Predicting Energy Futures Prices

By

Ioanna P. Nafplioti

Submitted to

Department of Banking and Financial Management
University of Piraeus

in Partial Fulfilment of the Requirements for the Degree of

Master of Science

Supervisor

Professor George Skiadopoulos

Committee

G. Skiadopoulos **D. Voliotis** **N. Kourogenis**

Piraeus, January 2018

Abstract

Forecasting energy commodity prices is of great importance for policymakers, individuals and researchers. Using end-of-month settlement prices of the first three shortest maturity NYMEX energy futures (i.e. WTI Crude Oil, Heating Oil, and Natural Gas) over the period Jan.1990-Dec.2016, this thesis examines whether the evolution of futures log-returns can be predicted across multiple forecast horizons and, if so, by which variables. Based on three alternative linear model specifications, in-sample and out-of-sample point forecasts are generated and evaluated under different performance measures, including the modified Diebold-Mariano Test. The economic model is constructed by means of macroeconomic and financial indicators which have been found to predict the time-varying risk-premia of traditional asset classes (i.e. equities and bonds). Three joint Principal Components (PCs) are also extracted from McCracken and Ng's (2016) large macroeconomic database and used as potential predictors in a latent factor model. The results are then compared to a uni-variate autoregressive AR(1) model. While the results provide evidence of significant in-sample predictability under the economic model, the benchmark AR(1) model outperforms both the economic and the PCA models out-of-sample.

Keywords: crude oil, commodities, energy, futures, heating oil, latent factors, natural gas, predictability, principal components analysis

Acknowledgements

First and foremost, I would like to express my deep and sincere gratitude to Professor George Skiadopoulos for agreeing to supervise me, as well as for his patient guidance, valuable suggestions and constructive comments, which have all contributed to the successful completion of my thesis work.

Special thanks to my sister, Dimitra, for her encouragement to accomplish my goals throughout my academic life. Completing this work would have been even more difficult were it not for her constant support. I also wish to express appreciation to Dr. George Halkos (University of Thessaly) for inspiring my interest in Environmental Economics and Applied Econometrics, as well as to Dr. Joaquin Vespignani (University of Tasmania) for directly providing me one data series requested.

Finally, I am deeply grateful to my parents, Panagiotis and Panagiota, for their faith, unconditional love and all their efforts to ensure that I would have access to a great education.

Dedicated to my family

Table of Contents

Acknowledgements 3
Table of Contents 5
List of Tables 7
List of Figures 7
Chapter 1: Introduction 8
1.1 The Framework 8
1.2 Purpose and Research Questions 9
1.3 Energy Markets 11
1.3.1 Energy Products as Physical Commodities 11
1.3.2 Properties of Energy Markets 11
1.3.3 Energy Futures Contracts 12
1.3.3.1 West Texas Intermediate (WTI) Crude Oil Futures Specification 13
1.3.3.2 NY Harbor ULSD (Heating Oil) Futures Specification 14
1.3.3.3 Henry Hub Natural Gas Futures Specification 14
1.4 Outline 14
Chapter 2: Literature Review 15
2.1 First Stream of Literature: The Two Traditional Theories 15
2.2 Proposed Economic Predictors 17
2.2.1 Predictors of Commodity (Spot and Futures) Markets 18
2.2.1.1 Short-Term Real Interest Rates 18
2.2.1.2 Default Spread 20
2.2.1.3 Term Spread 21
2.2.1.4 US Exchange Rates 21
2.2.1.5 Financialization of Commodities 22
2.2.1.6 Emerging Economies and Global Economic Activity 23
2.2.2 Predictors of Traditional Asset Classes 25
2.3 Common Latent Factors 27
Chapter 3: The Dataset 29
3.1 Energy Futures Prices 29
3.1.1 The Dataset 29
3.1.2 Data Characteristics 30
3.2 Economic Predictors of Energy Futures 32
3.2.1 Economic Predictors 32
3.2.2 Data Characteristics 33
3.3 Descriptive Statistics and Stationarity Tests 34

3.4 Principal Component Analysis (PCA) Dataset 40
3.4.1 McCracken and Ng's Dataset 40
3.4.2 Principal Components (PCs) Characteristics 41
3.4.3 Stability of the PCA results 44
Chapter 4: Methodology 45
4.1 In-Sample Return Prediction Models 45
4.1.1 Economic Variables Model 45
4.1.2 Univariate Autoregressive Models AR(1) 47
4.1.3 Principal Components Analysis (PCA) Models 48
4.1.3.1 Description 48
4.1.3.2 PCs as Predictors – Joint PCA Models 50
4.1.4 In-Sample Evaluation Criteria 51
4.2 Out-of-Sample (OoS) Forecasting Models 52
4.2.1 Forecasting Models 52
4.2.2 Out-of-Sample Evaluation Criteria 55
4.2.2.1 Standard Performance Measures 55
4.2.2.2 Test of Equal Predictive Accuracy 56
Chapter 5: Empirical Results and Discussion 59
5.1 In-Sample Evidence 59
5.2 Out-of-Sample Forecasting Performance 65
Chapter 6: Conclusions and Implications 73
References 76

List of Tables

Table 3.1 Correlation matrix of the first three shortest and concurrent futures contracts on the NYMEX WTI Crude Oil, Heating Oil, and Natural Gas 32

Table 3.2a Descriptive statistics of the three shortest maturity futures and economic predictors in levels and first differences 37

Table 3.2b Robustness tests for the stationarity of the economic predictors 39

Table 3.3 Descriptive statistics of the three standardized PCs retained from the Joint PCA and cumulative proportion of variance explained.. 42

Table 3.4 Proportion of explained variance of highly-correlated series................................. 43

Table 3.5 Descriptive statistics of the three standardized PCs retained from the Joint PCA and cumulative proportion of variance explained within the two sub-periods............................... 44

Table 5.1 In-sample one-step ahead (h=1) individual predictive regressions......................... 62

Table 5.2 Out-of-sample performance of the various models.. 67

Table 5.3a One-step ahead (h=1 month) out-of-sample forecast accuracy: Modified Diebold-Mariano Test and Theil's U 69

Table 5.3b Three-step ahead (h=3 months) out-of-sample forecast accuracy: Modified Diebold-Mariano Test and Theil's U 70

Table 5.3c Twelve-step ahead (h=12 months) out-of-sample forecast accuracy: Modified Diebold-Mariano Test and Theil's U 71

List of Figures

Fig. 3.1 Price evolution of the three shortest maturity NYMEX WTI Crude Oil, Heating Oil, and Natural Gas futures over the period Jan.1990 to Dec.2016... 31

Fig. 3.2 Evolution of the economic predictors series over the period Jan.1990 to Dec.2016.. 34

Fig. 3.3 Stationary series of the six economic predictor from Jan.1990 to Dec.2003 39

Fig. 3.4 Returns of the three shortest maturity NYMEX WTI Crude Oil, Heating Oil and Natural Gas futures from Feb.1990 to Dec.2003 40

Fig. 3.5 Graphical representations of the first three extracted PCs from Jan.1990 to Dec.2003 41

Chapter 1: Introduction

1.1 The Framework

Today, energy commodities as a whole, and particularly petroleum products, constitute the most dominant products relative to other commodity groups which, as major non-renewable energy sources, are essential to economic growth globally.

Given that oil prices have proved to be one of the fundamental factors affecting macroeconomic aggregates more broadly (Anzuini et al., 2010; Killian, 2008; Le Pen and Sévi, 2011), it is of great importance for various institutions and policymakers, including central banks, as well as individuals and researchers to investigate the issue of whether the evolution of oil prices can be predicted and, if so, by which variables. Given their liquidity, the evolution of petroleum futures prices is frequently investigated instead; futures prices are thought to reflect the market expectations for the future level of commodity spot prices (Gorton and Rouwenhorst, 2004; Mishra and Smyth, 2016).

Furthermore, another key nonrenewable energy source is natural gas which has recently gained increasing attention in the political community, especially in view of the Kyoto protocol and its binding commitments to limit and reduce carbon dioxide emissions (Apergis and Payne, 2010). The recent literature has also proved the effect of natural gas consumption and prices on macroeconomic aggregates and economic growth. Apergis and Payne (2010) have also provided statistical support for both short-term and long-term bidirectional causality between natural gas consumption and economic growth. As a consequence, focusing on natural gas futures prices, we examine their predictability, too.

According to Gorton and Rouwenhorst (2004), although being traded for over a century, commodity futures are considered to constitute a relatively "unknown" asset class when compared to the most prominent asset classes, i.e. equities and bonds. On the whole, the interest in understanding commodity futures pricing has led to the development of two streams of academic studies. One stream (see Subsection 2.*1*) concerns two traditional theories, the so called "Risk Premium Theory" (or the Unbiased Expectations Hypothesis), and the "Theory of Storage" (also known as the Cost-of-Carry Hypothesis), which lay the foundation for the development of theoretically correct models (Gorton et al., 2012).

