td>0.941</td>
<td>0.290</td>
</tr>
<tr>
<td>Maturity.10</td>
<td>139</td>
<td>83.286</td>
<td>22.253</td>
<td>77.150</td>
<td>49.620</td>
<td>146.530</td>
<td>0.966</td>
<td>0.340</td>
</tr>
<tr>
<td>Maturity.11</td>
<td>139</td>
<td>83.434</td>
<td>21.990</td>
<td>77.090</td>
<td>50.330</td>
<td>146.340</td>
<td>0.990</td>
<td>0.386</td>
</tr>
<tr>
<td>Maturity.12</td>
<td>139</td>
<td>83.573</td>
<td>21.740</td>
<td>77.420</td>
<td>51.020</td>
<td>146.100</td>
<td>1.012</td>
<td>0.428</td>
</tr>
<tr>
<td>Maturity.13</td>
<td>139</td>
<td>83.698</td>
<td>21.504</td>
<td>77.750</td>
<td>51.670</td>
<td>145.870</td>
<td>1.034</td>
<td>0.469</td>
</tr>
<tr>
<td>Maturity.14</td>
<td>139</td>
<td>83.815</td>
<td>21.283</td>
<td>77.910</td>
<td>52.280</td>
<td>145.640</td>
<td>1.054</td>
<td>0.508</td>
</tr>
<tr>
<td>Maturity.15</td>
<td>139</td>
<td>83.925</td>
<td>21.076</td>
<td>78.160</td>
<td>52.860</td>
<td>145.420</td>
<td>1.074</td>
<td>0.545</td>
</tr>
<tr>
<td>Maturity.16</td>
<td>139</td>
<td>84.030</td>
<td>20.882</td>
<td>78.490</td>
<td>53.390</td>
<td>145.200</td>
<td>1.093</td>
<td>0.579</td>
</tr>
<tr>
<td>Maturity.17</td>
<td>139</td>
<td>84.127</td>
<td>20.700</td>
<td>78.580</td>
<td>53.920</td>
<td>144.980</td>
<td>1.110</td>
<td>0.611</td>
</tr>
<tr>
<td>Maturity.18</td>
<td>139</td>
<td>84.223</td>
<td>20.526</td>
<td>78.750</td>
<td>54.450</td>
<td>144.750</td>
<td>1.126</td>
<td>0.637</td>
</tr>
</tbody>
</table>

Note: the table reports summary statistics of daily commodity future price data for different time-to-maturity from 01/01/2007 to 31/12/2009. The table first row is organized as the label for each column meaning: N (number of observations), Mean (mean value), SD (standard deviation), Median (median of price level), Max (maximum price value), Min (minimum price value), Skew (skewness of price) and Kurtosis (kurtosis of price).

Testing procedure via the DNS model factors indicates significant results on both rejection significance level and change point location. Robustness check via fully functional factors confirms the validity of testing results from the DNS model factors. Overall, consistency results are documented regardless of methods used, suggesting the existence of the real change points. It is clear to see that change point location for oil market (10/10/2008) is 3 months earlier than gold market (23/01/2009).
### Table 3.5: Empirical Functional Change Point Detection Results

<table>
<thead>
<tr>
<th>Breakpoint</th>
<th>Critical Value_90</th>
<th>Critical Value_95</th>
<th>Critical Value_99</th>
<th>Statistics</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Panel A: DNS Projected Factors</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gold Maturity 18</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.5</td>
<td>27.849</td>
<td>29.551</td>
<td>33.234</td>
<td>37.942***</td>
<td>2009/01/23</td>
</tr>
<tr>
<td>Gold Maturity 11</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.5</td>
<td>27.146</td>
<td>28.782</td>
<td>32.464</td>
<td>41.366***</td>
<td>2009/01/23</td>
</tr>
<tr>
<td>Light Crude Oil Maturity 18</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.5</td>
<td>5.357</td>
<td>5.719</td>
<td>6.385</td>
<td>7.664***</td>
<td>2008/10/10</td>
</tr>
<tr>
<td>Light Crude Oil Maturity 11</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.5</td>
<td>5.547</td>
<td>5.886</td>
<td>6.564</td>
<td>9.139***</td>
<td>2008/10/10</td>
</tr>
<tr>
<td><strong>Panel B: Fully Functional Factors</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gold Maturity 18</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.5</td>
<td>27.790</td>
<td>29.475</td>
<td>32.546</td>
<td>37.787***</td>
<td>2009/01/23</td>
</tr>
<tr>
<td>Gold Maturity 11</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.5</td>
<td>27.065</td>
<td>28.791</td>
<td>32.209</td>
<td>41.199***</td>
<td>2009/01/23</td>
</tr>
<tr>
<td>Light Crude Oil Maturity 18</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.5</td>
<td>5.365</td>
<td>5.707</td>
<td>6.347</td>
<td>7.671***</td>
<td>2008/10/10</td>
</tr>
<tr>
<td>Light Crude Oil Maturity 11</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.5</td>
<td>5.580</td>
<td>5.915</td>
<td>6.647</td>
<td>9.146***</td>
<td>2008/10/10</td>
</tr>
</tbody>
</table>

**Note:** the table reports functional change detection results via both Dynamic Nelson-Siegel model projected factors (Panel A) and Fully Functional Factors (Panel B) on both gold and light crude oil market with weekly frequency data from 04/05/2007 to 31/12/2009. The first row in table is labelled as: Breakpoint (break point setting statistics calculation, setting for this is fixed at 0.5), Critical Value\_90, Critical Value\_95, Critical Value\_99 (empirical critical value calculated for significance level 90%, 95% and 99%), Statistics (Kolmogorov-Smirnov functional statistics) and Location (sample location where maximum of functional statistics is reached). For each panel, detection results are listed for two assets (Gold and Light Crude Oil) with consideration on number of maturity selected in functional detection procedure (11 and 18 contracts with monthly time-to-maturity).
3.5.4 Empirical Regression Analysis

Based on the findings (change point location) above, change point implied economic situation change is further investigated in this section. Although it is possible to do functional regression test, employing scalar version is easier to interpret. Therefore, in the practical regression analysis, time-series latent factors\(^{35}\) that represent the term structure dynamics exposure on "Level", "Slope" and "Curvature" factors are used in this analysis as dependent variables. For each asset, latent factors samples are separated into sample before change and sample after change given the change points’ locations from the above section. Latent factors are then regressed on selected explanatory information set for both subsamples and full-sample based on weekly observations. Since idea here is to explore the underlying driving force and sign contribution, regression coefficients are not scaled in this part.

The "Level" factor is transferred to log difference innovations, denoted as \(\Delta \text{Level}\)\(^{36}\) in the regression analysis. Positive (negative) contribution from this explanatory variable can be explained as increasing (decreasing) shift value as well as volatility. For the "Slope" factor, backwardation (contango) is recorded when regression coefficient is positive (negative). Positive (negative) estimated coefficients given the "Curvature" factor contribute to increase (decrease) of middle term contract price movement.

In the interest of anchoring the driving force behind the DNS modelling, regression analysis is conducted under the framework of multivariate forecasting approach. Generally, three latent factors are assumed to be linearly forecasted by market information from equity, interest rate, foreign exchange market, market financial conditions and volatility.

Moreover, to explore how market participators’ trading behaviours contribute to latent factors

\(^{35}\)see figure 3.4 for their time-series dynamics
\(^{36}\)Augmented Dickey Fuller test suggests that unit root exists in "Level" projected factor, therefore, first order difference of "Level" projected factors and outcome of this is regarded as factor innovation
future movement, \( Trade \) is the proxy variable for market participation ratio change for specific type behaviour (open interest data from the Commodity Futures Trading Commission, CFTC).

\[
\begin{align*}
\text{Trade}^{\text{long}}_{i,j,t-1} &= \frac{\text{#positions}_{j,i,t-1}}{\sum_{j=1}^{\text{#positions}}_{j,i,t-1}} \\
\text{Trade}^{\text{net,long}}_{i,j,t-1} &= \frac{\text{#longpositions}_{j,i,t-1} - \text{#short positions}_{j,i,t-1}}{\text{#longpositions}_{j,i,t-1} + \text{#short positions}_{j,i,t-1}}
\end{align*}
\]

where, \( \text{Trade}^{\text{long}}_{i,j,t-1} \) is an set of explanatory variables to explore long only trading position taken by trading categories (\( j \) represented by either Managed Money, Swap Dealer, Producers, Hedgers and Speculators open interest in corresponding asset) data collected from the CFTC for asset \( i \) at time \( t - 1 \). \( \text{Trade}^{\text{net,long}}_{i,j,t-1} \) is for net long position calculation (long position minus short position).

\[
\text{Factor}_{i,t} = \alpha_i + \beta_{1,i}\text{Equity}_{t-1} + \beta_{2,i}\text{FinancialIndex}_{t-1} + \beta_{3,i}\text{CRBSPOT}_{t-1} + \beta_{4,i}\text{IR}_{t-1} \]

\[
+ \beta_{5,i}\text{TED}_{t-1} + \beta_{5,i}\text{EX}_{t-1} + \beta_{5,i}\text{Trade}_{i,j,t-1} + \epsilon_{i,t} \quad (3.5.1)
\]

where \( \text{Factor}_{i,t} \) is the projected factor with index \( i = 1 (\"\Delta \text{Level}\”) , 2 (\"\text{Slope}\”) and 3 (\"\text{Curvature}\”) \) and \( t \) is time dimension, \( \beta \) is the factor regression coefficients with index \( i = 1, 2, 3, 4, 5 \), \( \text{Equity}_{t-1} \) is the PCA transformed proxy variable for equity market in terms of the Dow Jones Industry Index and the S&P 500 Index at time \( t - 1 \), \( \text{FinancialIndex}_{t-1} \) is the PCA transformed proxy variable for financial conditions in terms of Aruboa Diebold Scotti index and Dry bulk shipping index at time \( t - 1 \), \( \text{CRBSPOT}_{t-1} \) is the volatility of commodity Research Bureau spot market price index at time \( t - 1 \), \( \text{IR}_{t-1} \) is the PCA transformed main component in terms of effective fund rate, 3-month secondary market T-bill and 6-month secondary market T-bill at time \( t - 1 \), \( \text{TED}_{t-1} \) is the differential between 3-month inter-bank rate and 3-month Libor rate at time \( t - 1 \), \( \text{EX}_{t-1} \) is Traded Weighted U.S Dollar Index for Major currencies index at time \( t - 1 \), \( \text{Trade}_{i,j,t-1} \) is universe version, delegating both \( \text{long} \) and \( \text{net,long} \) version above.
Table 3.6: Pairwise Correlation Matrix for Gold CFTC Participators Position

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>-0.496</td>
<td>1.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>-0.136</td>
<td>0.658</td>
<td>1.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>-0.673</td>
<td>0.935</td>
<td>0.636</td>
<td>1.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>0.733</td>
<td>-0.293</td>
<td>-0.133</td>
<td>-0.482</td>
<td>1.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>0.718</td>
<td>-0.937</td>
<td>-0.610</td>
<td>-0.977</td>
<td>0.473</td>
<td>1.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>-0.509</td>
<td>0.856</td>
<td>0.779</td>
<td>0.867</td>
<td>-0.403</td>
<td>-0.893</td>
<td>1.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>0.680</td>
<td>-0.886</td>
<td>-0.663</td>
<td>-0.907</td>
<td>0.295</td>
<td>0.921</td>
<td>-0.815</td>
<td>1.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>-0.622</td>
<td>0.853</td>
<td>0.759</td>
<td>0.902</td>
<td>-0.498</td>
<td>-0.905</td>
<td>0.939</td>
<td>-0.878</td>
<td>1.000</td>
<td>0.000</td>
</tr>
<tr>
<td>0.414</td>
<td>-0.581</td>
<td>-0.169</td>
<td>-0.620</td>
<td>0.577</td>
<td>0.635</td>
<td>-0.543</td>
<td>0.305</td>
<td>-0.475</td>
<td>1.000</td>
</tr>
</tbody>
</table>

Note: This table reports the pairwise correlation results for all gold market participators’ open interest data in both long only and net value measure. The first column and first row are standing for different participators’ category in both long only (denoted as L) and net (denoted as N) open interest measurement. Values in bold are at least 95% significant.

Table 3.7: Pairwise Correlation Matrix for Light Crude Oil CFTC Participators Position

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>0.212</td>
<td>1.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>0.462</td>
<td>-0.411</td>
<td>1.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>0.059</td>
<td>0.834</td>
<td>-0.264</td>
<td>1.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>0.242</td>
<td>0.466</td>
<td>-0.666</td>
<td>0.263</td>
<td>1.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>0.165</td>
<td>-0.376</td>
<td>-0.104</td>
<td>-0.490</td>
<td>0.206</td>
<td>1.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>0.015</td>
<td>-0.526</td>
<td>0.650</td>
<td>-0.321</td>
<td>-0.804</td>
<td>0.287</td>
<td>0.101</td>
<td>1.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>-0.101</td>
<td>0.267</td>
<td>0.187</td>
<td>0.596</td>
<td>-0.227</td>
<td>-0.782</td>
<td>0.708</td>
<td>-0.077</td>
<td>1.000</td>
<td>0.000</td>
</tr>
<tr>
<td>0.190</td>
<td>0.076</td>
<td>-0.302</td>
<td>-0.158</td>
<td>0.663</td>
<td>0.179</td>
<td>-0.409</td>
<td>-0.728</td>
<td>-0.324</td>
<td>1.000</td>
</tr>
</tbody>
</table>

Note: This table reports the pairwise correlation results for all light crude oil market participators’ open interest data in both long only and net value measure. The first column and first row are standing for different participators’ category in both long only (denoted as L) and net (denoted as N) open interest measurement. Values in bold are at least 95% significant.

Pairwise correlation matrices are reported for both gold and light crude oil market participators’ CFTC data. Variable names are listed in either the first column or the first row of these two tables. For either long position only or net position value, it is obvious that high correlation coefficient value (over 0.5 and some of them over 0.8 and 0.9), no matter positive or negative, can be observed and at least 95% significant. Therefore, in the following regression analysis, this chapter does not take into account of the scenario in which all variables listed in table 3.6 and 3.7 are included in explanatory variables for both gold and oil market.

From table 3.8, the first two factors, "\(\Delta\text{Level}\)" and "\(\text{Slope}\)”, have no exposure to nearly all common explanatory variables except for a little evidence on \(EX_{t-1}\). There is positive significant forecasting relations documented on "Curvature" factor from explanatory variables \(EX_{t-1}\) and \(\text{Equity}_{t-1}\).
### CHAPTER 3. TERM STRUCTURE MEAN CHANGE DETECTION

#### Table 3.8: Gold Subsample Regression Analysis under Hedging Pressure Effect

<table>
<thead>
<tr>
<th>Variable</th>
<th>Equation</th>
<th>R²</th>
<th>Adjusted R²</th>
<th>Residual Std. Error</th>
<th>F Statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Equity</td>
<td>0.80676</td>
<td>0.799</td>
<td>0.799</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Futures</td>
<td>0.80676</td>
<td>0.799</td>
<td>0.799</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Speculators</td>
<td>0.80676</td>
<td>0.799</td>
<td>0.799</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Hedgers</td>
<td>0.80676</td>
<td>0.799</td>
<td>0.799</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Panel B: CFCT Net Position</td>
<td>0.80676</td>
<td>0.799</td>
<td>0.799</td>
<td>0.000</td>
<td>0.000</td>
</tr>
</tbody>
</table>

Dependent variable: \( \Delta Level \)

### Panel B: CFCT Long Position

<table>
<thead>
<tr>
<th>Variable</th>
<th>Equation</th>
<th>R²</th>
<th>Adjusted R²</th>
<th>Residual Std. Error</th>
<th>F Statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Equity</td>
<td>0.80676</td>
<td>0.799</td>
<td>0.799</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Futures</td>
<td>0.80676</td>
<td>0.799</td>
<td>0.799</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Speculators</td>
<td>0.80676</td>
<td>0.799</td>
<td>0.799</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Hedgers</td>
<td>0.80676</td>
<td>0.799</td>
<td>0.799</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Panel B: CFCT Net Position</td>
<td>0.80676</td>
<td>0.799</td>
<td>0.799</td>
<td>0.000</td>
<td>0.000</td>
</tr>
</tbody>
</table>

Dependent variable: \( \Delta Level \)

Note: Panel A and panel B exact follows the same regression equation 3.5.1 with only difference on the usage of type of open interest data from CFTC. Panel A (B) use long only (net) open interest to calculate \( trade_{j,i,j-1} \) which is resented by Hedgers and Speculators listed in the first column.
Table 3.9: Light Crude Oil Subsample Regression Analysis under Hedging Pressure Effect

| Dependent variable | \( R^2 \) | \( F \) Statistic | \( F \) Statistic (df = 11; 79) | \( F \) Statistic (df = 11; 35) | \( F \) Statistic (df = 11; 126) | \( F \) Statistic (df = 11; 79) | \( F \) Statistic (df = 11; 35) | \( F \) Statistic (df = 11; 126) | \( F \) Statistic (df = 11; 79) | \( F \) Statistic (df = 11; 35) | \( F \) Statistic (df = 11; 126) | \( F \) Statistic (df = 11; 79) | \( F \) Statistic (df = 11; 35) | \( F \) Statistic (df = 11; 126) |
|-------------------|---------|------------------|-----------------|-----------------|-----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|
| Equities \( p = 1 \) | 0.091 | 41.387 | 41.387 | 41.387 | 41.387 | 41.387 | 41.387 | 41.387 | 41.387 | 41.387 | 41.387 | 41.387 | 41.387 | 41.387 | 41.387 |

Panel B: CFTC Net Position

| Dependent variable | \( R^2 \) | \( F \) Statistic | \( F \) Statistic (df = 11; 79) | \( F \) Statistic (df = 11; 35) | \( F \) Statistic (df = 11; 126) | \( F \) Statistic (df = 11; 79) | \( F \) Statistic (df = 11; 35) | \( F \) Statistic (df = 11; 126) | \( F \) Statistic (df = 11; 79) | \( F \) Statistic (df = 11; 35) | \( F \) Statistic (df = 11; 126) | \( F \) Statistic (df = 11; 79) | \( F \) Statistic (df = 11; 35) | \( F \) Statistic (df = 11; 126) |
|-------------------|---------|------------------|-----------------|-----------------|-----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|
| Equities \( p = 1 \) | 0.091 | 41.387 | 41.387 | 41.387 | 41.387 | 41.387 | 41.387 | 41.387 | 41.387 | 41.387 | 41.387 | 41.387 | 41.387 | 41.387 | 41.387 |

Note: Panel A and panel B exactly follow the same regression equation 3.5.1 with only difference on the usage of type of open interest data from CFTC.