As for the second stream of studies, it could be further separated into two subsections, both focusing on the identification of the common underlying factors which drive the prices. Concerning the first one (Subsection *2.2*), the futures prices are examined in terms of systematic economic factors that are found to price conventional (i.e. equities and bonds) as well as alternative (i.e. commodies) investment instruments rather successfully. A wide range of macroeconomic, financial and commodity-specific factors have been examined as potential predictors in an economic model set up. However, pricing energy futures prices is a much more complex process than modeling and pricing of bonds and equities, due to some distinguishing characteristics of the broader class of energy commodities (see Subsection *1.3.2*). Therefore, energy futures are thought to behave in a strikingly different way than traditional assets and even from other commodity classes (Gorton and Rouwenhorst, 2004). To this end, alternative pricing models might be required to account for the distinct behavior of energy futures contracts. A relatively newer approach focuses on the identification and use of common latent factors in a latent factor model set up (Subsection 2.*3*).

As far as we are concerned, with few exceptions, there is a paucity of relevant literature concerning the construction of pricing models capable of reliably describing individual energy futures. There has been found only limited and inconclusive evidence concerning the predictability of petroleum futures dynamics. According to Haase and Zimmermann (2013), the controversy regarding the optimum theoretical approach and, therefore, model specification lies in the 'hybrid role' of commodities and, especially, energy products; they constitute both consumption-production assets and alternative investment instruments.

1.2 Purpose and Research Questions

In view of this relatively little and inconclusive empirical evidence on the predictability of energy futures, the aim of this study is to investigate whether the prices of petroleum and natural gas futures traded across the first three shortest maturities can be forecasted and, if so, by which factors.

This appears to be an important issue because, although supply and demand conditions are thought to be the major driving forces of commodity prices, there are other forces that may at time account for these price fluctuations. To address our research aim, we need to consider improved forecasting variables. We employ various linear model specifications and evaluate their forecasting performance both in-sample and out-of-sample on the selected energy futures returns. In brief, we construct an economic model by means of potential macroeconomic and financial business-cycle indicators that have found to be statistically significant in traditional asset pricing models[1]. Apart from the factors that possess predictive power over the time-varying risk-premia of traditional assets, we also incorporate in the same model some commodity-specific factors which have also been confirmed in the literature. In order to avoid a possible omitted variable or irrelevant variable bias and account for possible energy market's segmentation, we construct two competing predictive models; for each individual futures, we examine a univariate first-order autoregressive model as well as a latent factor model, respectively. The latter employs as potential predictors the first three principal components (PCs) extracted by a large macroeconomic database of 134 US economic indicators created by McCracken and Ng (2016) through the Principal Component Analysis (PCA). We then compare the results of the three alternative model specifications both in-sample and out-of-sample.

To this end, we address the following research questions: First, do macroeconomic, financial, and commodity-specific variables account for the energy futures prices? Second, how does this predictability vary across the energy products? Third, how does this predictability vary across maturities? Fourth, how does the futures price predictability varies according to the forecast horizon, i.e. monthly ($h=1$), quarterly ($h=3$), annual ($h=12$)? Regarding the second and the third question, such differences in predictability could reflect the forms of convenience yields as well as storage costs that are affected by macroeconomic conditions; Alquist et al. (2013) postulate that longer-maturity futures returns are driven by information associated with (unspanned) macroeconomic risks (see also Gargano and Timmermann, 2014). As for the fourth research question, Gargano and Timmermann (2014) support that bottlenecks with respect to the supply and demand of various commodities can adversely affect the evidence of futures price predictability in the short-term (i.e. short-term forecasting horizons).

Our results suggest that there is in-sample evidence of predictability; the economic variables models employed appear to fit satisfactorily almost all the energy products examined (seven out of the nine futures). The risk-factors which are found to be statically significant for almost all the futures examined are the lagged median open interest growth rate, the lagged changes in default spread, and the lagged world steel production growth rate. The highlight that the in-sample predictability rises as the corresponding futures' maturity lengthens. This predictability is also found to be strongest for the natural gas futures (1%). There is

[1] If historical futures returns on a specific market are sensitive to the same instrumental variables as are traditional asset classes such as equities and bonds, a common factor risk premia across markets may be present. Consequently, the same set of state variables can be used to forecast futures markets returns (Bessembinder and Chan, 1992).

also evidence that the second best performing models in-sample are found to be the univariate autoregressive *AR(1)*, which account for almost half of the futures (four (4) out of the nine (9) futures examined). Nevertheless, the results show that the benchmark *AR(1)* model outperforms both the Economic and the PCA models out-of-sample (OoS). In a statistically significant sense, we find that the *AR(1)* is the optimum model specification in order to model the NYMEX energy futures returns. In light of this, we claim that the NYMEX energy market is not efficient even in its weak-form, while not violating the semi-strong form efficiency; technical analysis, relevant indicators and neural networks could be used to predict future returns via the previous realized returns. On the contrary, the informational content of the remainder economic risk-factors incorporated in the economic models is not beneficial enough to predict the futures expected returns.

1.3 Energy Markets

1.3.1 Energy Products as Physical Commodities

Fundamentally, products that fall into the energy group (energy commodities) are basic products that are characterised by their physical nature; they are created by natural forces and come out of the ground. Consequently, making these commodities saleable requires that we convert them into a more usable form and move them to where they can be used, at the time they are needed.

Broadly speaking, physical commodities fall into two categories: (i) Primary commodities, which are either extracted or captured directly from natural resources with their quality and characteristics to vary widely (e.g. crude oil and natural gas) and (ii) Secondary commodities, which are produced from primary commodities to satisfy specific market needs, e.g. crude oil is refined to make various liquid petroleum products such as gasoline, heating oil, kerosene, diesel as well as other fuels.

1.3.2 Properties of Energy Markets

Energy markets, however, demonstrate some distinguishing characteristics. Among others, pronounced seasonality and mean reversion are some of these characteristics, both describing the periodic behavior of commodity price levels and volatility in short-term and long-term periods.

As highlighted by Bashiri and Lawryshyn (2017), Hamilton (2009), Le Pen and Sévi (2011), Schalck and Chenavaz (2015), and Schofield (2007) the main factors generally affecting petroleum markets are distinct supply and demand conditions. More specifically, oil prices also tend to be characterized by strong seasonal patterns, which are greatly attributed to cyclical fluctuations of supply and demand. On the demand side, these fluctuations are mostly a result of different weather conditions and climate changes. For instance, the demand for heating oil storage is high during the fall and winter; on the contrary, demand appears to be low during the spring and summer. On the supply side, hurricanes or wars in regions where oil is produced (e.g. wars in the Middle East, Iraq war in 2003 or Gulf of Mexico in 1990) could admittedly affect supply. Moreover, technological advances could also affect the supply side; for instance, horizontal drilling and hydraulic fracturing in USA led to unprecedented levels of tight (or shale) oil and gas production after 2008 (Killian, 2017). These positive supply shocks could have at least in part triggered the subsequent drop in oil prices (2014-2016).

Eventually, despite the sharp rises or decreases during short periods of such events, oil prices tend to revert to a normal level. That is to say, oil prices tend to fluctuate around and drift over time to values determined by factors such as the cost of production and the level of demand.

1.3.3 Energy Futures Contracts

Following the agreements that OPEC-member states reached during the 1970s and 1980s, a new environment in the energy markets began to take shape. The newly shaped energy markets, which now operate in a free (Mileva and Siegfried, 2012), however exposed to a variety of risks (such as geopolitical and economic developments), environment, are characterized by extreme volatility and price shifts, rendering the need of protection against market risk even more urgent. Were it not for the development of risk management method, combating these risks would not have been possible (Medlock III and Jaffe, 2009).

Like other derivative instruments, energy futures can be used to hedge against the risk of adverse price movements of energy assets. More specifically, these contracts are contractual agreements between two counterparties to either to buy or sell a defined energy product (physical commodity) satisfying a broad range of needs. The quantity and the quality of the energy product, as well as the time and the place of its delivery, are all specified by the regulated futures exchange on which the futures contract is traded.

Interestingly, over the last decades petroleum markets have evolved to the biggest commodity markets in the world, covering a range of trading activities from primarily physical to sophisticated ones and attracting an even wider range of market participants. The underlying asset can be any energy product, which is currently traded on Commodity Exchange Markets, and the contracts are traded across a wide range of maturities. In this study, we focus on energy futures listed and traded on the New York Mercantile Exchange (NYMEX) which is part of the Chicago Mercantile Exchange Group (CME Group). In particular, we consider three closely related energy products as the underlying assets: 1. NYMEX West Texas Intermediate (WTI) Light Sweet Crude Oil, 2. NYMEX NY Harbor ULSD (Heating Oil) and 3. NYMEX Henry Hub Natural Gas. Next, we provide some details regarding these specific futures contracts used in the study.

1.3.3.1 West Texas Intermediate (WTI) Crude Oil Futures Specification

Concerning the NYMEX Light Sweet Crude Oil[2], also known as West Texas Intermediate (WTI) (hereafter WTI Crude Oil), the NYMEX WTI Crude Oil futures have been trading in the exchange markets since 1983 under the ticker symbol *CL* and is considered to be the most liquid and heavily traded commodity futures contract in the world. WTI is thought to be the benchmark for the majority of crude oil transactions carried out in U.S. economy (Alquist et al., 2013; Cavalcante, 2010).

Every futures contract is written on 1,000 barrels of crude oil (contract size) and its prices is quoted in U.S. dollars per barrel (Schofield (2007). The minimum price fluctuation has been set to $0.01 per barrel (minimum tick). Though the majority of the contracts tend to be closed before expiration, physical settlement is done with the hub of U.S oil trading in Cushing, Oklahoma, serving as the delivery point. Moreover, on any given day, there are contracts trading for the next 30 consecutive months as well as contracts for delivery in 36, 48, 60, 72, and 84 month (Chantziara and Skiadopoulos, 2008). At any point in time there are 35 contracts traded in total.

[2] Crude oil's quality is determined along two dimensions: *1*. Oil's sulphur content, ranging from sweet to sour (referring to low and high sulphur content, respectively), and *2*. Oil's density, ranging from light to heavy.

1.3.3.2 NY Harbor ULSD (Heating Oil) Futures Specification

As for the NY Harbor ULSD Futures Contracts, Heating oil (also known as No. 2 fuel oil under the ticker symbol *HO*) is supposed to be one of the most important refined products which trade in NYMEX in contracts of 42,000 US gallons (equal to 1,000 barrels). Its prices are quoted in US dollars and cents per gallon and settlement is done with physical delivery (Schofield, 2007). There are also contracts available for the next 18 months ahead.

1.3.3.3 Henry Hub Natural Gas Futures Specification

In addition, Henry Hub Natural Gas started trading at the NYMEX in 1990 under the ticker symbol *NG*. The underlying asset of one futures contract is 10,000 million British thermal units (MMBtu) [3]of natural gas and its price is quoted in U.S. Dollars and cents per MMBtu, a price that constitutes the benchmark for North American natural gas. The minimum price fluctuation has been set to $0.001 per MMBtu (Schofield, 2007). Its physical settlement and delivery takes place at Henry Hub in Louisiana, a central location with a large system of pipeline interconnects.

1.4 Outline

The remainder of this study is structured as follows: In the next section (Chapter 2), we provide a brief review of relevant studies in the related literature. Chapter 3 concerns the econometric specifications of the forecasting models employed within the context of this thesis, along with the evaluation methods of their forecasting performance. Next, a brief description and an analysis of the data series to be used in the study are presented in Chapter 4. Chapter 5 presents and summarizes the empirical findings, while the final section (Chapter 6) contains the concluding remarks as well as the implications of the study, suggesting some possible avenues for future research.

[3] The British thermal unit (Btu) constitutes one of the most common measures of energy content in USA and it is defined as the amount of heat required to increase the temperature of one-pound weight of water by one degree Fahrenheit.

Chapter 2: Literature Review

In this chapter, we provide a brief overview over the two strands of literature on commodity futures pricing and predictability. The first stream of studies, which concerns the two traditional theories, is briefly described in Subsection *2.1*. Concerning the second stream of studies, it focuses on the identification of common underlying risk-factors that drive the prices of various asset classes. This stream could be further separated into two subsections. In Subsection *2.2* we report the economic (macroeconomic and financial) factors that are systematically found to price and predict conventional (i.e. equities and bonds) as well as alternative (i.e. commodities) investment instruments. Next, in Subsection *2.3* we discuss a relatively newer development in predicting the dynamics of futures markets, i.e. predicting models that employ common latent factors as potential predictors instead of predetermined economic factors.

2.1 First Stream of Literature: The Two Traditional Theories

Concerning the two traditional theories, the existence and the origin of risk premia in commodity markets trace back to the 1930s. Concerning the former theory, Keynes was the first to introduce the concept of commodity futures risk premium. In its original form, the Theory of Normal Backwardation supports that futures prices increase over time in order to reward the storage of commodities and cover its costs (Melolinna, 2011). Due to some drawbacks, however, the Keynesian Theory was further extended by many authors.

Among others, the "Market Pressure Theory" is an alternative theory developed with the scope of understanding the behavior and the interactions among the two main types of traders participating in the commodities futures markets, i.e. the traditional commercial traders (or bona fide hedgers) and the non- commercial traders (i.e. speculators and financial intermediaries) (Medlock III and Jaffe, 2009). In particular, hedgers represent all those participants who have with direct commercial interests in the underling commodities. Hence, such participants may as well be direct consumers and holders of inventories (i.e. households and firms) or producers of commodities. In order to protect themselves against potential financial losses from future short-term spot price fluctuations, the former (latter) tend to hedge their short (long) cash market positions by going long (short) in commodity futures markets (Baker, 2012; Bashiri and Lawryshyn, 2017; Gorton et al., 2012). In other words, what the hedgers achieve by their entry into the futures market is a price risk transfer towards a counterparty who, as usual, does not have any commercial interest in these commodities (e.g. intermediaries such as market dealers or speculators).

Persuading speculators and other market participants to take the opposite position in commodity futures and to willingly hold these risky instruments rather than relatively safer (risk-free) ones, requires an expected appreciation on part of them in the dollar value of the asset (net of costs) in the foreseeable future (Basu and Miffre, 2013). That is to say, the current prices of commodity futures should be set such that they include a positive risk premium[4], which reflects the compensation (expected payoff) of speculators to a futures position (Frankel, 1984; Gorton and Rouwenhorst, 2004; Hamilton and Wu, 2013). Heterogeneous preferences, however, affect the risk premia level; the more risk-averse futures market participants are, the larger the risk premium (Baker, 2012).

Overall, according to the "Market Pressure Theory", hedgers are willing to pay for this insurance against a potential price risk exposure, with this payment being expressed in the form of positive risk premium of their counterparties' positions. Actually, in the commodity futures markets the aforementioned risk premia is expressed and interpreted as expected positive excess returns earned by speculators from securing these futures positions (Alquist et al., 2013; Hamilton and Wu, 2013). That is to say, speculators buy futures contracts, say in month t, which are later sold, say in month $t+h$, enjoying a positive capital gain at time $t+h$.

On the other hand, the second predominant theory of futures pricing is that initially introduced by Kaldor, Working and Brennan. The valuation method of this ends up pricing futures depending only on the current spot price, the cost-of-carry and the convenience yield. Specifically, many consumption commodities can be consumed or physically stored (hence the description "storable"), thus impacting their allocation of consumption and physical storage over time according to the opportunity cost; the opportunity cost of today's consumption is the commodity's value the following period. According to Alquist et al. (2013), the concept of short term "convenience yield", however, reflects an implicit marginal benefit (i.e. utility) derived from holding inventories for particular purposes (e.g. consumption or speculative selling of a commodity, such as oil, in periods of extreme shortage and exorbitant prices). However, a storage cost has to be borne and any interest that could be earned on this money has

[4] Positive risk-premia in commodity futures involves the futures prices initially being set lower than the expected future (spot) price of the underlying commodity. It therefore follows that futures markets are forward looking; they embed the investors' expectations regarding the future (spot) price of the underlying asset at expiry. The futures prices are set as the conditional expectation of the future (spot) price discounted at the appropriate, however time-varying, risk premium (Gorton et al., 2012); that rewards long positions (speculators' positions) because of their respective futures prices being gradually increased over time as expiration approaches. The "backwardation", therefore, arises as a result of the futures prices initially being set lower that the expected spot prices (Haase and Zimmermann, 2013).

also to be foregone; hence, the convenience yield declines as inventory levels of the respective individual commodity are high (Gorton et al., 2012; Melolinna, 2011). . The "backwardation", therefore, arises as a result of the convenience yield levels and the following negative relationship with the physical inventory levels (Haase and Zimmermann, 2013).

Even though both approaches employ specific factors that have some clear economic interpretation, most of the assumed factors are not observable (Chantziara and Skiadopoulos, 2008) or, if they are observable, random and unpredictable, thus resulting in modeling prices using stochastic approaches (Bashiri and Lawryshyn, 2017). Nevertheless, another stream of studies has related the futures risk premia variation to a set of potential underlying determinants (Melolinna, 2011).

2.2 Proposed Economic Predictors

Relying into the second strand of literature, there is a large body of research assuming that a set of systematic risk-factors that are found to price traditional asset classes (i.e. equities and bonds) successfully should also price the commodity futures contracts. Indicatively, Melolinna (2011) and Bessembinder and Chan (1992) address similar research questions to ours; they examine whether oil futures prices can be predicted by factors previously shown to possess significant predictive power over equity and/or bond returns.

Regarding this hypothesis, recent literature confirms the co-movement of commodity markets and the presence of some common underlying drivers. Particularly, Gross (2017) argues that macroeconomic and financial business-cycle variables constitute the main drivers of the common movement in commodity markets. Gargano and Timmermann (2014) provide a brief explanation on the reason why such drivers might contribute to forecasting commodity prices. Specifically, they clarify that commodity prices are widely believed to be driven by time-varying storage costs and convenience yields, both influenced by the underlying state of the economy.