Panel A (B) use long only (net) open interest to calculate \( t_{\text{trades}} \) which is represented by Hedgers and Speculators listed in the first column.
after financial crisis period. Generally, increase of the exchange rate and equity level at current time will push up the middle time-to-maturity contract price in next period after the financial crisis but have no impact before it. Although the equity coefficient result is consistent with the finding from Karstanje et al. (2017), result in this chapter specifies that this common finding is mainly due to data after the recent financial crisis.\footnote{Stronger trade weighted dollar exchange rate index causes higher factor level, implying positive forecasting power (Chen et al., 2010).}

In the view of trading behaviour impact, the proportion of the Hedgers and Speculators long only position does not help with identifying the driving force behind. However, the net positions of both Hedgers and Speculators explain the future variation of factors in terms of full sample and subsample before the financial crisis. Net positions increase will stabilize market level fluctuation and drive up short and middle term contract price level in the next period to backwardation. The results are consistent with the normal backwardation theory idea, discussed by Hamilton and Wu (2014), Gorton et al. (2007) and Gorton and Rouwenhorst (2004). Evidences are not consistent when the sample after financial crisis is taken into consideration, which states the fact that trading behaviours from hedging pressure effect is not able to explain structural change.

The same hedging pressure effect based on the light crude oil data is reported in table 3.9. The equity market shows different effects on the light crude oil market with positive forecasting relation on the "\Delta Level" and negative forecasting relation on the "Slope" and "Curvature". Increasing volatility calculated from the spot market price index reduces oil market level innovations across three samples, even though this effect is weakened when trading behaviour data type is moving from the long only to net effect.

Interest rate component is another key indicator positively (negatively) and significantly forecast-
ing the level innovation (slope and curvature dynamics). Different from Karstanje et al. (2017) studies, interest rate effect finding is more robust in this case given the pricing rationale behind (spot and future pricing formula from Gorton et al. (2007), the Carry idea from Koijen et al. (2013), stochastic interest rate model from Casassus and Collin D. (2005)).

Exchange rate plays the same role as it does in the gold market. TED spread is referred to market liquidity stating negatively significant forecasting relation to the "Slope" factor only. The larger the differential between 3-month interbank rate and 3-month Libor rate, the more contango will be (the short-term price lower than long term one). When the limitation of borrowing money from market is obvious, it may cause unwind of trading positions or liquidity is transferred to riskless asset, driving down short-term price level (Brunnermeier et al., 2008).

The long only open interest proportion from both Hedgers and Speculators successfully forecast the future movement of three factors for the sample after the financial crisis. For the "ΔLevel", this effect holds for all three samples with significant negatively relationship, reducing the level volatility. This finding is also consistent with literature of normal backwardation theory mentioned in gold market.

Overall, in the perspective of trading behaviour, up to now, light crude oil term structure change is more likely to be explained and forecasted by the long only positions’ Hedgers and Speculators in market, while gold market is only exposed to few fundamentals. This leads us to next analysis by including financial traders (marked by Managed Money and Swap Dealer).

In addition to the common information from table 3.8 and 3.9, the Managed Money and Producers traders are recorded with strong significant relations found in gold market in table 3.10 and no relation in light crude oil market in table 3.11. Results on oil market is matched with finding by Heidorn et al. (2015) who argue no forecasting effect from financial traders on WTI crude oil market.
Table 3.10: Gold Subsample Regression Analysis under Managed-Money Trading

**Panel B: CFTC Long Position**

<table>
<thead>
<tr>
<th>Equities</th>
<th>CRBSPOT</th>
<th>Slope</th>
<th>Constant</th>
</tr>
</thead>
<tbody>
<tr>
<td>Equity</td>
<td>TEDRATE</td>
<td>TEDRATE</td>
<td>CRBSPOT</td>
</tr>
<tr>
<td>-1.102</td>
<td>2,817.232</td>
<td>1,823.365</td>
<td>14.758</td>
</tr>
<tr>
<td>2.277</td>
<td>2,817.232</td>
<td>1,823.365</td>
<td>14.758</td>
</tr>
</tbody>
</table>

Panel B: CFTC Net Position

<table>
<thead>
<tr>
<th>Equities</th>
<th>CRBSPOT</th>
<th>Slope</th>
<th>Constant</th>
</tr>
</thead>
<tbody>
<tr>
<td>Equity</td>
<td>TEDRATE</td>
<td>TEDRATE</td>
<td>CRBSPOT</td>
</tr>
<tr>
<td>-1.102</td>
<td>2,817.232</td>
<td>1,823.365</td>
<td>14.758</td>
</tr>
<tr>
<td>2.277</td>
<td>2,817.232</td>
<td>1,823.365</td>
<td>14.758</td>
</tr>
</tbody>
</table>

Note: Panel A and panel B exact follows the same regression equation 3.5.1 with only difference in the usage of type of open interest data from CFTC. Panel A (B) use long only (net) open interest to calculate \( \text{trade}_{i,j-1} \) which is resented by \( \text{Producers} \) and \( \text{Managed Money} \) listed in the first column.
### CHAPTER 3. TERM STRUCTURE MEAN CHANGE DETECTION

**Table 3.11: Light Crude Oil Subsample Regression Analysis under Managed-Money Trading**

<table>
<thead>
<tr>
<th>Panel A</th>
<th>Panel B</th>
<th>Panel C</th>
</tr>
</thead>
<tbody>
<tr>
<td>Obs</td>
<td>Obs</td>
<td>Obs</td>
</tr>
<tr>
<td>126</td>
<td>126</td>
<td>126</td>
</tr>
</tbody>
</table>
| Panel A (B) use log (1 + return) to calculate crude oil, which is reused by Producers and Managed Money listed in the first column.

**Notes:**
- Panel A and panel B chart follows the same regression equation 3.1 with only difference on the usage of type of open interest data from CFTC.
Within gold market, among these two types of traders, the Producers’ long only position holds for all three samples with negative forecasting power in the "∆Level" and positive forecasting power in the "Slope" and "Curvature" factors. Impact from the Managed Money trader (according to CFTC report, the Managed Money trader is categorised as those who are commodity trading advisor service, commodity pool operators and hedge funds with conducting commodity trading for client) only hold for full sample.

From table 3.10, long only position held by the Managed Money trader has opposite sign compared with the Producers side, stating that speculation from managing money in commodity futures market contribute to increasing of market level fluctuation and convert market from backwardation to contango.

By considering the Swap Dealer effect in table 3.12 and 3.13, there is not significant effect found for both gold and light crude oil market. Overall, analysis of different types of traders’ effect in market shows that gold market is more controlled by the traditional Producers within financial crisis period and light crude oil market is more structured by the Hedgers and Speculators at the same sample period. Apart from this, gold market is less affected by fundamental information, except for equity and exchange rate market, while light crude oil market is well forecasted by equity, exchange rate, interest rate market, spot market volatility and TED spread. Therefore, there is a distinct difference (different fundamental and participators) among these two markets, which also support the different change locations from functional test part.

### 3.6 Summary

This chapter first extends recent research findings by Bardsley et al. (2017) in terms of functional change point detection by investigating the asymptotic property of a new functional change detection
\textbf{Table 3.12: Good Subsample Regression Analysis under Swap-Dealer Trading} 

<table>
<thead>
<tr>
<th>Panel A</th>
<th>Panel B: CFTC Net Position</th>
<th>Panel C: Swap Dealer</th>
<th>Panel D</th>
<th>Panel E: Swap Dealer</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \alpha )</td>
<td>( \beta )</td>
<td>( \gamma )</td>
<td>( \delta )</td>
<td>( \epsilon )</td>
</tr>
<tr>
<td>( -0.318 )</td>
<td>( -0.344 )</td>
<td>( -0.153 )</td>
<td>( -0.239 )</td>
<td>( -0.824 )</td>
</tr>
<tr>
<td>( -0.309 )</td>
<td>( -0.341 )</td>
<td>( -0.156 )</td>
<td>( -0.270 )</td>
<td>( -0.828 )</td>
</tr>
<tr>
<td>( -0.153 )</td>
<td>( -0.156 )</td>
<td>( -0.153 )</td>
<td>( -0.270 )</td>
<td>( -0.828 )</td>
</tr>
<tr>
<td>( -0.239 )</td>
<td>( -0.270 )</td>
<td>( -0.239 )</td>
<td>( -0.270 )</td>
<td>( -0.828 )</td>
</tr>
<tr>
<td>( -0.824 )</td>
<td>( -0.828 )</td>
<td>( -0.824 )</td>
<td>( -0.828 )</td>
<td>( -0.828 )</td>
</tr>
</tbody>
</table>

Note: Panel A and Panel B each follow the same regression equation 3.1, with only difference on the basis of type of open interest data from CFTC.
## Chapter 3. Term Structure Mean Change Detection

### Table 3.13: Light Crude Oil Subsample Regression Analysis under Swap-Dealer Trading

<table>
<thead>
<tr>
<th>Variable</th>
<th>Measure</th>
<th>Regression Equation</th>
<th>Coefficient</th>
<th>t-statistic</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Equity</td>
<td>$\alpha$</td>
<td>$\beta$</td>
<td>$\gamma$</td>
<td>$\delta$</td>
<td>$\epsilon$</td>
</tr>
<tr>
<td>Panel A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Panel B: CFTC Long Position

<table>
<thead>
<tr>
<th>Variable</th>
<th>Measure</th>
<th>Regression Equation</th>
<th>Coefficient</th>
<th>t-statistic</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Equity</td>
<td>$\alpha$</td>
<td>$\beta$</td>
<td>$\gamma$</td>
<td>$\delta$</td>
<td>$\epsilon$</td>
</tr>
<tr>
<td>Panel A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Panel B: CFTC Net Position

<table>
<thead>
<tr>
<th>Variable</th>
<th>Measure</th>
<th>Regression Equation</th>
<th>Coefficient</th>
<th>t-statistic</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Equity</td>
<td>$\alpha$</td>
<td>$\beta$</td>
<td>$\gamma$</td>
<td>$\delta$</td>
<td>$\epsilon$</td>
</tr>
</tbody>
</table>

Note: Panel A and panel B exact follows the same regression equation 3.5.1 with only difference on the usage of type of open interest data from CFTC.

Panel B (U) use long only (net) open interest to calculate $trade_{\text{weekly}}$, which is resented by $Producers$ and $Swap Dealer$ listed in the first column.
statistics based on Kolmogorov-Smirnov functional formula under both null and alternative hypothesis. After constructing and conducting simulations under same mean change settings, new statistics performs less conservative under the null but more powerful under the alternative. Apart from this, simulation results also confirm that break point (allowing the heterogeneous information included in functional error curve generation) needs to be considered and fail to do this will diminish testing power remarkably.

Sensitivity analysis for this new statistic is conducted with respect to the change point magnitude. Monotonic property is well observed under the scenario where change point is equal to the setting of break point and not be guaranteed under the scenario where change point is not equal to the setting of break point. Sample size does make a great contribution on the rate of convergence given the fact that testing power under alternative with 500 sample sizes converge to descent probability much quicker than testing power under 250 sample sizes. Results up to now are consistent with Bardsley et al. (2017) findings even though statistics functional form is different.

Real data application is also tested in this chapter with focus on commodity futures market with data from gold and light crude oil products. By smoothing data and performing functional change detection procedure, two significant change point on the mean of functional mean observations: 23/01/2009 for gold market and 10/10/2008 for light crude oil market are documented. Segmentation method is applied and no further change point is found. These results are robust to fully functional projection method which proposes the same change location and significance as Dynamic Nelson-Siegel method does.

Economic intuition behind these two change points are studied via subsample analysis in the framework of multivariate forecasting regression. In general, two markets show strong segmentations for change occurrence reasons. Regarding the common information explanation, future trend of
gold market on three factors are only explained by equity and exchange rate market current change, while light crude oil market future movement is nearly exposed to current movement of all common variables, especially significant larger impact from equity and spot market volatility sides.

As for the market heterogeneous effects, based on the open interest position data from the Commodity Futures Trading Commission (CFTC), participator recorded as the Producers makes consistent forecasting contribution across all subsamples for gold market. The larger the proportion of long only position for the Producers compared with other trading participators (Managed Money, Swap Dealer, Hedgers and Speculators), the lower the price level volatility and more stable of future backwardation market situation. While in light crude oil market, positions taken by both the Hedgers and Speculators have well forecasted and explained the future movement of three projected factors after financial crisis period. Current increasing of long only positions will leads to next period volatility reduction on level and trend for market backwardation. However, for sample before financial crisis, no market participator can forecast the future trend.
Chapter 4

Commodity Futures Basis Term Structure Forecasting

This chapter investigates the forecasting power of the functional predictive factors (functional autoregressive model with 1 period lag) on the futures basis term structure dynamics. Compared with other functional and non-functional forecasting models, functional predictive factor outperforms both statistically and economically. Moreover, the Model Confidence Set, used to select the best model set with respect to loss functions, shows that the functional predict factors have incomparable advantages on futures contracts forecasting with time-to-maturity date up to 10 months. The results are robust to different loss functions settings. Trading application based on forecasted futures basis is proposed and implemented showing superior advantages on strategy variance reduction and return magnification compared with traditional backwardation theory implied trading style.
4.1 Introduction

Motivated by the continuous fundamental important pricing theory, normal backwardation theory, see (Gorton et al., 2007 and Litzenberger and Rabinowitz, 1995), recent question of spot market hedging via futures contract when asymmetric futures basis effect matters, see (Lien and Yang, 2008), factors dynamics embedded in term structure idea, see convenience yield modelling from stochastic process (Schwartz, 1997 and Casassus and Collin D., 2005), figuring out and predicting the dynamics of futures factor, especially futures basis variation, plays a non-trivial role in understanding of futures market data dynamics. Meanwhile, inspired by the recent term structure data modelling, fitting and forecasting in the global commodity futures market, see (Karstanje et al., 2017, Grønborg and Lunde, 2016 and Barunik and Malinska, 2016), this chapter extends current literature futures basis factor modelling from the single time series to the term structure level. By observing the future basis on the term structure level, it is more natural to call this format data term structure futures basis curve, which is similar to the yield curve borrowed from bond market.


To formalise so called curve data idea, futures basis (logarithm price difference among two contracts with same underlying but different maturities) is then measured for all maturity range at each time point observation. Capturing and forecasting the dynamics of futures basis makes contribution to numerous aspects on futures research, for example pricing theory, risk factor trading application,
investment shocks and different market correlation analysis (Fernandez-Perez et al., 2017b; Gorton et al., 2007; Fuertes et al., 2015; Koijen et al., 2013; Yang, 2013 and Roll et al., 2007). Given the cross-sectional curve data, a new functional forecasting method, namely predictive factors method, introduced by Kargin and Onatski (2008) is used to depict their variations following the autoregressive process with lag order 1.

Real data application is conducted on the crude oil futures market over the past 20 years. This new functional predictive factor method shows superior advantage on forecasting error reduction with contract’s maturity up to 10 months, which is still valid and robust compared to all other functional and non-functional candidate models. It also holds the good advantages on forecasting shape keeping compared with other models, except for Naive approach which simply looks at past observation. However, forecasting futures basis under predictive factors method shows competitive merits on new variance minimization trading strategy, generating theoretically annualized average return (Sharpe ratio) without and with transaction at 40% (7.92) and 28% (5.44) maximum.