As a consequence, it could be reasonably assumed that predictors systematically found to account for the risk-premium of conventional asset classes should also possess a predictive power over movements in commodity futures markets and, subsequently, energy futures. As the literature has shown, there is a great range of macroeconomic and financial variables associated with the risk-premia in equity and bond markets which could be used as predictors in a similar asset pricing model. Next, we report all those common factors proposed and confirmed by the literature on commodity spot and futures predictability (Subsection *2.2.1*) as well as equity and bond return predictability (Subsection *2.2.2*).

2.2.1 Predictors of Commodity (Spot and Futures) Markets

Although a vast literature has examined the impact of commodity prices on macroeconomic aggregates or provide empirical evidence in favor of causality effects from oil futures to equities, far too little attention has been paid to the contrary direction (i.e. the effect of macroeconomic aggregates and financial determinants on commodity prices, such as oil prices).

From an asset pricing perspective with time-varying risk premia, the recent empirical literature has pointed that commodity prices can be determined by monetary conditions and macroeconomic risks (Alquist et al., 2013; Anzuini et al., 2010; Frankel, 1984, 2006; Schalck and Chenavaz, 2015). In general, the following factors have mostly been considered as potential predictors: commodity volatility, default return spread (also known as junk bond risk premium), equity dividend yields, industrial production, inflation, investment to capital ratio, long tern rate of returns, money stock, T-bill yields, term spread (term structure of interest rates), and unemployment (see among others Alquist et al., 2013; Bastianin et al., 2012; Bessembinder and Chan, 1992; Chen et al., 2008; Erb and Harvey, 2006; Gargano and Timmermann, 2014; Gorton and Rouwenhorst, 2004; Sadorsky, 2002; Shang, 2010). Gargano and Timmermann (2014) conclude that movements in commodity prices are partially predictable by means of the aforementioned set of determinants.

2.2.1.1 Short-Term Real Interest Rates

The short-term nominal interest rate reflects the current state of an economy, thus constituting a procyclical indicator; high rates indicate economic growth, while lower levels reflect periods of financial distress and turbulence. Specifically, in a high short-term real interest rate[5] environment, the four channels cited below reduce the demand of commodities. The corresponding lower prices of commodities reveal a negative relationship between futures and the short-term real interest rate.

On the demand side, according to Frankel and Hardouvelis (1983) and Frankel (2006), high real interest rates tend to increase the opportunity cost of storing commodities and carry-

[5] The short-term real interest rate is calculated by subtracting the expected inflation from the nominal interest rate, where the 3-month T-bill rate is the nominal interest rate.

ing inventories over into future periods. This, in turn puts downward pressure on the demand for storable commodities and the respective commodity futures prices. In line with this, Frankel (2006), Basu and Miffre (2013), and Mishra and Smyth (2016) also support that investors may prefer to pay a premium for the acquisition of a commodity at a future time rather than bearing the storage costs of acquiring the asset in the present, a case in which a contangoed[6] market is typically formed. This negative effect of the U.S. real interest rate on commodity returns has also been confirmed by Akram (2009), Bastourre et al. (2010).

On the supply side, since producers do not incur any storage costs, high interest rates decreases their incentive to leave commodities (e.g. crude oil) under the ground for future extraction. This is because any proceeds from selling the commodity at today's spot price could be invested and compounded at these higher interest rates[7]. As a result, they prefer pumping, say oil, to preserving, hence increasing the extraction and the supply of commodities today; this, in turn, leads to lower prices (Frankel, 2006, 2014; Pradhananga, 2015). This seems to apply to petroleum markets, due to the fact that there are no perfect substitute products for petroleum in the short run, and thus any fluctuations on the supply-side affect the petroleum price levels (Mileva and Siegfried, 2012).

As for the third channel, in the short term financial markets high real interest rates and the corresponding high borrowing costs discourage speculators from borrowing money and entering into the futures markets. On the contrary, the bond markets are rendered more attractive to enter into (Frankel and Hardouvelis, 1983; Frankel, 2006, 2014). Admittedly, the corresponding futures prices could fall through this third channel.

Additionally, another theoretical argument corresponding to the aforementioned decline of commodity prices has also been proposed by Barsky and Killian (2002, 2004). Based on the theory of monetary economics and asset's price determination in international financial markets, Frankel and Hardouvelis (1983) and Frankel (1984, 2014) also share the same view; rational expectations and information regarding the future monetary conditions are reflected in the current level of market prices. For instance, they suggest that tight monetary policy conditions, which increase real interest rates, form expectations of a future low-inflation environment. On the demand side, it serves as a self-fulfilling prophecy; expectation of lower than today's future spot prices could be discounted and negatively affect today's commodity prices through a lower demand level. That is because commodities are expected to cost relatively

[6] Contango refers to the situation in which the price of near month futures is lower than the price of those futures contracts expiring further in the future, thus generating an upward sloping term structure of futures prices.
[7] That is investing proceeds at interest rates higher than the expected rate of return that would be earned if they left commodities (e.g. oil) under the ground (Frankel, 2014).

less than what they cost in the present. As for speculators, lower than today's expected spot prices at the maturity of the commodity futures contract lead to futures prices being settled on a relatively higher level than that of the current spot prices (Erb and Harvey, 2006; Gorton and Rouwenhorst, 2004).

Regarding the empirical findings of the short term interest rate, Gargano and Timmermann (2014) provide evidence of modest out-of-sample predictability of the T-bill rate (however, nominal rather than real interest rate) over monthly commodity spot price indexes over the interval 1991-2010; the expected negative relationship is confirmed. Bessembinder and Chan (1992) also find limited predictive power of the T-bill over 12 futures markets.

2.2.1.2 Default Spread

Regarding the default return spread, it is commonly considered to be a measure of future business conditions and which incorporates default expectations as well as reflects the fear in the one of the conventional financial markets, i.e. the bonds market. As a countercyclical indicator, a negative relationship is expected between the default spread and the commodity returns; this is in contrast to the positive sign that is expected on equities and bonds markets (Hong and Yogo, 2012).

This is achieved by considering and comparing bonds of moderate risk (i.e. Moody's BAA corporate bonds) to the safest and most liquid ones (credited as AAA by the credit rating agency). Default spread is thus defined as the difference between the BAA bonds yields and the AAA bonds yields and also reflects the compensation for one to bear the risk of holding relatively riskier investment instruments. As a result, associated with the fear in traditional financial markets, default spread if found to exhibit a countercyclical pattern, rising during recessions and decreasing in economic expansions. That is because BAA bonds yields tend to increase (decrease) sharply during recessions (expansions). Consequently, default spread capturing fluctuations and business-cycle phases, may contribute to analyzing time-varying risk premia of traditional and alternative financial assets.

Gargano and Timmermann (2014) provide evidence of modest out-of-sample predictability of default return spread over monthly commodity spot price indexes during 1991-2010, however, with a positive sign. Bessembinder and Chan also provide evidence of the limited predictive power of the junk bond premium over 12 futures markets. On the other hand, Hong and Yogo (2012) find significant evidence of commodity returns predictability by the yield spread.

2.2.1.3 Term Spread

A number of studies have also documented the predictive power of the term spread over stock and commodity returns. That is because, as supported by Fama and French (1989), term spread reflects the compensation required to invest in longer-term interest sensitive financial assets, as it embeds longer-term inflation expectations and expectation for the short-term interest rates.

Specifically, the term spread is defined as the difference between yields on long-term (i.e. the US 10-year treasury note) and short-term (i.e. the 3-month T-bill) government securities, reflecting the slope of the term structure of interest rates (or yield curve). Term spread has been found capable of forecasting potential changes in economic activity and predicting recessions at a range of 6 to 12 month-horizons (Estrella and Hardouvelis, 1991), which, in turn, cause increased volatility in futures market prices. Moreover, empirical evidence also supports the countercyclical pattern of the term spread; term spread rises (decreases) during recession (expansive) periods, because short-term rates are commonly lowered (increased) in recessions (booms). As a consequence, term spread as a potential driving factor is considered to be capable of capturing fluctuations and business-cycle phases, therefore predicting and delivering countercyclical, hence time-varying, risk premia of financial assets (Alquist et al., 2013).

2.2.1.4 US Exchange Rates

Furthermore, the impact of exchange rates on commodity prices has also been examined and appears in the literature.

Authors argue that currency market is rather efficient and embeds forward looking information and reflect expectations regarding future macroeconomic fundamentals and price movements (see, among others, Chan et al., 2011; Chen et al., 2008). A positive relation has been proposed and should be expected between the exchange rates and the futures returns; investors demand higher expected returns (premium) as a compensation for their exposure to the exchange rate risk. Indicatively, Chen et al. (2008) verify the out-of-sample robust predictive power of bilateral US dollar exchange rates (relative to other exporting countries) over a commodity spot price index (see also Erb and Hervey, 2006). In line with this, Gargano and Timmermann (2014) and Reboredo and Revera-Castro (2013) also confirm the impact of exchange rates on oil prices.

Specifically, energy products (e.g. oil) and other primary commodities have mostly been invoiced in vehicles currencies when traded internationally in exchanges, the predominant one being the US dollar[8]. Subsequently, a decrease in the value of the US dollar with respect to foreign currencies (let it be expressed as a rise in the $S_{USD/foreign}$) is expected to be followed by an increase in commodity prices (see Akram, 2009; Bastianin, et al., 2012; Haase and Zimmermann, 2013; Sadorsky, 2002; Schalck and Chenavaz, 2015; Schofield, 2007; Shang, 2010). Employing the Trade-Weighted Dollar Index in a five-factor models over 2007-2016, Cummins et al. (2016) also provide empirical evidence of its significance and positive effect on the shortest maturity oil futures.

From the demand side point of view, that is rather reasonable since commodities expressed in foreign local currencies become relatively less expensive, hence more attractive to non-US consumers. This in turn drives demand and the corresponding commodity prices up. On the supply-side, producers experience reduced profits. As a result, they reduce their production and, hence, the overall supply of the commodity, which in turn drives prices up. Last but not least, the depreciated US dollar renders the dollar-denominated conventional financial assets less profitable (in terms of lowered return). On the contrary, this makes alternative assets, such as energy commodities, a more attractive financial instrument to invest into.

2.2.1.5 Financialization of Commodities

Moreover, Frankel (2014), Pradhananga (2015) and a significant number of other researchers allude to the fact that commodity-specific shocks are not sufficient enough to account for the 2007-2008 price surge and the 2008-2011 synchronized price movement within unrelated [9] commodity markets (see also Gross, 2017). In light of this, other underlying common drivers, thought to simultaneously affect multiple commodity markets, are investigated instead.

To this end, apart from the identifiable economic factors discussed in the Subsections *2.2.1.1 - 2.2.1.4*, financialization of commodity markets is an issue that has attracted increasing attention, especially after the early 2000s (Bashiri and Lawryshyn, 2017). Financialization refers to the voluminous, massive, and rapid financial trading of alternative assets, i.e. commodities, especially in periods of financial distress and economic uncertainty. That is because

[8] For the sake of enhanced price comparability and homogeneity, most of the commodities, and especially energy products such as crude oil, are traded internationally denominated in US dollars. That is because of the US economy's stability along with the US financial markets' depth (Mileva and Siegfried, 2012).

[9] Commodities are said to be related when they are substitute or complement commodities in either production process or consumption (Cummins et al., 2016; Pradhananga, 2015).

commodities are considered as relatively safer investment vehicles than conventional financial assets. Due to these diversification benefits, enhanced market liquidity and speculative demand shocks are supposed to artificially drive commodity prices up (Erb and Harvey, 2006; Hamilton and Wu, 2013). To this end, speculation has been proposed as a potential determinant of futures prices (see Medlock III and Jaffe, 2009; Pradhananga, 2015) and, especially, of oil prices or oil futures prices (Alquist et al., 2013; Le Pen and Sévi, 2011). However, the recent empirical findings are found to be partially contradictory.

On the one hand, Hong and Yogo (2012) confirm the predictive power of aggregate futures market dollar open interest[10] over commodity futures and bond returns, as well as currency and stock markets. They claim that dollar open interest is capable of explaining time-varying risk premia, because it is a procyclical indicator which embeds further information regarding the futures economic activity, inflation and economic conditions. Expectations of higher economic activity and demand increase the oil producers and consumers' incentive to obtain short and long positions in futures markets, respectively. This in turn drives the open interest, the futures prices, and the respective expected returns up.

Furthermore, Pradhananga (2015) also reports significant evidence of financialization of commodity futures market using two proxies, one being the sum of open interest across commodity futures markets (i.e. Total Open Interest in US dollars) which reflects the money inflow into the futures market. On the contrary, Bastianin et al. (2012) provide evidence that Working's T index (viz. a measure of excess speculation) is not statistically significant in explaining the commodity risk-premiums. Cummins et al. (2016) studying the predictability of oil futures returns over 2007-2016 also find that speculation is not a statistically significant factor in their fundamental five-factor model.

2.2.1.6 Emerging Economies and Global Economic Activity

Furthermore, many other commentators indicated that global economic activity and, especially, a possible demand "shock" stemming from important emerging and newly-industrialized economies could have at least in part triggered the previous price hike (see Frankel, 2014; Hamilton, 2009; Pradhananga, 2015 etc.). That is because such economies, viz. BRICS and Asian economies (particularly China, India, South Africa, Russia and Brazil), experienced a

[10] First, dollar open interest for each futures contract is calculated as the spot price times the correspondind open interest (viz. the amount of each futures contracts outstanding) and then the summation of the dollar open interest is taken.

relatively rapid recovery from the 2007-2009 global recession and had positive future prospects. At the same time, Killian (2008) interprets the sustained upward pressure on crude oil prices since 2003 as a consequence of the sustained strong demand driven by the global economic growth, and mostly, in Asia. Similarly, the abrupt economic slowdown underwent by the Asian economies and Russia could have at least in part caused the downward trend of oil prices experienced during 1997-1999 (He et al., 2010).

Meanwhile, several studies in the literature have further extended the aforementioned set of commodity predictors by including variables that capture macroeconomic aggregates outside the U.S. (e.g. global economic activity). That is because commodities constitute an integral part of global economy and are traded globally on a large scale. For instance, Le Pen and Sévi (2011) consider macroeconomic variables from developed and emerging economies rather than limiting their study to the U.S. economy exclusively. The authors conclude that this oil price surge was mainly attributed to the global supply and demand conditions (e.g. mostly stemming from emerging economies) and expectations about global economic activity rather than to the increased level of speculation.

Moreover, Frankel (2014) claims that demand of storable commodities (for the scope of being kept as inventories) lies behind economic activity; many commodities and, especially, energy commodities are typically used as inputs in the production process. As a result, increased economic activity today or expectations of future activity account for higher demand for storable commodities and, hence, their corresponding prices, thus revealing a positive relationship to be expected. Frankel (2014) also mentions that, although the Chicago Fed National Activity Index (CFNAI), US GDP growth, and industrial production indices growth are variables usually applied as a proxy for the US economic activity and future conditions[11], measures at a global level would be more appropriate in order to account for global shifts in demand and inventory holdings, something generally not taken into account by Gargano and Timmermann apart from two commodity related US dollar exchange rates (see also Groen, 2014).

Since such measures of global real economic activity are not readily available, literature has shown that the most widely used indicators by empirical researchers are the IMF's World real GDP and the OECD's index of world industrial production (aggregated industrial production for 34 OECD countries). However, GDP data are only available at a quarterly fre-

[11] Faccini et al. (2017) has developed a new predictor a US real economic activity, the S&P 500 option Implied Relative Risk Aversion (IRRA), which is found to predict US rea both in-sample and out-of-sample.

quency and OECD's index presents some drawbacks[12] as well. In light of this, alternative measures of global economic activity have been proposed in the literature. For instance, The MSCI Emerging Markets Index has also been considered by Gross (2017) as a proxy for the economic conditions in various emerging markets.

Killian (2008) proposed a monthly index, i.e. Killian's index of global real economic activity or the Killian's rea[13], which captures the global activity in industrial commodity markets and the corresponding cumulative demand pressures by a global index of dry cargo single-voyage ocean freight rates (see also Apergis and Miller, 2009; He et al., 2010; Killian, 2008; Ravazzolo and Vespignani, 2016, 2017). He argues that the increasing (decreasing) global demand drives these freight rates up (down), and hence the index is capable of capturing fluctuations of global real economic activity and the corresponding expansionary (recessionary) periods.

Furthermore, Ravazzolo and Vespignani (2016, 2017) also proposed another monthly measure of global real economic activity, i.e. the World steel production. Comparing their indicator to the Killian's index in an out-of-sample analysis for the World GDP, they conclude that both indicators are found to be sufficiently accurate. Ravazzolo and Vespignani (2017) also demonstrate strong evidence of the predictive power of their index over crude oil prices. Nevertheless, different disadvantages[14] also seem to apply to World steel production.

2.2.2 Predictors of Traditional Asset Classes

After having reviewed the key findings of the extant literature on commodity return predictability, we proceed with summarizing the main findings of the voluminous literarure on the long-standing debate regarding the predictability of conventional asset classes (e.g. stock indices, bonds, currencies etc.) by means of traditional asset pricing models. Specifically, we primarily focus on the cases of equity (equity risk premium) and bond return predictability.

To start with, a wide range of financial variables have been identified that display significant in-sample and out-of-sample predictive ability over equity risk premia (future

[12] Killian states that the OECD's index does not consider major emerging economies including China, India etc. Meanwhile, Ravazzolo and Vespignani (2016) argue that the OECD's index can be a good measure only on the grounds that the 34 OECD countries and the industrial-manufacturing sector represent satisfactorily the World Economy and each country's full economy, respectively. However, they comment that these prerequisites were met only until 1990.

[13] Killian's index is available on http://www-personal.umich.edu/~lkilian/paperlinks.html.