Up to this writing moment, there is no prior research talking about futures basis modelling on the term structure curve idea and this chapter is trying to fill this gap for the purpose of offering (1) a new perspective of practising real futures basis modelling and forecasting method compared with the current Dynamic Nelson-Siegel model, (2) forecasting comparison in terms of different aspects (overall forecasting error reduction, forecasting error dynamics, forecasting curve shape preserving, economically trading advantages) (3) new application idea, via variance minimisation framework, on multi-contract trading and robustness analysis via transaction cost consideration.

To sum up, this chapter is organised as follows: section 2 states literature related to both forecasting and related topics, section 3 goes through both data and methodology used in this chapter, section 4 covers all empirical analysis results as well as new real trading application, section 5 concludes all
findings in the end.

4.2 Literature Background

Different from other methods in high dimensional data modelling, functional data analysis acts as a more general way of defining observations as functions over certain sets, see more examples from Ramsay (2006). General theory is functional linear dependent (functional observation is linear represented by another one) and autoregressive model (functional observation is linear represented by itself lag value) is extensively studied and proposed by Bosq (2000) and recent version with more application, see Bosq (2012).

In this framework, all functional processes are defined in Hilbert and Banach spaces by assuming that functional data (error) observation is a mean zero (mean zero i.i.d with finite variation) element in spaces. Linear or autoregressive operator is an estimated variable linking and transferring the current functional observation to the future one.

For simplicity, this chapter considers only autoregressive process with order 1 (AR(1) therefore in the following content) for commodity futures data under the framework of functional analysis. From the literature view of the futures basis dynamics, mean-reversion idea is widely discussed on its time series modelling (Fong and See 2003, Monoyios and Sarno 2002 and Roll et al. 2007). This is mainly depending on single continuous contract series which does not account for maturity structure effect. Functional autoregressive model is somehow treated as a straightforward way of handling the structure effect. In the meantime, the long run mean of a mean-reversion process is constant when data stationary property is satisfied.

To estimate the functional linear operator, empirical functional principal components analysis (EFPCs therefore in the following content) is a widely used technique with comprehensive theoreti-
ical development in (Bosq, 2000). Given the well-estimated eigenfunctions and eigenvalues from functional observations, EFPCs method has good property on its convergence. More recent autoregressive modelling, forecasting and theoretical extensions can be seen from (Bosq, 2012) and (Antoniadis and Sapatinas, 2003). Besse et al. (2000) and Hörmann et al. (2010).

In this chapter, the key interest of modelling and forecasting method is dependent on the recent predictive factors model in autoregressive framework via Kargin and Onatski (2008). Different from Bosq (2000) who use EFPCs to construct the operator, predictive factors method filters out the factors that can minimize the forecasted error. Corresponding empirical functional forecasting performance comparison has been studied by Besse et al. (2000) and Didericksen et al. (2012). However, they argue that the Estimated Kernel method (the same idea to EFPCs) in the spirit of Bosq (2000) has the best performance. To fully explore predictive factors forecasting power, this kernel benchmark and other functional Naive, functional Random Walk method are included in this chapter. Other comprehensive review on forecasting comparison and application on these can be referred to Horváth and Kokoszka (2012).

Non-functional term structure level modelling and forecasting by Diebold and Rudebusch (2013), Dynamic Nelson-Siegel (DNS therefore in the following content), is also studied in this chapter due to recent increasing popularity with its application on commodity futures market. Barunik and Malinska (2016) extend DNS model via adding neutral network idea, Grönborg and Lunde (2016) model DNS factors based on copula framework, and Karstanje et al. (2017) generalize DNS model with consideration of seasonality factor). Following by Diebold and Rudebusch (2013), the same approach to estimate and forecast term structure data with assumption that lambda is fixed for whole sample period is used. Although there are some studies discussing the importance of time-varying lambda value in DNS model, see Hautsch and Ou (2012) and Koopman et al. (2010), this chapter will ignore this as
their tests do not show strong benefits by including this extra time-varying property.

For forecasting procedure, there is a strong assumption that linear operator is constant and not changed over time. There are studies discussing functional autoregressive and linear model change point detection theory (Horváth et al., 2014; Horváth et al., 2010; and Horváth and Kokoszka, 2012), which needs extra change detection implementation on estimated linear operator. This is out of the control and purpose of this chapter. In this study, linear operator is assumed to be constant at each estimation procedure. It also directs to the recent research, adaptive functional autoregressive forecasting method, proposed by Chen and Li (2017) who overcome this issue by estimating the time-varying operator based on the maximum likelihood method. For simplicity, rolling window method is used to minimize errors introduced in estimation by assuming each rolling sample estimator are constant and unbiased.

Storage and backwardation are important pricing theories in the commodity futures market, which differs from the futures index market due to physical inventory cost (Working, 1949; and Litzenberger and Rabinowitz, 1995). The basis differential (log difference between spot price and futures price) under backwardation or contango scenario are important for hedging and speculation activities.

Hedging and speculation for risk management, especially hedging spot market position via futures contract, indicate the necessity of trading in futures market by (Johnson, 1960; and Ederington, 1979). However, linear fitting with minimization of unconditional variance cause hedging risk due to the time-varying basis risk (hedging ratio). Recent progress on time-varying hedging ratio can be referred to GARCH modelling by (Park and Switzer, 1995; and Lien, 2009), regime switching model by Alizadeh and Nomikos (2004) and co-integration idea and non-matching hedging (time-to-maturity relating to spot market on futures contract is not consistent) by (Ghoddusi and Emamzadehfard, 2017) and reference inside).
In this chapter, an approach used differs from the traditional hedging strategy or recent state or regime modelling. A optimal way for trading two futures contracts simultaneously under the framework of portfolio variance minimization idea is proposed. New strategy implementation is purely dependent on the futures basis relation across different time-to-maturity range. Given the forecasted futures basis values predicted from models, a new trading strategy is implemented to capture the dynamic futures basis differential premium, which is stated in the last section of empirical analysis.

4.3 Data and Methodology

In this section, data selection and method implementation in empirical analysis are demonstrated in the following subsection respectively. Adaptive to the futures market characters, before transferring the discrete data to functional observation via basis function smoothing (Ramsay, 2006), futures data are first re-organized consistent with the literature. Functional forecasting following autoregressive method (Kargin and Onatski, 2008) and statistical test based on (Hansen et al., 2011) are illustrated in the second sections.

4.3.1 Data Description

Energy product, Light Crude Oil, from the New York Mercantile Exchange (NYMEX) with daily closing price data spanning from 01/01/1985 to 27/10/2016 are collected from the DataStream. The final in-sample training data is covering from 01/01/1985 to 15/04/1999, and left data period is for out-of-sample forecasting test, 16/04/1999 - 27/10/2016. Since futures contract data are subjected to different expiration dates, consistent with the literature, continuous time series price data is constructed by rolling over the nearest to maturity contract to the second nearest to maturity contract when the time left to expiration date of the nearest to maturity contract is less than 5 trading days.
(Koijen et al., 2013, Fuertes et al., 2015 and Fernandez-Perez et al., 2017b).

Data format is then organized in a matrix form with rows standing for time dimension daily observation and columns standing for different maturities. Roll over method is applied simultaneously for all maturities based contracts without generating any overlapping information. At the rolling date, when second nearest-to-maturity contract move to the first one, the third nearest-to-maturity contract automatically becomes the second one and this rolling continues to the last (longest) maturity contract. After the futures contracts roll-over, cross sectional number of contract is trimmed up to 20. So, time series data of futures contract prices are formatted in a matrix way and this idea is consistent with recent petroleum futures term structure forecasting study by (Chantziara and Skiadopoulos, 2008).

Incentives for selecting crude oil are mainly due to (1) perfect term structure with constant time to maturity, each maturity difference is equal to 1 month and the number of contract each year is consistent with no break (2) crude oil is "black gold" for many industries and countries and its dynamics on traded data is also crucial for market speculators and hedgers.

Futures basis, in this chapter, defined as the log price difference between two different time-to-maturity contracts with the same underlying at the same time. Consistent with formula in section E of (Routledge et al., 2000), futures basis calculation formula is represented as follows:

$$\text{Basis}_{t, \tau_j} = \log \left( \frac{F_{t, \tau_{j+1}}}{F_{t, \tau_j}} \right) \quad (4.3.1)$$

where, $t$ is the time dimension index, $\tau_j$ is maturity proxy with subscript $j$ standing for month to expiration, $j = 1, 2, 3, \ldots, 19$. Basis$_{t, \tau_j}$ is futures basis observation for at time $t$ with maturity $j$.

At each time point, 19 discrete values from futures basis calculation are obtained. Given this

---

38 Maturity over 20 introduces missing values at some time points, this study is trying to voiding interpolation for data manipulation.

39 Futures basis results cross-sectional are ended with 19 months to maturity (columns) due to price maturity trimmed ata 20 months
well-structured data frame, Dynamic Nelson-Siegel (DNS therefore in the following content) model can be easily applied to futures basis term structure data (Diebold and Rudebusch, 2013). Introducing DNS model forecasting results as one of benchmark models is inspired by some recent commodity future term structure fitting studies, see (Grønborg and Lunde (2016), Barunik and Malinska (2016) and Karstanje et al. (2017)).

4.3.2 Methodology

In this section, the general functional autoregressive model is presented in Hilbert and Banach spaces following the literature (Bosq, 2012):

\[ f_{t+1} = \rho f_t + \varepsilon_{t+1} \]  

(4.3.2)

where, \( f_{t+1} \) and \( f_t \) are functional curve observations at time \( t + 1 \) and \( t \) with mean assumed to be zero, \( \varepsilon_{t+1} \) is a strong H-white noise, \( \rho \) is the functional linear operator with

\[ \| \rho \| = \int \int \psi^2(t,s) \, dt \, ds \]  

(4.3.3)

where \( \psi(t,s) \) is a bivariate kernel in order to make \( \| \rho \| < 1 \). All functions are assumed to be elements defined on the Hilbert space \( L^2 \) of real square integral functions on the interval \([0,1]\) and inner products are satisfying \( \langle h, g \rangle = \int h(t)g(t) \, dt \).

For equation 4.3.2 to work properly, the key is about estimation of linear operator which transfers the current functional observation to the future ones. In this study, functional linear operator estimation is following recent functional auto-regression predictive factors model (PF therefore in the following contents) developed by Kargin and Onatski (2008). Different from the standard approach,
CHAPTER 4. COMMODITY FUTURES BASIS TERM STRUCTURE FORECASTING

functional principal component analysis (reducing data dimension and exploring the main functional components explaining data variation), the PF method estimate linear operator by selecting those factors that can minimize the mean squared forecasted error.

Given the purpose of minimization $E \| f_{t+1} - \rho f_t \|$, forecasting functional observation under functional predictive model is presented:

$$\hat{f}_{t+1} = \sum_{i=1}^{k} \langle f_t, \hat{b}_{\alpha,i} \rangle \hat{C}_1 (\hat{b}_{\alpha,i})$$

(4.3.4)

where,

$$\hat{b}_{\alpha,i} = \sum_{j=1}^{p} \hat{\lambda}_j^{-1/2} \langle \hat{x}_{\alpha,i}, \hat{v}_j \rangle \hat{v}_j + \alpha \hat{x}_{\alpha,i}$$

(4.3.5)

$\hat{x}_{\alpha,i}$ is a linear combinations of empirical functional principal components (EFPCs), denoted as $\hat{v}_j$, $\hat{\lambda}_j$ is the empirical eigenvalue corresponding to EFPCs in deceasing order ($\lambda_1 > \lambda_2 > \cdots > \lambda_p$), $\hat{C}_1$ is the empirical data covariance operator with lag 1, $\alpha$ and $k$ are two control parameters to be fixed before the forecasting procedure, which is computed from the in-sample training data via the cross-validation process with fold equal to 10. The predictive factor method is supposed to outperform other methods as long as the model parameters ($\alpha$ and $k$) are well specified somehow, see Didericksen et al. (2012) and Horváth and Kokoszka (2012).

Other functional candidate models for comparison are listed in the following: functional principal component based kernel estimation method (EK therefore in the following content) by Bosq (2012) and Besse et al. (2000),

$$\hat{f}_{t+1} = \sum_{k=1}^{p} \left( \sum_{l=1}^{p} \hat{\psi}_{k,l} \langle f_t, \hat{v}_l \rangle \right) \hat{v}_k(t)$$

(4.3.6)

where

$$\hat{\psi}_{k,l} = \hat{\lambda}_l^{-1} (N - 1) \sum_{n=1}^{N-1} \langle f_{t+1}, \hat{v}_j \rangle \langle f_t, \hat{v}_i \rangle$$

(4.3.7)
functional naive approach (Naive therefore in the following content) which simply uses the last period functional observation as the predicted value

\[ \hat{f}_{t+1} = f_t \]  
(4.3.8)

and functional random walk (RW therefore in the following content) which uses the last period functional observation plus a random H-White noise as an approximation inspired by the empirical functional models’ forecasting comparison study from Didericksen et al. (2012).

\[ \hat{f}_{t+1} = f_t + \varepsilon_{t+1} \]  
(4.3.9)

Apart from functional method family, motivated by recent commodity futures price term structure fitting and forecasting studies (Baruník and Malinska, 2016, Grønborg and Lunde, 2016 and Karstanje et al., 2017), non-functional forecasting method Dynamic Nelson-Siegel model (three latent factors are assumed to be vector autoregressive process with lag 1) by Diebold and Rudebusch (2013) is considered in this chapter for completeness.

\[ f_{t+1} = L_{t+1} + S_{t+1} \left( \frac{1 - e^{-\lambda \tau}}{\lambda \tau} \right) + C_{t+1} \left( \frac{1 - e^{-\lambda \tau}}{\lambda \tau} - e^{-\lambda \tau} \right) + \varepsilon_t(\tau) \]  
(4.3.10)

where, \( \tau \) is an index vector for maturities, \( L_{t+1} \) is the estimated first latent factor with observed factor is constant term which does not depends on the time, \( S_{t+1} \) is the estimated second factor with observed model factor \( \left( \frac{1-e^{-\lambda \tau}}{\lambda \tau} \right) \) in the DNS model focusing on the fitting of short time contract variation, \( C_{t+1} \) is the estimated third factor with factor loading \( \left( \frac{1-e^{-\lambda \tau}}{\lambda \tau} - e^{-\lambda \tau} \right) \) measuring the middle term of contract data variation. \( \lambda \) is a decaying parameter controlling the shape of observed model factor,
which is optimized before out-of-sample forecasting procedure. Dynamics of three latent factors are assumed to be vector autoregressive process with lag 1,

\[
\begin{align*}
\begin{bmatrix} L_{t+1} \\ S_{t+1} \\ C_{t+1} \end{bmatrix} &= \begin{bmatrix} \alpha_1 + \beta_1 L_t + \xi_{1,t+1} \\ \alpha_2 + \beta_2 S_t + \xi_{2,t+1} \\ \alpha_3 + \beta_3 C_t + \xi_{3,t+1} \end{bmatrix} \\
\end{align*}
\]

where, \( \alpha_1, \alpha_2 \) and \( \alpha_3 \) are intercept values, \( \beta_1, \beta_2 \) and \( \beta_3 \) are estimated from the in-sample training sample data and continued updated when new data observation becomes available. \( \xi_{1,t+1}, \xi_{2,t+1} \) and \( \xi_{3,t+1} \) are error terms. Although the first latent factor is often modelled as a first order difference stationary process, all factors in this study are stationary on levels which is mainly due to the stationary property on the futures basis.

Following the standard forecasting performance comparison, both overall and individual time point level due to functional data format. For overall forecasting error comparison, calculations are referred to functional root mean square error (FRMSE), functional mean absolute error (FMAE), functional Theil inequality coefficient (FTIC). For individual time level, mean square error (MSE) and mean absolute error (MAE) are used following [Kargin and Onatski (2008), Chantziara and Ski-adopoulos (2008) and Didericksen et al. (2012)].

\[
\text{FRMSE} = \sqrt{\frac{1}{N} \sum_{t=1}^{N} \left( \int_{0}^{1} (f_t(x) - \hat{f}_t(x))^2 \, dx \right)} \tag{4.3.11}
\]

\[
\text{FMAE} = \frac{1}{N} \sum_{t=1}^{N} \left( \int_{0}^{1} |f_t(x) - \hat{f}_t(x)| \, dx \right) \tag{4.3.12}
\]

\[
\text{FTIC} = \frac{\text{FRMSE}}{\sqrt{\frac{1}{N} \sum_{t=1}^{N} (\int_{0}^{1} f_t(x)^2 \, dx) + \sqrt{\frac{1}{N} \sum_{t=1}^{N} (\int_{0}^{1} \hat{f}_t(x)^2 \, dx)}}} \tag{4.3.13}
\]
CHAPTER 4. COMMODITY FUTURES BASIS TERM STRUCTURE FORECASTING

\[
MSE_t = \sqrt{\int_0^1 (f_t(x) - \hat{f}_t(x))^2 \, dx} \tag{4.3.14}
\]

\[
MAE = \left( \int_0^1 |f_t(x) - \hat{f}_t(x)| \, dx \right) \tag{4.3.15}
\]

where, \( f_t(x) \) and \( \hat{f}_t(x) \) are the real and estimated curve observation respectively, \( N \) is the number of observation in out-of-sample.

To access the out-of-sample forecasting performance statistically, the Model Confidence Set method (MCS therefore in the following contents, [Hansen et al., 2011]) is applied in this chapter. Regarding the degree of flexibility, MCS method has taken the great advantages as its interface framework can allow not only the out-of-sample forecasting error comparison but also permit more general comparison between two objectives as long as the loss function is well specified. Generally, MCS is performed via a sequential testing procedure on forecasting equivalence checking, model elimination criterion and continuing updating algorithm. Statistical test under the null hypothesis is carried out among two different models as follows:

\[
H_0 : E(d_{ij,t}) = 0, \text{ for all } i, j \in M
\]

where \( M \subset M^0, M^0 \) is the candidate model set specified before test, \( d_{ij,t} \) is the loss function differential between model \( i \) and model \( j \) at time \( t \). MCS test will stop when null hypothesis is not able to be rejected, indicating two models in this set are measured as the same performance given pre-determined loss function. The final superior models set is defined as:

\[
M^* = \{ i \in M^0 : E(d_{ij,t}) \leq 0 \text{ for all } j \in M^0 \}
\]

In this chapter, statistical test based on MCS method is conducted via different settings with controlling effects from window lengths selection and loss functions formation. Results are shown in following sections.
4.4 Empirical Analysis

All empirical data analysis results are summarized in this section from different perspectives. In terms of forecasting error, FRMSE, FMAE, FTI and MSE are reported to measure the forecasted value distance against the observed real value. To investigate the forecasted futures basis term structure shape variation in the out-of-sample period, shape analysis is then decomposed by (1) the time series sign matching on the level of individual time-to-maturity (whether the sign of forecasted futures basis is the same as the real observed one at time $t$ for each maturity contract), (2) the cross-sectional first order difference sign shape matching over time (whether backwardation or contango continues over time).\(^{40}\)

Apart from those forecasting error aspects, economic meanings behind forecasting is analysed via making trading investment on underlying assets given known forecasted futures basis. Borrowing the time series futures basis factor studies from the literature, trading strategy decision is following backwardation theory idea (Fuertes et al., 2015). Similar to their ideas, the nearest time-to-maturity contract price within each futures basis calculation pair is regarded as the spot price approximation. When futures basis is positive (negative), referred to market scenario of contango (backwardation), trading strategy is implemented via selling (buying) the nearest-to-maturity future contract. Under the functional forecasting framework, economic exploration results are explored by repeating above idea on each contract pairs constructed futures basis.

Concerning about whether forecasting suggested by these models are statistically different, the MCS method (Hansen et al., 2011) is used to select the superior model set regarding model forecasting performance for which this chapter considers both forecasting error and economic meanings under different loss functions and parameters.

\(^{40}\)For details on specific application, please refer to the section: Term Structure Shape Preserving
In the end, a new trading application and discussion on how to use forecasted futures basis from the view of practical implementation in financial market is proposed. Application is considered on futures contracts only without inclusion any information from spot market, but this can be easily adjusted.

4.4.1 Futures Basis Term Structure Summary Statistics

To give a general view about how input data are being organized, summary statistics for futures basis in terms of different maturities are reported in this part see table 4.1. According to the maturity based mean value, crude oil shows non-consistent pattern across maturities. Specifically, in table 4.1, starting from the third row, mean value is negative and continues to be negative afterwards.

Stylized fact on volatility is observed in table 4.1: a decreasing trend from the short maturity contract to the long maturity contract which is consistent with the Samuelson effect (Samuelson, 1965) as short-term maturity contract price is more volatile than long term maturity contract due to high volume of market participation in the short term one.

From the distribution moments’ calculation, futures basis has a wide range on short maturity contract compared with long maturity contract, large value on kurtosis (fail tail evidence). Skewness does generally not vary too much, with almost negative values.

4.4.2 Forecasting Performance Comparison

Before comparing the final results among models, a specific forecasting principal needs to be determined to initialize the following procedure.

Firstly, for all functional models, discrete data is smoothed by 99 Fourier Basis functions to construct curve data. New data is then demeaned to make sure that zero mean assumption is satisfied.
### Table 4.1: Summary Statistics of Future Basis Term Structure on Individual Maturity Level

<table>
<thead>
<tr>
<th>Maturity</th>
<th>Obs</th>
<th>Mean</th>
<th>SD</th>
<th>Median</th>
<th>Min</th>
<th>Max</th>
<th>Skew</th>
<th>Kurtosis</th>
<th>SE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maturity.1</td>
<td>6716</td>
<td>0.00162</td>
<td>0.01904</td>
<td>0.00319</td>
<td>-0.10562</td>
<td>0.16695</td>
<td>0.37114</td>
<td>5.82425</td>
<td>0.00023</td>
</tr>
<tr>
<td>Maturity.2</td>
<td>6716</td>
<td>0.00003</td>
<td>0.01483</td>
<td>0.00157</td>
<td>-0.07796</td>
<td>0.07876</td>
<td>-0.10856</td>
<td>1.51314</td>
<td>0.00018</td>
</tr>
<tr>
<td>Maturity.3</td>
<td>6716</td>
<td>-0.00079</td>
<td>0.01251</td>
<td>0.00050</td>
<td>-0.05407</td>
<td>0.04606</td>
<td>-0.13445</td>
<td>0.36260</td>
<td>0.00015</td>
</tr>
<tr>
<td>Maturity.4</td>
<td>6716</td>
<td>-0.00126</td>
<td>0.01119</td>
<td>0.00000</td>
<td>-0.13186</td>
<td>0.03351</td>
<td>-0.45048</td>
<td>2.88719</td>
<td>0.00014</td>
</tr>
<tr>
<td>Maturity.5</td>
<td>6716</td>
<td>-0.00148</td>
<td>0.00992</td>
<td>-0.00051</td>
<td>-0.03357</td>
<td>0.09556</td>
<td>-0.05218</td>
<td>1.49132</td>
<td>0.00012</td>
</tr>
<tr>
<td>Maturity.6</td>
<td>6716</td>
<td>-0.00169</td>
<td>0.00880</td>
<td>-0.00082</td>
<td>-0.03382</td>
<td>0.02360</td>
<td>-0.24168</td>
<td>-0.14454</td>
<td>0.00011</td>
</tr>
<tr>
<td>Maturity.7</td>
<td>6716</td>
<td>-0.00181</td>
<td>0.00805</td>
<td>-0.00104</td>
<td>-0.03039</td>
<td>0.03169</td>
<td>-0.20377</td>
<td>-0.17787</td>
<td>0.00010</td>
</tr>
<tr>
<td>Maturity.8</td>
<td>6716</td>
<td>-0.00176</td>
<td>0.00744</td>
<td>-0.00128</td>
<td>-0.04111</td>
<td>0.02066</td>
<td>-0.17790</td>
<td>-0.03924</td>
<td>0.00009</td>
</tr>
<tr>
<td>Maturity.9</td>
<td>6716</td>
<td>-0.00166</td>
<td>0.00700</td>
<td>-0.00129</td>
<td>-0.05033</td>
<td>0.01982</td>
<td>-0.17360</td>
<td>0.21521</td>
<td>0.00009</td>
</tr>
<tr>
<td>Maturity.10</td>
<td>6716</td>
<td>-0.00159</td>
<td>0.00655</td>
<td>-0.00117</td>
<td>-0.02337</td>
<td>0.02662</td>
<td>-0.12691</td>
<td>0.08033</td>
<td>0.00008</td>
</tr>
<tr>
<td>Maturity.11</td>
<td>6716</td>
<td>-0.00157</td>
<td>0.00617</td>
<td>-0.00110</td>
<td>-0.03234</td>
<td>0.04677</td>
<td>-0.05090</td>
<td>0.82453</td>
<td>0.00008</td>
</tr>
<tr>
<td>Maturity.12</td>
<td>6716</td>
<td>-0.00148</td>
<td>0.00583</td>
<td>-0.00110</td>
<td>-0.05474</td>
<td>0.03457</td>
<td>-0.21324</td>
<td>2.15106</td>
<td>0.00007</td>
</tr>
<tr>
<td>Maturity.13</td>
<td>6716</td>
<td>-0.00142</td>
<td>0.00541</td>
<td>-0.00105</td>
<td>-0.04267</td>
<td>0.02193</td>
<td>-0.15510</td>
<td>0.78585</td>
<td>0.00007</td>
</tr>
<tr>
<td>Maturity.14</td>
<td>6716</td>
<td>-0.00129</td>
<td>0.00508</td>
<td>-0.00102</td>
<td>-0.03051</td>
<td>0.02822</td>
<td>-0.08072</td>
<td>0.76535</td>
<td>0.00006</td>
</tr>
<tr>
<td>Maturity.15</td>
<td>6716</td>
<td>-0.00120</td>
<td>0.00480</td>
<td>-0.00082</td>
<td>-0.05774</td>
<td>0.02827</td>
<td>-0.37205</td>
<td>3.53935</td>
<td>0.00006</td>
</tr>
<tr>
<td>Maturity.16</td>
<td>6716</td>
<td>-0.00105</td>
<td>0.00458</td>
<td>-0.00056</td>
<td>-0.04049</td>
<td>0.03596</td>
<td>0.12955</td>
<td>2.47980</td>
<td>0.00006</td>
</tr>
<tr>
<td>Maturity.17</td>
<td>6716</td>
<td>-0.00085</td>
<td>0.00461</td>
<td>-0.00056</td>
<td>-0.02643</td>
<td>0.03549</td>
<td>0.32144</td>
<td>2.08605</td>
<td>0.00006</td>
</tr>
<tr>
<td>Maturity.18</td>
<td>6716</td>
<td>-0.00051</td>
<td>0.00497</td>
<td>-0.00053</td>
<td>-0.03216</td>
<td>0.03091</td>
<td>0.55926</td>
<td>2.67244</td>
<td>0.00006</td>
</tr>
<tr>
<td>Maturity.19</td>
<td>6716</td>
<td>-0.00018</td>
<td>0.00535</td>
<td>-0.00050</td>
<td>-0.05483</td>
<td>0.04464</td>
<td>0.79109</td>
<td>6.72323</td>
<td>0.00007</td>
</tr>
</tbody>
</table>

Summary statistics of empirical crude oil future basis term structure within whole sample period (30/01/1991 to 03/11/2016) is reported here with the first column standing for maturity date measured in month, from 1 month to maturity up to 19 months to maturity. From the second column afterwards, on the first row, data statistics are listed as: Obs. (number of observation on each time index on out-of-sample period, daily frequency), Mean (average value across all observations), SD (standard deviation), Median (median), Min (minimum), Max (maximum), Skew (third moment of sample series), Kurtosis (fourth moment of sample series), SE (standard error).
within sample period. Given demeaned functional curve data, forecasting procedure is applied. In the end, in-sample curve mean is added back to obtain the final forecasted curve data.

Before the forecasting process, optimized parameters from the cross-validation optimization in the initial training sample need to be settled (specifically, $\alpha$ and $k$ for functional predictive factors, number of empirical principal components for functional estimated kernel method and $\lambda$ value for Dynamic Nelson-Siegel method).

During the functional forecasting procedure, forecasting rules need to be determined. Four types of forecasting rules are proposed here: (1) standard in-sample forecasting test (2) use 70% of whole sample as in-sample training data and continue forecasting last 30% sample data without linear operator and mean curve re-calculation (3) given initial in-sample data, mean and coefficient are re-calculated when one new observation becomes available, at the same time, in-sample training data is expanded with new observation adding (4) given the initial in-sample training data, rolling window length fixed for 500, mean curve and linear operator are re-calculated at each fixed window data. To distinguish the best forecasting rule, functional root mean square error and functional $R^2$ are used as criterion.

Determination of forecasting rule is then operated by testing these four ideas in different subsample data. Results for forecasting error and goodness of fit are reported in table 4.2.

Regardless of values of two parameters, FRMSE indicates the best choice is Rolling method with the lowest forecasting error recorded across different samples. 70/30, also known as the naive forecasting approach, performs the worst under all cases. With consideration on PredictiveFactors and value PenalizedParameters, using out-of-sample forecasting method does not vary too much. By conclusion, Rolling is the best candidate for out later forecasting rule reference.

Given rolling window forecasting type and optimized parameters from in-sample training data,
Table 4.2: Subsample Forecasting Principals Comparisons

<table>
<thead>
<tr>
<th></th>
<th>FRMSE</th>
<th>Predictive Factors</th>
<th>Penalized Parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel A: 1989-1999</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>In-sample</td>
<td>0.01153</td>
<td>9</td>
<td>0.001</td>
</tr>
<tr>
<td>70/30</td>
<td>0.04503</td>
<td>1</td>
<td>0.001</td>
</tr>
<tr>
<td>Expanding</td>
<td>0.00876</td>
<td>7</td>
<td>0.01</td>
</tr>
<tr>
<td>Rolling</td>
<td>0.00011</td>
<td>9</td>
<td>0.0109</td>
</tr>
<tr>
<td><strong>Panel B: 1992-1999</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>In-sample</td>
<td>0.00715</td>
<td>6</td>
<td>0.001</td>
</tr>
<tr>
<td>70/30</td>
<td>0.04213</td>
<td>2</td>
<td>0.001</td>
</tr>
<tr>
<td>Expanding</td>
<td>0.0062</td>
<td>7</td>
<td>0.001</td>
</tr>
<tr>
<td>Rolling</td>
<td>0.00739</td>
<td>8</td>
<td>0.007</td>
</tr>
<tr>
<td><strong>Panel C: 2002-2006</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>In-sample</td>
<td>0.00504</td>
<td>4</td>
<td>0.005</td>
</tr>
<tr>
<td>70/30</td>
<td>0.04441</td>
<td>1</td>
<td>0.001</td>
</tr>
<tr>
<td>Expanding</td>
<td>0.00433</td>
<td>4</td>
<td>0.002</td>
</tr>
<tr>
<td>Rolling</td>
<td>0.0043</td>
<td>4</td>
<td>0.001</td>
</tr>
<tr>
<td><strong>Panel D: 2002 - 2010</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>In-sample</td>
<td>0.00571</td>
<td>4</td>
<td>0.004</td>
</tr>
<tr>
<td>70/30</td>
<td>0.04782</td>
<td>2</td>
<td>0.001</td>
</tr>
<tr>
<td>Expanding</td>
<td>0.0071</td>
<td>4</td>
<td>0.003</td>
</tr>
<tr>
<td>Rolling</td>
<td>0.00562</td>
<td>4</td>
<td>0.009</td>
</tr>
<tr>
<td><strong>Panel E: 2010 - 2014</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>In-sample</td>
<td>0.0026</td>
<td>3</td>
<td>0.003</td>
</tr>
<tr>
<td>70/30</td>
<td>0.0315</td>
<td>2</td>
<td>0.001</td>
</tr>
<tr>
<td>Expanding</td>
<td>0.00251</td>
<td>3</td>
<td>0.002</td>
</tr>
<tr>
<td>Rolling</td>
<td>0.00251</td>
<td>3</td>
<td>0.002</td>
</tr>
</tbody>
</table>

Note: This table reports forecasting error comparison for four different forecasting principals across different sum-samples. Two important parameters are validated when different forecasting principals are applied. The first column is forecasting principals: In-sample using all sample data for estimation and in-sample forecasting method, 70/30 70% training for parameter estimation and continue forecasting 30% out-of-sample method; Expanding 70% training sample and continue re-estimating parameters as new observed data available and forecasting left 30% out-of-sample data; Rolling fixed 500 observation to initialize parameter estimation and rolling forward with new parameter estimation and forecasting. From the second column to the last, it reports forecasting principals’ corresponding FRMSE (functional root means square error), PredictiveFactors (number of predictive factors used in forecasting method), PenalizedParameters (parameter \( \alpha \) in equation 4.3.4 and 4.3.5).
Forecasting procedure is completed for all models. Forecasting error distribution is shown via boxplot in figure 4.1 and average forecasting error on maturity level is in figure 4.2.