[14] Ravazzolo and Vespignani's critique of their newly-developed index is its being based exclusively on one commodity (i.e. steel), an important input.

stock returns), an evidence in favor of what is called "time-varying risk premia" (Gargano, 2013 etc.). Based on Goyal & Welch (2008), Kolev and Karapandza (2017), and Rapach and Wohar (2006) among others, we find that the most prominent predictors are variables primarily related to stocks (such as valuation ratios) along with other macro and interest-rate related variables. Based on an extensive body of empirical literature, we comprehensively summarize the macro-financial factors that have been widely found to predict premia: default yield spread or dfs (Campbell,1987; Fama and French, 1989; Goyal & Welch, 2008; Kolev and Karapandza, 2017), inflation rate or $infl$ (Goyal & Welch, 2008; Kolev and Karapandza, 2017), long-term yield or ltr (Goyal & Welch, 2008; Kolev and Karapandza, 2017), short-term interest rate or T-bill rate, tbl (Campbell,1987; Campbell and Thompson, 2008; Goyal & Welch, 2008; Kolev and Karapandza, 2017), and term spread or tms (Campbell,1987; Campbell and Thompson, 2008; Fama and French, 1989; Goyal & Welch, 2008; Kolev and Karapandza, 2017).

However, Goyal and Welch (2008) show that some of these variables fail to record a robust out-of-sample predictive ability over equity returns. On the contrary, Kolev and Karapandza (2017) point that most of these commonly employed variables exhibit a rather satisfactory out-of-sample predictability for the excess stock market returns. Moreover, Kostakis et al. (2015) using the same set of traditional forecasting variables over the period 1927-2012 also confirm their significant predictive power with respect to 1-step-ahead returns. Campbell and Thompson (2008) also confirm the satisfactory in-sample and out-of-sample predictive ability of several standard predictors over simple stock returns (i.e. T-bill, and term spread) once weak restrictions are imposed on the signs of coefficients and return forecasts. Furthermore, Rapach et al. (2013) also conclude that the short interest index is arguably the strongest predictor of equity premium, exhibiting stronger predictive power than the standard predictors of Goyal and Welch both in-sample and out-of-sample.

Gargano (2013), examining the equity return predictability in recessions, ends up with evidence in favor of highly predictable returns during recessions, while largely unpredictable during economic expansions. Particularly, cyclical variations in predictors generate cyclical (or time-varying) risk premia. Based on the concept of time-varying risk premium and its linkage to the states of economy, Gargano reports a countercyclical pattern for term and default spreads, which tend to be increasing in recessions and decreasing during expansions. He explains that this is so due to the fact that T-bills are in general lowered during recessions, thus driving up the term spread. Concerning the default spread counter-cyclicality, the default spread measures the so-called flight-to-quality which is mainly associated with fear in financial markets.

Regarding the second most prominent asset class, i.e. the bond, the variables that have been empirically found to display significant in-sample and out-of-sample predictive ability over bond returns have already been discussed in previous subsections. To this end, we highlight few alternative approaches. Indicatively, Gargano et al. (2016) make an attempt to empirically study the predictability of expected bill returns in an out-of-sample analysis. In order to account for time-varying potential factors, they use the Fama-Bliss forward spreads, forward rates and a latent macroeconomic factor extracted based on Ludvigson and Ng's (2009) methodology; Ludvigson and Ng had also extracted few latent factors to examine the predictability of bond excess returns. They both end up documenting significant stastistical evidence in favor of bond return predictability. More details regarding the latent factors are provided in the next subsection (*2.3*).

2.3 Common Latent Factors

Even though the variables described in Subsection *2.2* have some clear economic interpretation, the selection process of the most suitable economic predictors out of a wide range of potential driving factors is considered to be rather subjective. Moreover, the recent empirical evidence verifies that the commodity market is itself a segmented and heterogeneous market (see among others Daskalaki et al., 2014; Erb and Harvey, 2006). As a result, factors which are found to drive the prices of some traditional asset classes might not be capable of explaining the behaviour of individual commodities equally well.

To this end, the recent literature on predicting the dynamics of prices and returns in futures markets adopts a relatively newer approach; it involves predicting futures employing extracted common latent factors rather than predetermined common economic factors as the explanatory variables. That is because the set of predetermined economic factors may as well lead to omitted and/or irrelevant variable bias (Cummins et al., 2016). A statistical method commonly used in the literature to extract a few common unobserved (latent) factors out of a large dataset is the Principal Components Analysis (PCA hereafter), hence their name principal components (PCs). We provide a brief overview over the two strands of literature on the application of PCA on commodity futures. The first strand is associated with the application of the PCA on commodity price changes or returns, while the second performs the PCA on large economic datasets. The derived PCs are then used as predictors in a linear regression equation set up.

Concerning the first strand, Chantziara and Skiadopoulos (2008) employ joint PCA over the period 1993-2003 and extract three (3) PCs from the daily changes of futures prices. NYMEX WTI Crude Oil, Heating Oil, Gasoline, and International Petroleum Exchange (IPE) Brent Crude Oil futures of mutiple maturities comprise their large dataset. They provide evidence that the second PC posseses a significant predictive power over the heating oil and gasoline futures price changes. This predictive power, however, seems to apply for all maturities but the shortest. Finally, the estimated coefficients demonstrate a negative sign. Similarly, Cummins et al. (2016) also extract three PCs from WTI crude oil futures log returns over the period 2007-2016. In contrast to Chantziara and Skiadopoulos, they find significant evidence of oil futures predictability by the first extracted PC for any given maturity.

Regarding the second strand, Ludvigson and Ng (2009) and Gargano et al. (2016), among others, collect 131 monthly macroeconomic series from Global Insight database and extracted eight (8) common latent factors in order to predict the bond risk premia. Similarly, in an attempy to analyze oil prices, Le Pen and Sévi (2011) extract nine (9) common latent factors from a large dataset comprising of 187 real and nominal macroeconomic variables from both developed and emerging economies. The extracted common factors are then used in a typical predictive regression model.

Chapter 3: The Dataset

In this chapter, we discuss the energy futures prices as well as the selected macroeconomic, financial, and commodity-specific factors used as predictors in this study. Moreover, we provide a brief description of each data series characteristics.

3.1 Energy Futures Prices

3.1.1 The Dataset

The dataset is constituted by end-of-month settlement prices of energy futures traded on the New York Mercantile Exchange (NYMEX) which is part of the Chicago Mercantile Exchange Group (CME Group). In particular, we consider three closely related energy products as the underlying assets: *1*. NYMEX West Texas Intermediate (WTI) Light Sweet Crude Oil, *2*. NYMEX NY Harbor ULSD (Heating Oil) and *3*. NYMEX Henry Hub Natural Gas. Monthly[15] data on such futures contracts were obtained from Bloomberg database using the corresponding ticker symbols (i.e. *CL*, *HO* and *NG*, respectively).

However, within the context of this thesis, it is urgent that we use fixed maturity time series of futures prices. Bloomberg is not only capable of providing us with raw data on futures contracts for any maturity, but also gives us the possibility to roll over contracts yet keeping the same time to maturity. In this way, it creates generic series that actually represent contracts with almost a fixed time to maturity standing at any point in time (Chantziara and Skiadopoulos, 2008). Thus, we use generic series obtained from the first three nearest to maturity WTI Crude Oil futures contracts (labeled *CL1, CL2, CL3*), Heating Oil futures (labeled *HO1*, *HO2*, *HO3*) and Natural Gas futures (labeled *NG1*, *NG2*, *NG3*)[16]. Therefore, the chosen futures are those with the closest to delivery.

Furthermore, due to data availability constraints the time horizon of the dataset of all assets spans the period from January 1, 1990 to December 31, 2016, which accounts for 324 monthly observations. Consequently, the sample period captures both bearish and bullish regimes in commodity markets as well as potential break dates, i.e. the 1997 Asian financial crisis, the 1998 Russian Crisis, the early-2000s recession (Mar. 2001 - Nov. 2001) along with

[15] We use monthly observations to match with macroeconomic variables frequency, which are used as predictors.
[16] The *CL1*, *HO1* and *NG1* comprise the front month contracts or the shortest maturity (i.e. 1-month) series. They are the most liquid contracts traded and their prices are the closest ones to the spot prices. Similarly, *CL2*, *HO2* and *NG2* as well as *CL3*, *HO3* and *NG3* comprise the second and the third shortest maturity series, respectively.

the 2001 terrorist attack, the early-2000s US stock market rally, the 2003 Gulf War in Iraq, the 2003-2008 commodity boom, the 2007-2009 global financial crisis as well as the subsequent three rounds of US Quantitative Easing, i.e. QE1, QE2, and QE3 (see among others Radetzki, 2006; Mishra and Smyth, 2016).

Considering this 2000s commodities boom, we next proceed to splitting the initial dataset into smaller subsamples; the first subsample covers the period from January 1, 1990 to December 31, 2003, whereas the second subset covers the period from January 1, 2004 to December 31, 2016. This seems to be in line with Baker (2012), Chantziara and Skiadopoulos (2008), Hamilton and Wu (2013), who split their sample into two subsamples, i.e. the 1990-2003 pre financialization sample and the 2004-2011 post financialization sample. We note here that the first subsample (the so-called in-sample period) is used for parameter estimation. By contrast, the second subsample, which is called the out-of-sample period, is then used for producing forecasts across different forecasting horizons (up to one year) and the evaluation of their accuracy.