**Figure 4.1: Boxplot of Forecasting Error**

*Distribution Forecasted Error Model Comparisons*

Note: Boxplot of forecast error based on Mean Square Error (MSE) under all models (Functional Predictive Factors (PF), Functional Estimated Kernel (EK), Functional Naive Model (Naive), Functional Random Walk (RW) and Dynamic Nelson-Siegel (DNS)) are reported here for crude oil market within out-of-sample period, 1999/01/01 to 2016/12/31, daily frequency with 1 day forecasting procedure. Two different maturity lengths are considered with 18-month time-to-maturity on the top panel and 10-month time-to-maturity on the bottom panel.
From figure 4.1, forecasting error distribution for each model are illustrated in 19 months full term structure idea and first 10 months term structure idea in upper panel and lower panel separately. Conclusion is that PF method has slightly more accurate forecasting performance (less forecasting error) compared with EK method. DNS and RW method are in the worst group, while Naive falls in the middle. Results are also consistent from figure 4.2 in which PF method has the lowest forecasting error in both short and long time maturity and overlaps with EK method in the middle-term.
4.4.3 Dynamics of Forecasting Errors

To track the forecasting performance along out-of-sample period, forecasting error on each time observation is calculated via MSE, yielding a time series forecasting error. This procedure is applied on all candidate models and plotted to show forecasting performance across time. One interest is to show how models forecasting performance behaves under certain extreme market situation, especially during financial crisis period. Outperforming model is supposed to be able to control forecasting error in this scenario. This argument is from the perspective that model should has ability to adjust linear operator correspond.

Figure 4.3: Forecasting Error Dynamics

Note: this figure plots the forecasting error dynamics for all models (Functional Predictive Factors (PF), Functional Estimated Kernel (EK), Functional Naive Model (Naive), Functional Random Walk (RW) and Dynamic Nelson-Siegel (DNS)) across out-of-sample period from 1999/01/01 to 2016/12/31, daily frequency with 1 day forecasting horizon. Forecasting error is calculated based on Mean Square Error (MSE) described in methodology part. PF is selected and plotted in four panels to show dynamic comparison with other four models in a pairwise way.

Forecasting error dynamics are plotted in figure 4.3, where pair-wise comparison approach is
adopted to show relative performance with PF method (red line) against other candidate models (blue line). Overall, PF method has small variation on forecasting error dynamics across whole out-of-sample period compared with other models. PF method is also well behaved as it is supposed to be when market experiences extreme fluctuation, evidenced by lowest forecasting error across all models in earlier 2000 and recent sub-prime crisis 2008. DNS, non-functional method, marked as the worst candidate (nearly along whole out-of-sample period), peaks over 0.12 forecasting error at 2008.

4.4.4 Term Structure Shape Preserving

Forecasting futures basis term structure shape is investigated from both time-series and cross-sectional level spanning the whole out-of-sample period. On the time-series dimension, sign is matched between forecasted and real futures basis, showing whether forecasting methods preserve sign shape for single maturity based contract over time.

From table 4.3, PF, EK and Naive have consistent individual sign preserving ability across the whole out-of-sample period. High percentage value is observed for DNS model at the short time-to-maturity but falls when maturity goes longer. RW is never on the comparable case as it falls into the worst case in both short and long maturity. During the recent financial crisis period, PF outperforms nearly in call maturity cases.

On another dimension, for each curve observation, first order difference is calculated for both forecasted and real curve data. Sign matching is then conducted between these two first order differenced curve data on each time point. If all maturities’ signs of first order difference curve are matched, then a value is marked at this time point (specifically, 1 for all matching on 19-month time-to-maturity contracts \(TS 1-19\), 2 for matching on the first 9-month time-to-maturity contracts \(TS 1-9\) and 3 for matching on the last 10-month time-to-maturity contracts \(TS 10-19\)) as success of curve shape
### Table 4.3: Futures Basis Term Structure Shape Preserving on Time Series Dimension

#### Panel A: Whole Out-of-Sample Period

<table>
<thead>
<tr>
<th>Maturity</th>
<th>PF</th>
<th>EK</th>
<th>Naive</th>
<th>RW</th>
<th>DNS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.98</td>
<td>0.98</td>
<td>0.99</td>
<td>0.98</td>
<td>0.94</td>
</tr>
<tr>
<td></td>
<td>0.99</td>
<td>0.98</td>
<td>0.99</td>
<td>0.98</td>
<td>0.97</td>
</tr>
<tr>
<td></td>
<td>0.99</td>
<td>0.98</td>
<td>0.98</td>
<td>0.98</td>
<td>0.96</td>
</tr>
<tr>
<td></td>
<td>0.98</td>
<td>0.98</td>
<td>0.97</td>
<td>0.97</td>
<td>0.79</td>
</tr>
<tr>
<td></td>
<td>0.98</td>
<td>0.97</td>
<td>0.97</td>
<td>0.96</td>
<td>0.95</td>
</tr>
<tr>
<td></td>
<td>0.97</td>
<td>0.97</td>
<td>0.96</td>
<td>0.96</td>
<td>0.94</td>
</tr>
<tr>
<td></td>
<td>0.96</td>
<td>0.95</td>
<td>0.98</td>
<td>0.98</td>
<td>0.81</td>
</tr>
<tr>
<td></td>
<td>0.97</td>
<td>0.95</td>
<td>0.97</td>
<td>0.96</td>
<td>0.77</td>
</tr>
<tr>
<td></td>
<td>0.98</td>
<td>0.96</td>
<td>0.96</td>
<td>0.96</td>
<td>0.95</td>
</tr>
<tr>
<td></td>
<td>0.98</td>
<td>0.97</td>
<td>0.97</td>
<td>0.97</td>
<td>0.95</td>
</tr>
<tr>
<td></td>
<td>0.99</td>
<td>0.98</td>
<td>0.98</td>
<td>0.98</td>
<td>0.95</td>
</tr>
</tbody>
</table>

#### Panel B: Financial Crisis Period

<table>
<thead>
<tr>
<th>Maturity</th>
<th>PF</th>
<th>EK</th>
<th>Naive</th>
<th>RW</th>
<th>DNS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.98</td>
<td>0.96</td>
<td>0.98</td>
<td>0.98</td>
<td>0.93</td>
</tr>
<tr>
<td></td>
<td>0.98</td>
<td>0.97</td>
<td>0.97</td>
<td>0.97</td>
<td>0.83</td>
</tr>
<tr>
<td></td>
<td>0.98</td>
<td>0.98</td>
<td>0.97</td>
<td>0.97</td>
<td>0.87</td>
</tr>
<tr>
<td></td>
<td>0.99</td>
<td>0.99</td>
<td>0.99</td>
<td>0.98</td>
<td>0.96</td>
</tr>
<tr>
<td></td>
<td>0.98</td>
<td>0.98</td>
<td>0.98</td>
<td>0.98</td>
<td>0.81</td>
</tr>
<tr>
<td></td>
<td>0.96</td>
<td>0.93</td>
<td>0.93</td>
<td>0.93</td>
<td>0.77</td>
</tr>
<tr>
<td></td>
<td>0.96</td>
<td>0.94</td>
<td>0.93</td>
<td>0.93</td>
<td>0.77</td>
</tr>
<tr>
<td></td>
<td>0.96</td>
<td>0.96</td>
<td>0.96</td>
<td>0.95</td>
<td>0.95</td>
</tr>
<tr>
<td></td>
<td>0.96</td>
<td>0.96</td>
<td>0.96</td>
<td>0.96</td>
<td>0.96</td>
</tr>
</tbody>
</table>

Note: time series futures basis shape matching is conducted for both whole out-of-sample (from 16/04/1999 to 27/10/2016) and financial crisis period (from 30/05/2008 to 30/12/2008) in daily frequency. Maturity is measured from 1-month up to 19-month to expiration. Measurement behind this idea is to matching the sign on each maturity forecasted values against the realised observations in individually. Then count 1 when matching is achieved and 0 otherwise and sum all counts over total number of days in two different sample periods. Procedure are repeated for all candidate models (PF, EK, RW, Naive and DNS).

Their dynamics for the interests of behaviour is plotted for demonstration.
Note: based on 1 day head forecasting futures basis from all models, this figure plots the number of shape preserving times across out-of-sample period, from 1604/1999 to 27/10/2016. Shape preserving idea is to match the term structure curve from forecasted futures basis against realised one. Upward (downward) sloping is marked as contango (backwardation) market. To count the times of same curve matching, all futures basis are taken first order difference across all maturities at each daily observation. Then mark 1 when model forecasted futures basis are equal to realised futures basis on every maturity point in terms of sign, otherwise, mark 0. This tables considers three curves matching with different maturity lengths, whole curve (TS 1-19), first 9 months (TS 1-9) and last 10 months (TS 10-19). To avoid overlapping, value 1 is for matching in TS 1-19, 2 for matching in TS 1-9 and 3 for matching in TS 10-19.
If a perfect model is recorded, 4575 will be the upper boundary to be observed, meaning forecasting futures basis are exactly matched to realised futures basis at each time point. From figure 4.4, it is apparent to see that Naive model is the best on shape preserving across different scenarios. Taken Naive approach formation into account, this implies that crude oil futures’ shape are generally depended on last structure shape, with total matching days 1627, 2902 and 2362 for (TS 1-19, TS 1-9 and TS 10-19) respectively. PF and EK have similar results around 500 days matching for TS 1-19 and TS 10-19, while PF dominates other models for TS 1-9 at 1960 days. Observing matching counts across time, PF has relatively better results in recent financial crisis.

### 4.4.5 Forecasting Economic Scale

Economic scale exploration is explained by applying trading strategies on futures contracts. The fundamental idea of trading strategy is following future market backwardation and contango theory ([Litzenberger and Rabinowitz, 1995](#)) and ([Fernandez-Perez et al., 2017b](#)). According to futures basis calculation in this chapter, positive (negative) futures basis implies market contango (backwardation). Trading strategy is applied with shorting (longing) future contract correspondingly.

For completeness, under backwardation theory framework, spot price is approximated by future contract with nearest-to-maturity within each futures basis calculation pair and return is calculation based on logarithm form with position fully collateralized via nearest-to-maturity contract closing price. Single contract long and short investing principals is as follows:

\[
\begin{align*}
\text{if } & Basis_{t+1, \tau_j} > 0, \quad \text{weight} = -1, \quad \text{Return}_{t+1, \tau_j} = \log \left( \frac{F_{t+1, \tau_j}}{F_t, \tau_j} \right) \times \text{weight} \\
\text{if } & Basis_{t+1, \tau_j} < 0, \quad \text{weight} = 1, \quad \text{Return}_{t+1, \tau_j} = \log \left( \frac{F_{t+1, \tau_j}}{F_t, \tau_j} \right) \times \text{weight},
\end{align*}
\]

where, \( j = 1, 2, 3, ..., 19 \), portfolio performance is described by: (1) trading strategy return data series
first four moments with mean and volatility annualized by trading length, (2) Sharpe ratio, sortino ratio
and omega sharpe ratio for strategy’s risk adjusted compensation, (3) Cornish-Fisher VaR (Value-
at-Risk) and maximum drawdown measure strategy’s potential risk and loss. Results of portfolio
performance is reported in table 4.4 and cumulative portfolio performance under different maturities
level is shown in figure 4.5 below:
Figure 4.5: Strategy Performance based on Backwardation Theory for All Maturities

Note: this figure plots the forecasting error in terms of maturity level for all models (Functional Predictive Factors (PF), Functional Estimated Kernel (EK), Functional Naive Model (Naive), Functional Random Walk (RW) and Dynamic Nelson-Siegel (DNS)) across out-of-sample period from 01/01/1999 to 31/12/2016, daily frequency with 1 day forecasting horizon. Forecasting error is calculated based on Function Root Mean Square Error Method (FRMSE) described in methodology part. Results are shown with X-Axis standing for smoothed maturity level and Y-Axis standing for FRMSE level.
Note: performance of backwardation strategy on crude oil futures contracts is reported in this table with daily out-of-sample period from 16/04/1999 to 27/10/2016. The first row stands for time-to-maturity from 1-month to 19-month and panels in table separates forecasting models (Simple holding return for comparison, PF, EK, RW, Naive and DNS model sequentially ordered). For each panel, descriptive statistics are listed on the first column: Mean (annualized average value by multiplying sample mean with 250), StDev (standard deviation), Sharpe ratio (mean adjusted to standard deviation), SortinoRatio (mean adjusted to downside deviation with minimum acceptable return is 0), OmegaRatio (probability weighted sharp ratio), VaR (Cornish-Fisher adjusted 99% Value at Risk), % of positive months (percentage of positive return in portfolio series).

From table 4.4, simple holding returns has consistent annual mean value across past 16 years out-of-sample testing, around 6% to 7%. Compared with the benchmark holding return, all models, except for RW, present relative better performance when time-to-maturity is over 10 months (long maturity forecasting makes difference). For maturity less than 10, non-functional idea from DNS method does not show strong attractiveness due to weak strategy performance. PF, EK and Naive, on the contrary, generally beat all other models in terms of strategy performance for all maturities where short time-to-maturity (up to 2-month) is the best record.
Specifically, annualized average returns from strategies based on PF predicted values are over 13% for all maturities, with Sharpe ratio starting from 0.37 up to 0.73 recorded as the best value for all methods. However, by going through all strategies’ performance across all maturities, strategy volatility and maximum drawdown (maximum loss) are recorded with unacceptable value which can wipe out all strategy performance in the real world trading, even though there is a large improvement on the long-time maturity side from predicting models. An optimal variance reduction strategy with simultaneously buying and shorting two contracts is proposed later section to overcome this big loss and large volatility issue.

4.4.6 Forecasting Statistical Test: Model Confidence Set

To statistical conclude forecasting validity and distinguish the best forecasting models, the general model selection framework based on loss functions proposed by Hansen et al. (2011) is studied in this part. Procedure is based on T-Max statistics and bootstrap replication times are 5000. From the traditional aspect, single time series root mean square error and mean absolute error are calculated as loss functions to accommodate forecasting error idea, while from trading side, loss functions are obtained from investing strategy’s maximum drawdown (maximum cumulative loss within a certain training period) and downside deviation (return series deviation when positive value is excluded as practical traders care more about negative risk side).

\[
MaxDrawdown_{i,t} = \min\{MaxDrawdown_{i,t-1}, C(r_{i,t-D:t}) - C(r_{i,t-D-1:t-1})\}
\]

\[
DownsideDeviation_{i,t} = \sqrt{\frac{\sum_{\ell=t-D}^{t} \min\{r_{i,\ell} - MAR, 0\}^2}{n}}
\]
where, $t$ is the time index from first observation to time length $n$, $i$ is the index standing for different investing results, $r_{i,\ell}$ is the asset $i$ return at time $\ell$, $r_{i,i:j}$ is the asset $i$ return from time $i$ to $j$, $D$ is the window length for data time dimension, $C$ is a function calculating the cumulative value of the input, $MaxDrawdown_{i,t}$ is the maximum loss (minimum value) for asset $i$ at time $t$, $DownsideDeviation_{i,t}$ is the downside deviation value for product $i$ at time $t$, $t_{i,t}$ is the time series return on series $i$ at time $t$, $MAR$ (Minimum Acceptable Return) is set to be 0 in this chapter.

To control and adapt to functions’ measurements accuracy, different window lengths are settled up when computing the specific loss functions values. Window length is set to be 5, 30, 60, 90, 120 and 252 days for forecasting error measurement, from one week (short-run forecasting error) to one year (long-run forecasting error). Then 120, 252, 500, 1000 and 2000 days are used to calculate maximum loss and downside volatility because maximum loss normally stays unchanged within a short time period and volatility estimation is not accurate when sample size is small. Results are reported based on time-to-maturity, organized on the column of tables.

MCS is conducted given loss functions (MSE and MAE) and results are reported in table 4.5 and 4.6. Combining these two results, PF method has distinct advantages on forecasting error reduction with maturity before 7-month on 99% significance level. For maturity after 13-month, Naive method stands out as the best model on error reduction, indicating the fact that there is no necessity of predicting long-term maturity futures basis dynamics and the best way is to use its past one-period information. Results are robust to both MSE and MAE measurements as well as different window lengths on loss function values computation. On another side, results in terms of middle-term contracts are not easy to conclude a clear pattern on which model statistically outperforms others or whether there is a better forecasting model.

From economic point of view on discussing models’ forecasting ability, MCS loss functions is
### Table 4.5: Model Confidence Set Testing with Root Mean Square Error Loss Function

<table>
<thead>
<tr>
<th>Maturity</th>
<th>Model</th>
<th>Maturity</th>
<th>Model</th>
<th>Maturity</th>
<th>Model</th>
<th>Maturity</th>
<th>Model</th>
<th>Maturity</th>
<th>Model</th>
<th>Maturity</th>
<th>Model</th>
<th>Maturity</th>
<th>Model</th>
<th>Maturity</th>
<th>Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Pf</td>
<td>2</td>
<td>Pf</td>
<td>3</td>
<td>Pf</td>
<td>4</td>
<td>Pf</td>
<td>5</td>
<td>Pf</td>
<td>6</td>
<td>Pf</td>
<td>7</td>
<td>Pf</td>
<td>8</td>
<td>Pf</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Panel A: 5 Days Rolling RMSE</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model</td>
<td>Pf</td>
<td>Pf</td>
<td>Pf</td>
<td>Pf</td>
<td>Pf</td>
<td>Pf</td>
<td>Pf</td>
<td>Pf</td>
<td>Pf</td>
<td>Pf</td>
<td>Pf</td>
<td>Pf</td>
<td>Pf</td>
<td>Pf</td>
<td>Pf</td>
</tr>
<tr>
<td>Prob. 0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.005</td>
<td>0.123</td>
<td>0.81</td>
<td>0.099</td>
<td>1.606</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Statistics 14.126</td>
<td>15.217</td>
<td>39.572</td>
<td>25.817</td>
<td>18.514</td>
<td>22.704</td>
<td>2.804</td>
<td>1.49</td>
<td>0.2</td>
<td>0.099</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Confidence Interval 0.0616,1.9546</td>
<td>0.0611,1.951</td>
<td>0.0621,1.8733</td>
<td>0.0548,1.939</td>
<td>0.062,1.9399</td>
<td>0.0643,1.9298</td>
<td>0.0807,1.9868</td>
<td>0.0531,1.9195</td>
<td>0.0578,1.9205</td>
<td>0.0937,1.891</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Panel B: 30 Days Rolling RMSE</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model</td>
<td>Pf</td>
<td>Pf</td>
<td>Pf</td>
<td>Pf</td>
<td>Pf</td>
<td>Pf</td>
<td>Pf</td>
<td>Pf</td>
<td>Pf</td>
<td>Pf</td>
<td>Pf</td>
<td>Pf</td>
<td>Pf</td>
<td>Pf</td>
<td>Pf</td>
</tr>
<tr>
<td>Prob. 0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.079</td>
<td>0.038</td>
<td>0</td>
<td>0</td>
<td>4.992</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Statistics 15.446</td>
<td>17.489</td>
<td>51.378</td>
<td>13.936</td>
<td>17.774</td>
<td>7.67</td>
<td>1.72</td>
<td>2.453</td>
<td>2.104</td>
<td>0.099</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Confidence Interval 0.0762,2.046</td>
<td>0.0762,1.9509</td>
<td>0.0599,1.9169</td>
<td>0.0572,1.9917</td>
<td>0.0551,1.9235</td>
<td>0.0546,1.9364</td>
<td>0.0651,1.9445</td>
<td>0.071,1.9427</td>
<td>0.0476,2.0182</td>
<td>0.0636,1.9236</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Panel C: 60 Days Rolling RMSE</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model</td>
<td>Pf</td>
<td>Pf</td>
<td>Pf</td>
<td>Pf</td>
<td>Pf</td>
<td>Pf</td>
<td>Pf</td>
<td>Pf</td>
<td>Pf</td>
<td>Pf</td>
<td>Pf</td>
<td>Pf</td>
<td>Pf</td>
<td>Pf</td>
<td>Pf</td>
</tr>
<tr>
<td>Prob. 0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.05</td>
<td>0.361</td>
<td>0.847</td>
<td>0.745</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Statistics 11.31</td>
<td>11.081</td>
<td>10.337</td>
<td>9.51</td>
<td>8.665</td>
<td>4.845</td>
<td>2.023</td>
<td>0.906</td>
<td>0.194</td>
<td>0.312</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Confidence Interval 0.0697,1.9134</td>
<td>0.0593,1.9392</td>
<td>0.0878,1.9337</td>
<td>0.0769,1.9881</td>
<td>0.0661,1.9551</td>
<td>0.0595,1.9014</td>
<td>0.0612,1.9995</td>
<td>0.0641,1.935</td>
<td>0.0515,1.946</td>
<td>0.058,1.9319</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Panel D: 90 Days Rolling RMSE</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model</td>
<td>Pf</td>
<td>Pf</td>
<td>Pf</td>
<td>Pf</td>
<td>Pf</td>
<td>Pf</td>
<td>Pf</td>
<td>Pf</td>
<td>Pf</td>
<td>Pf</td>
<td>Pf</td>
<td>Pf</td>
<td>Pf</td>
<td>Pf</td>
<td>Pf</td>
</tr>
<tr>
<td>Prob. 0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.079</td>
<td>0.017</td>
<td>0.038</td>
<td>0</td>
<td>4.992</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Statistics 12.917</td>
<td>52.801</td>
<td>30.664</td>
<td>22.72</td>
<td>19.182</td>
<td>5.81</td>
<td>1.987</td>
<td>0.739</td>
<td>0.1</td>
<td>0.557</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Confidence Interval 0.0811,1.9529</td>
<td>0.0633,1.9173</td>
<td>0.0631,1.956</td>
<td>0.0657,1.9242</td>
<td>0.0662,2.0351</td>
<td>0.0827,1.9679</td>
<td>0.0569,1.9208</td>
<td>0.0613,1.9719</td>
<td>0.0473,1.9131</td>
<td>0.0551,1.9878</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Panel E: 120 Days Rolling RMSE</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model</td>
<td>Pf</td>
<td>Pf</td>
<td>Pf</td>
<td>Pf</td>
<td>Pf</td>
<td>Pf</td>
<td>Pf</td>
<td>Pf</td>
<td>Pf</td>
<td>Pf</td>
<td>Pf</td>
<td>Pf</td>
<td>Pf</td>
<td>Pf</td>
<td>Pf</td>
</tr>
<tr>
<td>Prob. 0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.079</td>
<td>0.038</td>
<td>0.847</td>
<td>0.745</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Statistics 13.6</td>
<td>14.47</td>
<td>37.841</td>
<td>24.808</td>
<td>19.216</td>
<td>21.209</td>
<td>1.701</td>
<td>0.7</td>
<td>0.024</td>
<td>0.626</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Confidence Interval 0.0651,1.9883</td>
<td>0.0675,1.9733</td>
<td>0.0629,1.9551</td>
<td>0.0583,1.9043</td>
<td>0.0598,1.9014</td>
<td>0.0678,1.9899</td>
<td>0.0595,2.0318</td>
<td>0.0692,1.9348</td>
<td>0.0711,1.9543</td>
<td>0.0576,1.9711</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: Hansen Model Confidence Set (MCS) statistical testing results based on root mean square error (RMSE) loss function is reported in this table with daily out-of-sample period from 16/04/1999 to 27/10/2016. We calculate RMSE for all candidate models (PF, EK, Naive, RW and DNS) on each maturity based time series data with rolling window: 5, 30, 60, 90, 120 and 250 days. For each testing procedure, 95% is the significance level and bootstrap replication number is 5000. Results are formatted in maturity level with first column showing: Model (best model suggested from MCS), Prob. (corresponding probability value), Statistics (T-Max statistics value) and Confidence Interval (95% confidence level from bootstrap). replaced by calculating strategy return (via single contract backwardation theory) rolling characteristics: maximum drawdown and downside deviation. By passing these rolling loss function values to
From these two tables, PF and EK methods are concluded of providing significant merits on
strategy loss reduction interchangeably. This means, for contracts in the short and long time-to-maturity, PF methods stand on the success role of demonstrating best results on maximum loss and downside volatility control while EK method shifts to this role when time-to-maturity for contracts are in the middle. This evidence for maximum drawdown loss measurement is valid at maturity less than 5 months and larger than 15 months for PF and within this range for EK method. On the side of

Note: Hansen Model Confidence Set (MCS) statistical testing results based on maximum drawdown (Max Drawdown) loss function is reported in this table with daily out-of-sample period from 16/04/1999 to 27/10/2016. Max Drawdown is calculated and reported for all candidate models (PF, EK, Naive, RW and DNS) on each maturity based time series data with rolling window: 120, 250, 500, 1000 and 2000 days. For each testing procedure, 95% is the significance level and bootstrap replication number is 5000. Results are formatted in maturity level with first column showing: Model (best Naive, RW and DNS) on each maturity based time series data with rolling window: 120, 250, 500, 1000 and 2000 days. For each testing procedure,
### Table 4.8: Model Confidence Set Testing with Downside Deviation Loss Function

<table>
<thead>
<tr>
<th>Maturity,1</th>
<th>Maturity,2</th>
<th>Maturity,3</th>
<th>Maturity,4</th>
<th>Maturity,5</th>
<th>Maturity,6</th>
<th>Maturity,7</th>
<th>Maturity,8</th>
<th>Maturity,9</th>
<th>Maturity,10</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel A: 120 Days Rolling Downside Deviation</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model</td>
<td>PF</td>
<td>PF</td>
<td>PF</td>
<td>PF</td>
<td>PF</td>
<td>EK</td>
<td>EK</td>
<td>EK</td>
<td>EK</td>
</tr>
<tr>
<td>Prob.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Confidence Interval</td>
<td>0.0667,2.0075</td>
<td>0.0575,1.0573</td>
<td>0.0482,1.9477</td>
<td>0.0631,1.9385</td>
<td>0.0596,1.9302</td>
<td>0.0605,1.9386</td>
<td>0.08251,1.9543</td>
<td>0.06951,1.9175</td>
<td>0.06911,2.0191</td>
</tr>
</tbody>
</table>

| **Panel B: 250 Days Rolling Downside Deviation** |
| Model     | PF        | PFEK       | PFEK       | PF         | PFEK       | EK         | EK         | EK         | EK         |
| Prob.     |           |            |            |            |            |            |            |            |            |
| Statistics| 2.484     | 1.925      | 2.304      | 1.723      | 4.007      | 1.744      | 6.553      | 3.089      | 1.928      |
| Confidence Interval | 0.0516,2.0273 | 0.0749,1.9702 | 0.0647,1.9168 | 0.23531,2.1336 | 0.0611,1.9601 | 0.0589,1.9862 | 0.06141,1.9371 | 0.07271,1.9567 | 0.19621,2.022 | 0.06552,0.063 |

| **Panel C: 500 Days Rolling Downside Deviation** |
| Model     | PF        | PF         | PF         | PF         | PF         | EK         | EK         | EK         | EK         |
| Prob.     |           |            |            |            |            |            |            |            |            |
| Statistics| 32.551    | 2.566      | 3.275      | 6.31       | 5.924      | 2.497      | 4.121      | 3.596      | 3.852      |
| Confidence Interval | 0.0621,1.8702 | 0.0517,2.0133 | 0.0547,1.9643 | 0.0546,1.9714 | 0.0492,1.9506 | 0.0527,1.9068 | 0.0642,1.9477 | 0.0661,1.9465 | 0.0551,1.9279 | 0.0585,1.9976 |

| **Panel D: 1000 Days Rolling Downside Deviation** |
| Model     | PF        | PF         | PF         | PF         | PF         | EK         | EK         | EK         | EK         |
| Prob.     |           |            |            |            |            |            |            |            |            |
| Statistics| 32.561    | 2.566      | 3.275      | 6.31       | 5.924      | 2.497      | 4.121      | 3.596      | 3.852      |
| Confidence Interval | 0.0621,1.8702 | 0.0517,2.0133 | 0.0547,1.9643 | 0.0546,1.9714 | 0.0492,1.9506 | 0.0527,1.9068 | 0.0642,1.9477 | 0.0661,1.9465 | 0.0551,1.9279 | 0.0585,1.9976 |

| **Panel E: 2000 Days Rolling Downside Deviation** |
| Model     | PF        | PF         | PF         | PF         | PF         | EK         | EK         | EK         | EK         |
| Prob.     |           |            |            |            |            |            |            |            |            |
| Statistics| 32.561    | 2.566      | 3.275      | 6.31       | 5.924      | 2.497      | 4.121      | 3.596      | 3.852      |
| Confidence Interval | 0.0621,1.8702 | 0.0517,2.0133 | 0.0547,1.9643 | 0.0546,1.9714 | 0.0492,1.9506 | 0.0527,1.9068 | 0.0642,1.9477 | 0.0661,1.9465 | 0.0551,1.9279 | 0.0585,1.9976 |

Note: Hansen Model Confidence Set (MCS) statistical testing results based on Downside Deviation loss function is reported in this table with daily out-of-sample period from 16/04/1999 to 27/10/2016. Downside Deviation is calculated and reported for all candidate models (PF, EK, Naive, RW and DNS) on each maturity based time series data with rolling window: 120, 250, 500, 1000 and 2000 days. For each testing procedure, 95% is the significance level and bootstrap replication number is 5000. Results are formatted in maturity level with first column showing: Model (best model suggested from MCS), Prob. (corresponding probability value), Statistics (T-Max statistics value) and Confidence Interval (95% confidence level from bootstrap).

PF, downside deviation holds the same results on short time-to-maturity (less than 5 months), while long time-to-maturity is only applied to 17-month and 18-month. All these results are robust under 99% significant and all window lengths calculation scenarios.

---

**Chapter 4. Commodity Futures Basis Term Structure Forecasting**

144
CHAPTER 4. COMMODITY FUTURES BASIS TERM STRUCTURE FORECASTING

4.4.7 Real Time Trading Application based on Forecasted Futures Basis

Up to now, model forecasting economic scale is only explored on the sign of the futures basis within the framework of contango and backwardation theory, which leaves the discussion on the future basis magnitude economic effect missing. In this section, a new idea is proposed for the sake of (1) how to use the forecasted futures basis in real world practical application and (2) to what extent does forecasting accuracy matter.

Inspired by Cecchetti et al. (1988), this future contract trading idea is conducted in the spirit of variance reduction. Different from the traditional hedging strategy which uses the future contract to cover the risk from underlying spot market transaction, new approach here is trying to find out the optimal strategy by trading futures contracts only. Performance is tested for all models’ forecasted futures basis from the previous section for comparison.

To ensure this new application purely depend on futures basis, this chapter assumes no explicitly inventory cost or convenience yield effect (Fama and French, 2016, Gorton and Rouwenhorst, 2004 and Gorton et al., 2007) in futures contract pricing model. Specifically, in the application, futures basis contains all the information among two contracts pricing relationship. For simplicity, interest rate is constant or deterministic which is far from stochastic process proposed by Schwartz (1997) and Casassus and Collin D. (2005) who price future contract by assuming interest rate is a mean-reverting process and a key factor influencing convenience yield.

All these effects are taken into account by assuming that futures basis process has been well controlled and its dynamics follows the functional autoregressive AR(1) process. Consider a maturity based futures contract for a certain product with price $F_{\tau_j}$ and time left to expire $\tau_j$. New trading application is completed by trading two future contracts simultaneously, rebalancing frequently in order to obtain the future basis risk premium and reducing the concern from basis risk (Tomek and
Peterson (2001)

Portfolio return of this trading strategy is formatted in an expectation way:

\[
E(R_{t+1, \text{strategy}}) = E(R_{t+1, \tau_1}) + h E(R_{t+1, \tau_2})
\]  (4.4.1)

where, \(E\) is the expectation operator with conditional information set time subscript \(t\) is ignored, \(h \in [-1, 1]\) is the position (ratio) for trading on the same underlying but with maturity \(\tau_2\) contract, for \(h = -1/1\), fully short/long futures contract on \(F_{t, \tau_2}\), \(E(R_{t+1, \text{strategy}})\) is the expected trading strategy return from trading activity, \(R_{t, \tau}\) is log return calculated based on the same underlying,

\[
R_{t+1, \tau_1} = \log \left( \frac{F_{t+1, \tau_1}}{F_{t, \tau_1+1}} \right)
\]  (4.4.2)

for more general application, time-to-maturity differential \((\tau_2 - \tau_1)\) is taken into account following the spirit from Koijen et al. (2013) and basis spread definition in Yang (2013),

\[
F_{t, \tau_2} = F_{t, \tau_1} \exp(Basis_{t, \tau_1} (\tau_2 - \tau_1))
\]  (4.4.3)

the strategy return is then re-written as follows,

\[
E(R_{t+1, \text{strategy}}) = \log E(F_{t+1, \tau_1}) - \log (F_{t, \tau_1+1}) + h \log E(F_{t+1, \tau_2}) - h \log (F_{t, \tau_2+1})
\]  (4.4.4)

for the sake of calculating trading position \(h\), variance minimization on portfolio return is presented