3.1.2 Data Characteristics

Having classified the futures contracts with respect to maturity, we show the dynamics of the three commodities futures prices, measured in US dollar and observed with a monthly frequency over the period 1990-2016.

Indicatively, Figure 3.1 plots prices of the three nearest maturity contracts for each commodity over time, giving a visual representation of the high correlations between these commodity prices. The graphs show simultaneous price movements of the first, second and third nearest contracts of the three energy commodities, therefore depicting a very similar pattern. The data appear to be non-stationary, with occasional spikes that indicate extreme volatility patterns. The actual US recessions are also shown as shaded grey areas. During the period of 2003-2008, and especially from 2006 until mid-2008 when the global financial crisis breaks out, prices increase sharply (commodities boom). The financial crisis seems to be followed by a rather sharp price reduction until prices recover again in 2009 and rocket in 2011.

Additionally, it is typical of prices of different crudes and refined products to correlate, a fact that could be easily detected by the same figure (Fig. 3.1). A common movement, and hence correlation, among the first three shortest maturity series of heating and crude oil is rather discernible. In particular, Table 3.1 shows correlation of these concurrent generics for the three commodities under examination. As perceived through the graphical representation of Fig.3.1, Table 3.1 spots nearly-perfect correlations among the crude and heating oil concurrent generics.

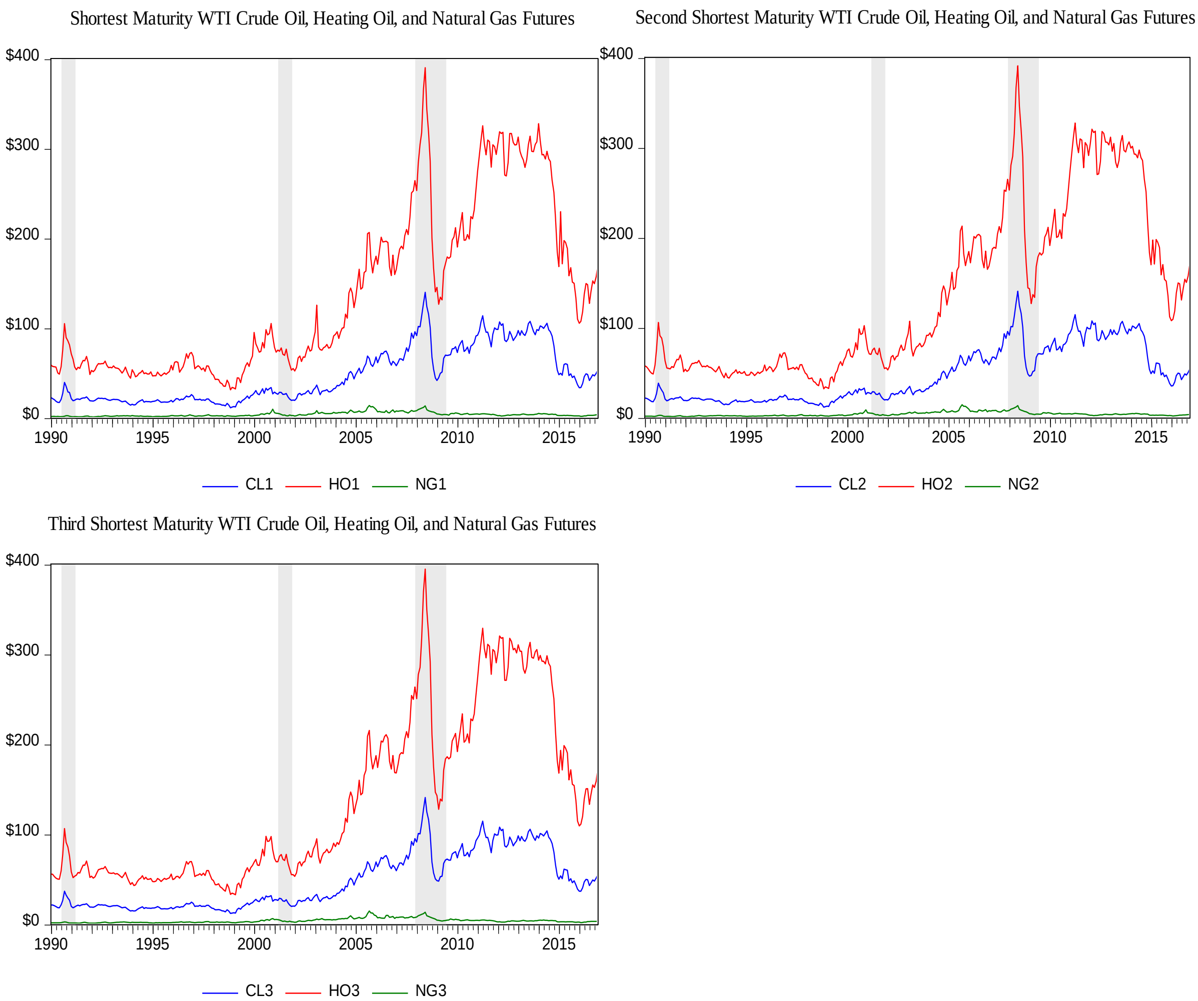

Fig. 3.1 *Price evolution of the three shortest maturity NYMEX WTI Crude Oil, Heating Oil, and Natural Gas futures over the period Jan.1990 to Dec.2016.*

Table 3.1
Correlation matrix of the first three shortest and concurrent futures contracts on the NYMEX WTI Crude Oil, Heating Oil, and Natural Gas.

	Panel A: Shortest Maturity		
	CL1	HO1	NG1
CL1	1.000		
HO1	0.989	1.000	
NG1	0.526	0.482	1.000
	Panel B: Second Shortest Maturity		
	CL2	HO2	NG2
CL2	1.000		
HO2	0.991	1.000	
NG2	0.537	0.495	1.000
	Panel C: Third Shortest Maturity		
	CL3	HO3	NG3
CL3	1.000		
HO3	0.992	1.000	
NG3	0.551	0.510	1.000

This table presents correlation among concurrent generics for the NYMEX WTI Crude Oil, Heating Oil, and Natural Gas futures sampled at monthly frequency. The results are reported for each maturity, i.e. shortest, second shortest, and third shortest maturity series in Panel A, Panel B, and Panel C, respectively. The (full) sample period is from January 1, 1990 to December 31, 2016.

3.2 Economic Predictors of Energy Futures

3.2.1 Economic Predictors

Concerning our economic model, we consider a set of six variables that have been documented to predict commodity, bond and equity returns (see Chapter 2) and use them as control variables. In short, the economic predictors employed in this study are the open interest median growth rate (*oi_gr*), the short-term real interest rate (*rir*), the term spread (*ts*), the default spread (*dfs*), the trade-weighted US dollar index (*twdi*), and the world steel production (*wsp*). With the exception of the world steel production, data on these potential driving factors are obtained from the Federal Reserve Economic Data (FRED) database[17] of the Federal Reserve Bank of St. Louis.

In particular, we first collect monthly not seasonally adjusted data on the Trade Weighted U.S. Dollar Index (with index being Jan 1997=100) in order to account for the US dollar exchange rate fluctuations (for further reasoning see Subsection *2.2.1.4 Exchange*

[17] Data are available at https://fred.stlouisfed.org .

Rates). Trade Weighted U.S. Dollar Index constitutes a weighted average of the foreign exchange value of the US dollar against major US trading partners. A positive relation should be expected between this factor and the futures prices or returns. To account for monetary conditions, we also calculate the term spread (see Subsection *2.2.1.3 Term Spread*) obtaining monthly data on the Long-Term Government Bond Yields (10-year) and the 3-Month Treasury Bill (Secondary Market Rate). Monthly default spread (see Subsection *2.2.1.2 Default Spread*) is also calculated by collecting data on the Moody's Seasoned BAA Corporate Bond Yields and the Moody's Seasoned AAA Corporate Bond Yields as well.

Moreover, following Shang (2010), we go on calculating the short-term real interest rate as the difference of the nominal interest rate (i.e. the 3-Month Treasury Bill rate) and the expected inflation. Expected inflation, constructed by the Survey Research Center of University of Michigan, constitutes the median expected change of prices during the following 12 months. A negative relation should be expected between the short-term real interest rate and the futures prices or returns (see Subsection *2.2.1.1 Short-Term Real Interest Rate*).

Second, in order to account for financialization in the NYMEX energy futures market, we collect data on each futures contract open interest (i.e. the amount of futures contracts outstanding) for the same sample period from the Commodity Futures Trading Commission. Following Gong and Yogo's (2012) methodology, we originally aimed at constructing a similar predictor, the aggregate dollar open interest as described in Subsection *2.2.1.5*. However, due to data availability constraints on the Henry Hub Natural Gas Spot Prices (spot prices have been available since Jan. 1997), we consider the growth rate of (simple instead of dollar) open interest for each energy commodity futures. After that, we proceed to calculating a median growth rate out of all the commodities examined at any given point in time. As discussed in Subsection *2.2.1.6*, a positive relation should be expected.