\footnote{trading strategy based on Naive model forecasting result is not tested here due to no signal generation}
here,

\[
\text{Minimize}\{\text{VAR}[E(R_{t+1,\text{strategy}})]}\} = \text{Minimize}\{\text{VAR}[\log E(F_{t+1,\tau_1}) - \log (F_{t,\tau_{t_1}+1}) + h \log E(F_{t+1,\tau_2}) - h \log (F_{t,\tau_{t_2}+1})]\} \tag{4.4.5}
\]

ignoring the variance operator and simplifying items inside square bracket,

\[
E(R_{t+1,\text{strategy}}) = \left[\log E(F_{t+1,\tau_2}) - \log E(F_{t+1,\tau_1})\right] + (h-1) \log E(F_{t+1,\tau_2}) + 2 \log E(F_{t+1,\tau_1}) + [\log (F_{t,\tau_{t_2}+1}) - \log (F_{t,\tau_{t_1}+1})] - (1+h) \log (F_{t,\tau_{t_2}+1}) \tag{4.4.6}
\]

\[
E(R_{t+1,\text{strategy}}) = \left[\log E(F_{t+1,\tau_2}) - \log E(F_{t+1,\tau_1})\right] + [\log (F_{t,\tau_{t_2}+1}) - \log (F_{t,\tau_{t_1}+1})] - 2 \left[\log E(F_{t+1,\tau_2}) - \log E(F_{t+1,\tau_1})\right] - (1+h) \log (F_{t,\tau_{t_2}+1}) + (h+1) \log E(F_{t+1,\tau_2}) \tag{4.4.7}
\]

replacing the pricing formula between futures contracts via different maturities will be,

\[
E(R_{t+1,\text{strategy}}) = E(B_{t+1,\tau_1}) (\tau_{2} - \tau_1) + B_{t,\tau_{t_1}+1}(\tau_2 - \tau_1) - 2E(B_{t+1,\tau_1}) (\tau_{2} - \tau_1) - (1+h) \log (F_{t,\tau_{t_2}+1}) + (h+1) \log E(F_{t+1,\tau_2}) \quad \text{(4.4.8)}
\]
replacing the expectation operator with forecasted value in future basis and combining same values,

\[ E(R_{t+1, \text{strategy}}) = \hat{B}_{t+1, \tau_1}(\tau_2 - \tau_1) + B_{t, \tau_1+1}(\tau_2 - \tau_1) - (1 + h) \log(F_{t, \tau_2+1}) + (h + 1) \log E(F_{t+1, \tau_2}) \]  

under the portfolio minimum variance hedging framework, the target is now constructed as,

\[ \text{Minimize}\left\{ \text{VAR}\left[ E(R_{t+1, \text{strategy}}) \right]\right\} = \] 

\[ \text{Minimize}\left\{ \text{VAR}\left[ -\hat{B}_{t+1, \tau_1}(\tau_2 - \tau_1) + B_{t, \tau_1+1}(\tau_2 - \tau_1) - (1 + h) \log(F_{t, \tau_2+1}) + (h + 1) \log E(F_{t+1, \tau_2}) \right]\right\} \]  

(4.4.9)

since \( \hat{B}_{t+1, \tau_1}, B_{t, \tau_1+1} \) and \( \log(F_{t, \tau_2+1}) \) are known at time \( t \), the only uncertainty adding to this trading strategy is due to variation from expectation of \( \tau_2 \) maturity contract log price \( \log E(F_{t+1, \tau_2}) \), therefore minimum variance target is now transferred to minimizing the variance from,

\[ \text{Minimize}\left\{ \text{VAR}\left[ E(R_{t+1, \text{strategy}}) \right]\right\} = \] 

\[ \text{DeterministicPart} + \text{Minimize}\left\{ \text{VAR}\left[ (h + 1) \log E(F_{t+1, \tau_2}) \right]\right\} \]  

(4.4.10)

it easy to see that trading position equal to -1 will automatically get rid of the unknown variation part from this strategy return.

To formalise, trading rules is organised under the condition that the expectation of the futures
basis is obtained as follows,

\[
\begin{align*}
\text{if} & \quad \text{Basis}_{t+1} \left( \tau_2 - \tau_1 \right) - \hat{\text{Basis}}_{t+1} \left( \tau_2 - \tau_1 \right) > 0, \\
\text{weights} & = \left[ 1, -1 \right], E \left( \text{Return}_{t+1, \text{strategy}} \right) = E \left( R_{t+1, \tau_1+1} \right) - E \left( R_{t+1, \tau_2+1} \right) \\
\text{if} & \quad \text{Basis}_{t+1} \left( \tau_2 - \tau_1 \right) - \hat{\text{Basis}}_{t+1} \left( \tau_2 - \tau_1 \right) < 0, \\
\text{weights} & = \left[ -1, 1 \right], E \left( \text{Return}_{t+1, \text{strategy}} \right) = -E \left( R_{t+1, \tau_1+1} \right) + E \left( R_{t+1, \tau_2+1} \right),
\end{align*}
\]

trading strategy is then applied on crude oil future market in terms of different time-to-maturity contracts based on forecasted futures basis, reported in table 4.9.

### Table 4.9: Performance of Trading Strategy on Individual Maturity Level

<table>
<thead>
<tr>
<th>Panel</th>
<th>Simple Holding Return</th>
<th>Panel A: Simple Holding Return</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>0.06</td>
<td>0.06</td>
</tr>
<tr>
<td>StDev</td>
<td>0.02</td>
<td>0.02</td>
</tr>
<tr>
<td>% of positive months</td>
<td>0.01</td>
<td>0.01</td>
</tr>
</tbody>
</table>

| Panel B: Implied Strategy Return |
| Mean  | 0.05 | 0.05 | 0.05 | 0.05 | 0.05 | 0.05 | 0.05 | 0.05 | 0.05 | 0.05 | 0.05 | 0.05 | 0.05 | 0.05 | 0.05 | 0.05 | 0.05 | 0.05 | 0.05 |
| StDev | 0.02 | 0.02 | 0.02 | 0.02 | 0.02 | 0.02 | 0.02 | 0.02 | 0.02 | 0.02 | 0.02 | 0.02 | 0.02 | 0.02 | 0.02 | 0.02 | 0.02 | 0.02 | 0.02 |
| % of positive months | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 |

| Panel C: EK Implied Strategy Return |
| Mean  | 0.04 | 0.04 | 0.04 | 0.04 | 0.04 | 0.04 | 0.04 | 0.04 | 0.04 | 0.04 | 0.04 | 0.04 | 0.04 | 0.04 | 0.04 | 0.04 | 0.04 | 0.04 | 0.04 |
| StDev | 0.02 | 0.02 | 0.02 | 0.02 | 0.02 | 0.02 | 0.02 | 0.02 | 0.02 | 0.02 | 0.02 | 0.02 | 0.02 | 0.02 | 0.02 | 0.02 | 0.02 | 0.02 | 0.02 |
| % of positive months | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 |

| Panel D: DNS Implied Strategy Return |
| Mean  | 0.03 | 0.03 | 0.03 | 0.03 | 0.03 | 0.03 | 0.03 | 0.03 | 0.03 | 0.03 | 0.03 | 0.03 | 0.03 | 0.03 | 0.03 | 0.03 | 0.03 | 0.03 | 0.03 |
| StDev | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 |
| % of positive months | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 |

Note: performance of trading strategy on crude oil futures contracts is reported in this table with panels separating forecasting models (Simple holding return for comparison, PF, EK, RW and DNS model sequentially ordered) and first row standing for different time-to-maturity. For each panel, descriptive statistics are listed on the first column with Mean is annualized, StDev (standard deviation), SharpRatio (mean adjusted to downside deviation with minimum acceptable return is 0), OmegaRatio (probability weighted sharp ratio), VaR (Cornish-Fisher adjusted 99% Value at Risk), % of positive months (percentage of positive return in portfolio series).

Comparing with simple buy-and-hold strategy return (Simple Holding Return in table), in general, annualized strategy return from trading activity does not show strong improvement except for certain
specific maturity contracts in table 4.9. Holding and rolling nearby contracts for different time-to-
maturity will make on average 6% return across over the past 16 years for crude oil market.

Consistent trading purpose, PF and EK implied futures basis indeed reduce strategy volatility to a
large extent which is reflected by high risk adjusted measure (Sharpe ratio, Sortino Ratio and Omega
Ratio). For those contracts with high risk adjusted ratio, annualized mean returns are also larger than
buy-and-hold strategy. Specifically, for crude oil product, Sharpe ratio is peaked at 8.5 for 3-month
time-to-maturity contract and on average over 1.64 across all maturities for PF method only. EK
method has also achieved 22% annual return in maximum with Sharpe ratio 4.05.

Functional Naive model is not considered in this part as signal generation step (basis difference
between forecasted and realised one) is not available, which implies no trading at all time. Functional
random walk falls into the worst performance group, which could be due to its predicting accuracy.
Dynamic Nelson-Siegel model overall does not show strong competitive ability against functional
models (mainly PF and EK). In conclusion, intuition behind this hedging strategy is attributed to risk
premium generated by corrected futures basis value prediction.

For better description of how trading activity makes difference when future basis is correctly
forecasted, strategy cumulative returns are plotted for the purpose of showing variance reduction idea
and significant trading return in figure 4.6.
Figure 4.6: Futures Contracts based New Trading Strategy Performance

Note: this figure plots the forecasting error in terms of maturity level for all models (Functional Predictive Factors (PF), Functional Estimated Kernel (EK), Functional Naive Model (Naive), Functional Random Walk (RW) and Dynamic Nelson-Siegel (DNS)) across out-of-sample period from 01/01/1999 to 31/12/2016, daily frequency with 1 day forecasting horizon. Forecasting error is calculated based on Function Root Mean Square Error Method (FRMSE) described in methodology part. Results are shown with X-Axis standing for smoothed maturity level and Y-Axis standing for FRMSE level.
Evidence is clear to document across all panels and all maturities. Compared with simple holding strategy, all forecasted futures basis based trading strategies have significant impact on return volatility reduction, witnessed by less volatile and stable return series. Meanwhile, PF model based trading strategies present non-trivial return across both time and maturity dimensions against other models. With maturity up to 3 months, PF forecasted futures basis based trading strategy shows incomparable return at each maturity point.

Another interesting point for this strategy is focused on the financial crisis period 2008 when the simple holding strategy cumulative return series experienced the biggest drop, while the new trading strategy on the contrary magnifies their returns quickly. This is supported by observing a slight return jump for each PF line in all panels, even though this becomes less visible as maturity goes longer. Trivial cumulative returns are recorded for other model suggesting trading strategies given the fact that all cumulative return lines are under the benchmark model along the whole out-of-sample period. This is also consistent due to their large forecasting error observed before.

To robust check trading application results, inspired by Goulas and Skiadopoulos (2012), trading strategy is further tested when transaction cost is taken into account following Locke and Venkatesh (1997) with per trade cost 0.066%. Computation is repeated for each maturity based trading application and results are reported in table 4.10.

Comparing with table 4.9, where no transaction cost is embedded, table 4.10 shows that the maximum cost for transaction in real trading strategy is around 12% return annually, reducing 40% annual mean return to 28% when cost is added on the shortest time-to-maturity contract. Risk adjusted compensation is still attractive peaking at 5.44 after cost control. EK method has a relatively less return drop by around 7% at short time-to-maturity contract trading. For maturity larger than 5 months, strategy returns have been wiped out completely and even negative afterwards.
in general has better performance than non-functional method. In details, functional predictive fac-
real world economic investment decision. Based on this new trading strategy proposal, purely relying
from PF method has the highest accuracy. This magnitude consideration has non trivial effect on the
sequentially ordered) respectively and first row standing for different time-to-maturity. Transaction cost is calculated based on (Locke and Venkatesh,
Note: performance of trading strategy with transaction cost is taken into account on crude oil futures contracts from 16/04/1999 to 27/10/2016 in daily
frequency. In this table, panels (A, B, C, D and E) describe forecasting models (Simple holding return for comparison, PF, EW, RW and DNS model
sequentially ordered) respectively and first row standing for different time-to-maturity. Transaction cost is calculated based on [Locke and Venkatesh
1997], per trade is 0.066%. For each panel, descriptive statistics are listed on the first column with Mean is annualized, StDev (standard deviation),
Sharp ratio (mean adjusted to standard deviation), SortinoRatio (mean adjusted to downside deviation with minimum acceptable return is 0), OmegaRatio
(probability weighted sharp ratio), VaR (Cornish-Fisher adjusted 99% Value at Risk), % of positive months (percentage of positive return in portfolio
series).

Overall, the new trading application concludes that forecast magnitude of futures returns basis
from PF method has the highest accuracy. This magnitude consideration has non trivial effect on the
real world economic investment decision. Based on this new trading strategy proposal, purely relying
on the accuracy of forecasted futures basis risk premium generation, functional forecasting method
in general has better performance than non-functional method. In details, functional predictive fac-
tors and functional estimated kernel method are outperforming other models with remarkable returns recording, especially on the short time-to-maturity contracts. These significant annual returns are mainly due to larger variation on the short time-to-maturity contract and forecasting accuracy.

4.5 Conclusion

Crude oil futures basis term structure forecasting via the functional predictive factors (PF in the following content) method shows competitive merits among other benchmark models (the functional estimated kernel, functional naive, functional random walk and Dynamic Nelson-Siegel models) proposed in the literature. In general, PF model, overall, has the lowest forecasting error value for all maturities up to 19 months. Over the past 16 years, 16/04/1999 to 27/10/2016, forecasting error dynamics are stable from PF model while certain extreme significant errors are witnessed for other models. From the idea of forecasting curve shape preserving, PF outperforms other candidate models for different maturities lengths consideration, although functional naive approach tends to be the best model on the shape holding.

Model Confidence Set, equipped with different loss functions (root mean square error and mean absolute error from the forecasting side, maximum drawdown and downside deviation from the trading strategy side), indicates that PF method has strong advantages on the short time-to-maturity (normally less than 10 months) forecasting ability. For the longer maturity, the best way to predict the futures basis is by simply looking at the past observation, named functional naive approach.

Following the backwardation theory, trading strategies given the signal generated from the forecasted futures basis shows that PF method has the best performance, beating the market holding returns. This is mainly due to individual contract sign matching. Extending futures contract trading with the variance minimization idea, this study successfully captures the forecasted risk premium
from long and short position on two futures contracts simultaneously. Results suggest the best performance from PF method, with annualized average return 40% and 28% before and after taking transaction cost into consideration separately. Sharpe ratio is also achieved and marked at the highest level 7.92 and 5.44 respectively. This is mainly because forecasted futures basis has the most forecast accuracy in PF method.

Overall, PF method with rolling window forecasting procedure successfully generate less forecasting error statically and dynamically. It also provides the best shape matching on both time-series and cross-sectional dimension. Backwardation supported trading strategy suggested by PF method outperform both simple buy-and-hold return and all other candidates. This success continues for new proposed variance minimisation trading strategy in which remarkably return and Sharpe ratio are recorded for PF model.
Chapter 5

Conclusions, Limitations and Future Researches

All current findings in this thesis are summarized here for a comprehensive conclusion, which is followed by some discussions on potential limitations from both data and model perspectives and some ideas on future research possibilities.
5.1 Conclusions

In the past decades, witnessed by the remarkable performance of investing in futures market with annualised return 12.2% and low correlation with equity market roughly -0.03 since 1969, see (Erb and Harvey, 2005, Gorton and Rouwenhorst, 2004 and Bhardwaj et al., 2015), doing academic and practical research on commodity futures becomes more attractive. The CAPM model (Sharpe, 1964, Lintner, 1965 and Mossin, 1966), originally introduced to explain the asset movement in the equity market, fails to tell the story behind the futures market movement (Jagannathan, 1985 and Erb and Harvey, 2005).

The strong market segmentation evidence between equity and futures market is documented. More specifically, market backwardation (contango) and hedging pressure theory, not from equity market, successfully capture the futures market movement. Under the asset pricing framework, above theories are transferred to risk factors pricing relation that term structure, hedging pressure as well as momentum constitutes the baseline pricing model of futures market. For these researches details, studies are referred to (Koijen et al., 2013, Erb and Harvey, 2005, Szymanowska et al., 2014, Bessembinder, 1992, De Roon et al., 2000 and Basu and Miffre, 2013).

Far more than this baseline model, new factors, e.g. idiosyncratic volatility from Fuertes et al. (2015), co-skewness factor from Junkus, 1991 and Christie-David and Chaudhry, 2001 and realised skewness factor from Fernandez-Perez et al. 2018 become prevalent recently. Among these, skewness factor is firstly argued by investment skewness preference idea in the literature via the CAPM framework Arditti and Levy (1975) as well as behaviour finance idea from Kraus and Litzenberger (1976). They state that investors incline to select positively skewed assets in their portfolio holdings as they expect extreme positive return compensation.

Inspired by the recent promising risk-neutral moments estimator pricing ability in the equity mar-
ket, this thesis first introduces the Risk-Neutral Skewness into the global commodity futures asset pricing frame. The pricing mechanisms are well discussed for the sake of accommodating the commodity futures market. Specifically, the theories of demand-based option pricing theory (Garleanu et al., 2009 and Bollen and Whaley, 2004), heterogeneous belief idea (Friesen et al., 2012 and Han, 2008) and selective hedging (Stulz, 1996) demonstrate the pricing transmission process underneath.

In addition to this single time-series assets data variation explanation study, term structure (the daily observations given different maturities) modelling, is also taken into account in this thesis. This thesis first tests the current prevalent term structure modelling method in literature, namely the Dynamics Nelson-Siegel model. Followed the functional data change point detection procedure frame from Horváth and Kokoszka (2012) and motivated by the recent projection test method by Bardsley et al. (2017), a new test statistics is proposed with well discussion on its asymptotic property, simulation and empirical studies.

Since the current proposal requires some pre-determined factors to project the functional observations, three factors from the Dynamic Nelson-Siegel model are selected. This is because this specific model has widely acceptance in the literature and clear economic interpretation behind. After identifying the model instability (mean change) in the Dynamic Nelson-Siegel three factors, this thesis moves further with novel idea of using functional autoregressive model to predict term structure dynamics.

Rather than dealing with the return data (normally in time-series data) and the term structure price data (Dynamic Nelson-Siegel model suggest, see GrØnborg and Lunde (2016), Baruník and Malinska (2016) and Karstanje et al. (2017) in futures literature), futures basis term structure (standard futures basis with calculation applied on all maturities calculation) acts as the research interest due to its strong connection to backwardation and contango theory. This thesis then uses a more natural
functional autoregressive model to fit and predict the term structure data futures basis.

To sum up, in general, there are several new contributions to the global futures market literature from different perspectives. Each part of this thesis is strongly connected with data modelling and term structure related theories’ understanding. By the meaning of it, data modelling has three branches: (1) whether the new factor is superior to the term structure factor, (2) term structure price modelling and (3) testing and term structure futures basis fitting and forecasting. Each part is closely concerned on the market backwardation and contango theory. In the meantime, methods employed are spanning from asset pricing, statistical inference and functional data process modelling.

In the first part, a new common risk factor, Risk-Neutral Skewness, is statistically documented on the global futures market under the asset pricing framework. Given the estimation process on the options market data from 2007 to 2016, Risk-Neutral Skewness does positively price the future return at least during the past 10 years.

Time-series factor exposure analysis states that the Risk-Neutral Skewness offers an extra 14.6% annual return when considering all the traditional baseline factors (e.g. term structure, momentum and hedging pressure). Practically, the new proposed Risk-Neutral Skewness does superior to both the traditional risk factors and its counterparts, the realised skewness estimated from the past historical data. Trading strategy by longing the highest Risk-Neutral Skewness group assets and shorting the lowest Risk-Neutral Skewness group assets, points out the most attractive risk adjusted performance measure, sharp at 1.39.

In the second part, research interest moves onto the futures price term structure modelling test. Modelling on the term structure price in futures market has obtained less attention and most of completed studies have focused on the extensions of the Dynamic Nelson-Siegel Model (a model is originally proposed to handle yield curve data fitting problem). Although there are similarities between
these two markets data pattern, but some heterogeneous characteristics accounts, for example, season-
ality effect, cost of carry, underlying demand and supply effect are somehow unique in the commodity
futures market.

For testing the Dynamic Nelson-Siegel modelling stability, a new statistics is developed on detect-
ing term structure modelling uncertainty (term structure mean change). This new testing procedure
naturally refers to functional frame with well discussion on its asymptotic behaviour with proof and
simulation analysis. Testing power under the alternative case (term structure mean indeed changes in
the simulated sample) shows descent results and certain consistence with the literature.

The empirical testing procedure is then applied on the futures market (crude oil and gold term
structure price data), indicating the existence of a real change point during the recent financial crisis
period. Multivariate forecasting regression is applied on the Dynamic Nelson-Siegel three factors
sample data before and after change point to figure out the driving force behind the change.

With the market participators’ trading position data from the Commodity Futures Trading Com-
mission (CFTC), long only trading position data on Producers (market participators who produce the
physical commodity products) statistically stabilise the market volatility and contribute market shift
to backwardation for gold product. For the crude oil futures market, Hedgers and Speculators have
the same impact.

In the third part, the term structure modelling is moving onto the futures basis, but considering the
log price difference on all available maturities’ contract at each observation time point. This thesis
first uses the functional data analysis method to fit the term structure discrete futures basis data and
then uses the functional autoregressive predictive factor method to forecast its dynamics.

The empirical data analysis via the crude oil futures market data shows that functional predic-
tive factor model outperforms other functional models (e.g. functional principal component analysis
or estimated kernel method) and non-functional method (e.g. the Dynamic Nelson-Siegel). The new method superiority is obtained by comparing the forecasting error, term structure shape forecast preserving, trading strategy economical intuition as well as further variance minimization strategy performance (an intra-contract trading strategy with trading signal conditional on forecasted futures basis).

5.2 Limitations and Future Researches

In terms of global commodity futures market asset pricing, Risk-Neutral Skewness estimation may potentially automatically introduce bias. One potential limitation could be the usage of Black-Scholes model when converting between market price and implied volatility mutually as Black does not truly reflect the real market data modelling. Data curve fitting process via natural and hermite cubic spline may not account for the day-to-day dynamics pattern. There might be a potential change on the fitting method, say today spline method works while tomorrow linear method works. This thesis tries to control this influence on final results with other fitting method robustness test, however, it is still worth mentioning that data mining might exists somehow.

Another potential limit is due to market data capacity as options in commodity futures market is not as active as equity and foreign exchange market. Given the filtration setting in the second chapter, errors may generate when filtered market data is too sparse (e.g. observed points have large interval). Utilising further high frequency option data could be a potential way as risk-neutral moments curve is able to be constructed and then fitting method can apply to obtain specific values.

In addition to some limits in the risk-neutral moments estimation procedure, from the view of further potential research possibilities, one extension can be conducted is to analyse the role of difference between the Risk-Neutral Skewness and the realised skewness or more generally the spread
between model-free moments and historical moments in the global futures market.

Since the risk-neutral moments are estimated in terms of market participators’ risk homogeneity and forward-looking idea, its distance relative to historical moments can reflect market participators’ expectation correction value. Pricing this differential can be potential further works, which is inspired by Kozhan et al. (2013) who argue that trading the differential between risk-neutral moments and realised moments can be interpreted as purchasing forward look moment movement with financing from historical moment value, like swap products. Rather than copying their intuition, this differential value can be referred to heterogeneous belief effect or another idiosyncratic factor risk premium.

On another hand, the historical moments can be modelled in order to accommodate forecasting property. Under this framework, with proper forecasting models (e.g. expected skewness calculation procedure (Boyer et al., 2010)), it is able to introduce the differential between risk-neutral moments and model forecasted moments as model interpretation error or model innovation. Modelling either expectation value and innovations can deep the understanding of both risk-neutral moments behaviour and forecasting model dynamics down to their roots and then contribute the literature further.

Regarding the functional mean change detection procedure, there is potential limit on test statistics as current method requires pre-determined orthogonal factors to project. The reason of employing this DNS model projection method refers to this model advantages of economical presentations. It is reasonable to extend the current scope to include more general functional mean change case without selected factors to project for the sake of statistical completeness idea.

Another limit and potential impact on the current model factors, Dynamic Nelson-Siegel, is caused by decaying factor $\lambda$. The $\lambda$ is currently set to be fixed following the literature, however, there is a possibility that $\lambda$ itself may contain change point. If this is the case, projecting functional curve data on model factors (with parameter $\lambda$ including change point) will diminish the final testing power.
Therefore, one further work will be the new statistics construction with consideration of change point on $\lambda$.

In the line of new statistics proposal, regardless of the effect of $\lambda$, new weighted functional mean change statistics is worth exploring. One idea could be the adjustment of the convergence rate while another one is concerned of these projected factors’ relative importance. The former one on weight method is applied to all factors while the later one is to put special weight on some selected factors (e.g. market participators may concern more on "Slope" factor rather than "Level" factor in Dynamic Nelson-Siegel, then new weighting function should be able to accommodate it).

For the term structure futures basis forecasting framework, one limit will be the information abundance embedded in the futures basis calculation. In terms of the futures basis calculation formula, futures basis has involved dynamics from the interest rate, cost of carry, convenience yield and other potential effects. In the current case, these impact on the dynamics of futures basis is ignored by assuming they are synchronised and represented by futures basis only.

One possibility is to isolate the interest rate term structure effect from the current futures basis term structure, in functional meaning, not discrete form. It is worth mentioning that maturities across two markets data need to be adjusted. It is guaranteed that market maturity structure in interest rate is same to that in commodity futures market. The clearer separation to have, the more useful the further dynamics modelling based on functional will be.

Another limit is the number of empirical testing product tested here. Employing crude oil has several concerns with discussion on the fourth chapter data part, but it can be possible to extend the current research to more products testing idea. In the meantime, new forecasting methods, such as the adaptive Dynamics Nelson-Siegel (Chen and Niu, 2014) and functional Dynamics Nelson-Siegel model (Hays et al., 2012), are valuable to do the further test. It is ignored here for the purpose is to
compare the performance of functional method with the benchmark and implementing these methods involves high level of complexity both in time and techniques.
Appendices
Appendix A

Risk-Neutral Estimation from BKM Method

\[
RNSK(B)_{i,t}(\tau) = \frac{e^{\tau r} W_{i,t}(\tau) - 3\mu_i(\tau)e^{\tau V_{i,t}(\tau)} + 2\mu_i(\tau)^3}{[(e^{\tau r} V_{i,t}(\tau) - \mu_i(\tau))^2]^2}
\]  
(A.0.1)

where, \(RNSK^Q_{i,t}(\tau)\) is the risk-neutral measure of skewness of underlying \(i\) at time \(t\) with \(\tau\) maturity. \(V_{i,t}(\tau), W_{i,t}(\tau)\) and \(X_{i,t}(\tau)\) are the time \(t\) prices of \(\tau\) maturity quadratic, cubic and quartic contracts, respectively. \(r\) is the risk-free rate, which is three-month treasury bill.

\[
V_{i,t}(\tau) = \int_{S_{i,t}}^\infty 2\left(1 - \ln\left(\frac{K_i}{S_{i,t}}\right)\right)K_i^2C_{i,t}(\tau,K_i)dK_i + \int_0^{S_{i,t}} 2\left(1 + \ln\left(\frac{S_{i,t}}{K_i}\right)\right)K_i^2P_{i,t}(\tau,K_i)dK_i,
\]  
(A.0.2)

\[
W_{i,t}(\tau) = \int_{S_{i,t}}^\infty 6\ln\left(\frac{K_i}{S_{i,t}}\right) - 3\left(\ln\left(\frac{K_i}{S_{i,t}}\right)\right)^2K_i^2C_{i,t}(\tau,K_i)dK_i - \int_0^{S_{i,t}} (6\ln\left(\frac{S_{i,t}}{K_i}\right) + 3\left(\ln\left(\frac{S_{i,t}}{K_i}\right)\right)^2K_i^2P_{i,t}(\tau,K_i)dK_i,
\]  
(A.0.3)
\[ X_{i,t}(\tau) = \int_{S_{i,t}}^{\infty} \frac{12(\ln(\frac{K_i}{S_{i,t}})^2 - 4(\ln(\frac{K_i}{S_{i,t}}))^3)}{K_i^2} C_{i,t}(\tau, K_i) dK_i + \int_{0}^{S_{i,t}} \frac{12(\ln(\frac{S_{i,t}}{K_i})^2 + 4(\ln(\frac{S_{i,t}}{K_i}))^3)}{K_i^2} P_{i,t}(\tau, K_i) dK_i, \]  

(A.0.4)

\[ \mu_{i,t}(\tau) = e^{r\tau} - 1 - e^{r\tau} V_{i,t}(\tau)/2 - e^{r\tau} W_{i,t}(\tau)/6 - e^{r\tau} X_{i,t}(\tau)/24, \]  

(A.0.5)

where, \( C_{i,t}(\tau, K_i) \) and \( P_{i,t}(\tau, K_i) \) are the time \( t \) prices of European out-of-money calls and puts written on the underlying product with strike price \( K \) and expiration \( \tau \) periods from time \( t \), \( S_{i,t} \) is the \( i^{th} \) underlying security’s price, in the commodity future market, standardized nearest to maturity contract price is a proxy variable.
Appendix B

Functional Change Point Detection Theorem

Proof

B.1 Proof of Theorem 3.1

It is easy to see that

\[ \sum_{\ell=1}^{m} z_{\ell} - \frac{m}{N} \sum_{\ell=1}^{N} z_{\ell} = \sum_{\ell=1}^{m} (z_{\ell} - E(z_{\ell})) - \frac{m}{N} \sum_{\ell=1}^{N} (z_{\ell} - E(z_{\ell})) + \sum_{\ell=1}^{m} E(z_{\ell}) + \frac{m}{N} \sum_{\ell=1}^{N} E(z_{\ell}), \]  \hspace{1cm} (B.1.1)

the equation (3.3.21) implies that

\[ \max_{1 \leq m \leq N} N^{-\frac{1}{2}} \left\| \sum_{\ell=1}^{m} (z_{\ell} - E(z_{\ell})) - \frac{m}{N} \sum_{\ell=1}^{N} (z_{\ell} - E(z_{\ell})) \right\| = O_p(1) \]  \hspace{1cm} (B.1.2)

and the equation (3.3.16) or (3.3.17) yields that,

\[ \max_{1 \leq m \leq N} N^{-\frac{1}{2}} \left\| \sum_{\ell=1}^{m} E(z_{\ell}) + \frac{m}{N} \sum_{\ell=1}^{N} E(z_{\ell}) \right\| \xrightarrow{p} \infty, \]  \hspace{1cm} (B.1.3)
The proof of **THEOREM 3.1** is complete.
B.2 Proof of Theorem 3.2

It is following the equation (3.3.16) that

\[ K_N(t, m, x) = (m-t)^{-\frac{1}{2}} \left( \frac{(m-t)x}{m-t} \sum_{\ell=t+1}^{[m-t]x+t} z_\ell - \frac{(m-t)x}{m-t} \sum_{\ell=t+1}^{m} z_\ell \right) \]

\[ = (m-t)^{-\frac{1}{2}} \left( \sum_{\ell=t+1}^{[m-t]x+t} (z_\ell - E(z_\ell)) - \frac{(m-t)x}{m-t} \sum_{\ell=t+1}^{m} (z_\ell - E(z_\ell)) \right) \]

\[ + g_N(t, m, x), \quad (B.2.1) \]

with,

\[ g_N(t, m, x) = (m-t)^{-\frac{1}{2}} \left( \sum_{\ell=t+1}^{[m-t]x+t} E(z_\ell) + \frac{(m-t)x}{m-t} \sum_{\ell=t+1}^{m} E(z_\ell) \right) \]

\[ (B.2.2) \]

for all \( 1 \leq t < m \leq N \). Since the following equation is satisfied,

\[ \sup_{0 \leq x \leq 1} (m-t)^{-\frac{1}{2}} \left( \sum_{\ell=t+1}^{[m-t]x+t} (z_\ell - E(z_\ell)) - \frac{(m-t)x}{m-t} \sum_{\ell=t+1}^{m} (z_\ell - E(z_\ell)) \right) \rightarrow \infty, \]

\[ (B.2.3) \]

therefore, it is easy to see that for all \( 1 \leq t < m \leq N \), \( \| g_N(t, m, x) \|^2 \) is 0 or a sequence of monotonic broken lines. The function \( \| g_N(t, m, x) \|^2 = 0 \) if and only if there is no change point between the \( t^{th} \) and \( m^{th} \) sample observation. Under the condition of the equation (3.3.17), if there is a change point between \( \lfloor Na \rfloor \) and \( \lfloor Nb \rfloor \) for \( 0 \leq a < b \leq 1 \), then,

\[ \sup_{0 \leq x \leq 1} \| g_N(\lfloor Na \rfloor, \lfloor Nb \rfloor, x) \| \rightarrow \infty, \quad (B.2.4) \]

For the sake of simplicity, assuming the \( R = 3 \) and combing the equation (B.2.2) and (B.2.3), it is able to conclude that the first step will find a change point \( \hat{r}_1 \) and this estimate is close to one of change points. Following the argument in Bardsley et al. (2017), it is easy to obtain that \( K_N(0, \hat{r}_1) \)
and $K_N(\hat{r}_1, N)$ are asymptotically independent. If there is at least one change point between the first and the $\hat{r}_1^{th}$ observation, one must be found on the account of equation (B.2.2) and (B.2.3). If there is no change point on the interval of $[1, \hat{r}_1]$, a change might be found but this probability is less than $\alpha$ asymptotically. If change is found, locating the time of change can result in the subsets $[1, \hat{r}_2]$ and $[\hat{r}_2 + 1, \hat{r}_1]$. And then the following step will test change on this interval, so continuing in this case will identify at least $R$ change points with probability closing to 1. Finding $k$ "artificial change points" cannot have larger probability than $\alpha^k$ due to the asymptotic independence of the statistics on non-overlapping.
B.3 Proof of Theorem 3.3

Following the statement from the equation (B.2.2) and (B.2.3), the estimates for the existing change points are close to the points where the function $\|g_N(t,m,x)\|$ will take its largest value. Note that $\|g_N(t,m,x)\|$ is small at “artificial change points”, e.g. where committing an error rejecting for $H_0$. It followings from equation (B.2.2) and (B.2.3) that the estimate where the corresponding test statistics is large, close to a change point and difference between them is bounded by $OP(N)$. This is the statement in THEOREM 3.3.
References


Amaya, D., Christoffersen, P., Jacobs, K. and Vasquez, A. (2011), Do realized skewness and kurtosis predict the cross-section of equity returns?, Creates research papers, Department of Economics and Business Economics, Aarhus University.


REFERENCES


REFERENCES


REFERENCES


**URL**: [https://ssrn.com/abstract=2558014](https://ssrn.com/abstract=2558014) or [http://dx.doi.org/10.2139/ssrn.2558014](http://dx.doi.org/10.2139/ssrn.2558014)


REFERENCES