Finally, aiming at capturing the global real economic activity, we employ Ravazzolo and Vespignani's (2016) newly-developed proxy, i.e. the World steel production. Seasonally adjusted monthly data on the World steel production (measured in thousand metric tonnes, TMT) has been directly provided to us by Mr. Vespignani. A positive relation should also be expected between the World steel production and the futures prices or returns.

3.2.2 Data Characteristics

Figure 3.2 plots the level prices of the *oi_gr*, *rir*, *ts*, *dfs* and *twdi* as well as the logarithmic prices of the *wsp* (i.e. *lwsp*) over the sample period Jan.1990 to Dec.2016. From this graphical representation, one could deduce that nearly all the data series are non-stationary with the exception of the *oi_gr* series.

Indicatively, real interest rate exhibits a downward trend during the early 1990s and the early 2000s as well as a rather abrupt reduction in the wake of the global financial crisis of 2007, maintaining persistently low levels and even reaching negative values. According to FRED[18], these periods correspond to US recessions; during recession periods the nominal interest rates are decreased. The term spread seems to undergo a rather sharp increase during the same periods, obviously as a consequence of the low short-term interest rates. Default spreads are also sharply increased after the 2007 financial crisis, apparently reflecting the increased fear in the financial markets and reaching almost four basis points in Dec. 2008 (4bps ≈ 0.0338%). Moreover, the logarithmic series of World steel production maintains an overall sustained upward trend throughout the whole period with the exception of a rather small reduction during 1997-1998 (presumably as a result of the Asian and Russian crises) and a relatively sharper reduction during 2008.

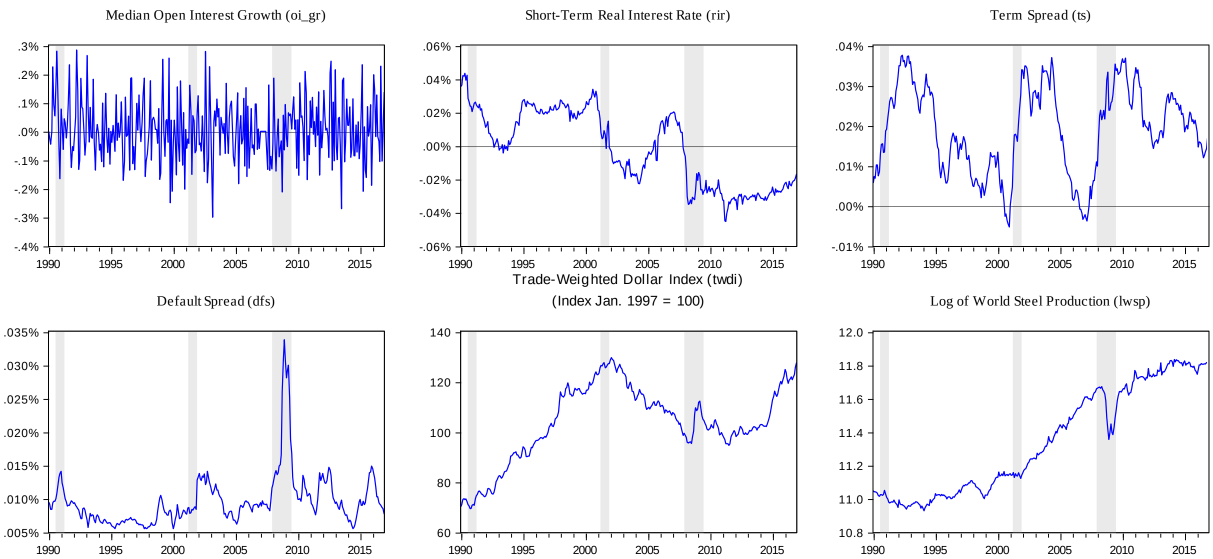

Fig. 3.2 *Evolution of the economic predictors series over the period Jan.1990 to Dec.2016.*

3.3 Descriptive Statistics and Stationarity Tests

It's urgent that the futures prices series as well as the economic predictors obtained undergo all the necessary statistical tests.

[18] See https://fred.stlouisfed.org/series/JHDUSRGDPBR#0 .

We thus continue by presenting the descriptive statistics for the end-of-month settlement prices of the NYMEX WTI Crude Oil, Heating Oil, and Natural Gas futures contracts and the economic variables as well. Table 3.2a summarizes the distribution of these series by reporting the mean, standard deviation, coefficient of skewness, and kurtosis of the series both in levels (Panel A and C) and first differences (Panel B and D).

We first check the data series for stationarity conducting an Augmented Dickey-Fuller (ADF) unit root test. In addition, in order to ensure the robustness of our results, we also apply the Dickey-Fuller test with GLS Detrending (DF-GLS) of Elliott et al., the Phillips-Perron (PP) tests, and the Zivot-Andrews Breakpoint Unit Root Test (see Table 3.2b). The aforementioned three tests check the null hypothesis (H_0) of a unit root existence which indicates non-stationarity. As for the Zivot-Andrews test, it tests the null hypothesis of a unit root existence while allowing for a structural break in both the intercept and trend.

Indicatively, for any series y_t we conduct the Augmented Dickey-Fuller test based on three different forms. The first form does not contain a constant or a trend term, i.e.

$$\Delta y_t = \delta\, y_{t-1} + \sum_{i=1}^{q} \gamma_i\, \Delta y_{t-i} + \varepsilon_t \qquad (3.1a)$$

the second contains a constant term, i.e.

$$\Delta y_t = a_0 + \delta\, y_{t-1} + \sum_{i=1}^{q} \gamma_i\, \Delta y_{t-i} + \varepsilon_t \qquad (3.1b)$$

and the third contains both constant and trend, i.e.

$$\Delta y_t = a_0 + \delta\, y_{t-1} + a_1 t + \sum_{i=1}^{q} \gamma_i\, \Delta y_{t-i} + \varepsilon_t \qquad (3.1b)$$

where q denotes the appropriate - optimal number of lag differences, defined by the Schwarz-Bayesian information criterion (SIC).

Results from the application of these tests indicate that not enough evidence is available to reject the null hypothesis at the significance level of 5%. It follows that most of the current futures prices are not stationary series in levels (see Panel A). As for the economic predictors, the only stationary series is that of the median open interest growth rate (see Panel C), while there is evidence that the rest of the predictors are integrated of order one, i.e. *I(1)*. To overcome the problem of non-stationarity, we proceed to creating new time series.

Regarding the futures contracts, we first transform the initial price series into logarithmic prices. For each commodity i (i=*CL*, *HO*, *NG*) expiring at T months (T=1,2,3), the new series is computed as follows:

$$lnF_t^{iT} = log(F_{t,T}^i) \tag{3.2}$$

where F_t^{iT} stands for the i-th commodity futures price observes at time t. We then compute their first differences, i.e. the first difference of the log prices at $t+h$ and t:

$$\begin{aligned} dlnF_{t+h}^{iT} &= lnF_{t+h}^{iT} - lnF_t^{iT} \\ &= \log\left(F_{t+h}^{iT}\right) - \log\left(F_t^{iT}\right) \\ &= \log\left(\frac{F_{t+h}^{iT}}{F_t^{iT}}\right) \Rightarrow \end{aligned}$$

$$d\left(\log\left(F_{t+h}^{iT}\right)\right) = \log\left(\frac{F_{t+h}^{iT}}{F_t^{iT}}\right) \tag{3.3a}$$

Note that in economic science small changes in the natural log of a variable, i.e. $d(\log(\bullet))$, approximate continuously compounded returns.

$$R_{t+h}^{iT} = \log\left(\frac{F_{t+h}^{iT}}{F_t^{iT}}\right) \tag{3.3b}$$

Indicatively, for a futures contract written on the NYMEX WTI Crude Oil (i=*CL*) maturing at one month (T=1) the $t+h$ return is computed as $R_{t+h}^{CL1} = \log\left(\frac{F_{t+h}^{CL1}}{F_t^{CL1}}\right)$. Actually, in the commodity futures markets the aforementioned returns reflect the (positive or negative) h-month (log) excess returns earned at time $t+h$ by securing a futures position; a futures market participant buys a T-month futures contract in month t and sells it h-months later, i.e. in month $t+h$ (just before it expires), thus experiencing a capital gain/loss of $R_{t+1:t+h}^{CL1}$ (see among others, Alquist et al., 2013).

To this end, Table 3.2a (Panel B) also summarizes the distribution of the futures returns. The results show that all energy futures earned positive mean returns during the whole period; *NG1* and *NG2* exhibit the highest monthly mean returns (0.008 or 0.80%) and, hence, standard deviation (see 16.4% and 14.3% for R_{NG1} and R_{NG2}, respectively). The results also indicate that the return distributions of natural gas futures contracts are negatively skewed (skewed to the left), whereas the distributions of the crude and heating oil futures returns are positively skewed and exhibit positive excess kurtosis (leptokurtic or fat tailed). Particularly, the Jarque-Bera (hereafter JB) Normality Test (with the null hypothesis being the H_0: the series is normally distributed) indicated that the all the return series (with the exception of R_{NG3}) are normally distributed. The application of the ADF test confirms the stationarity of the newly-created return series, which is in line with Zagaglia (2010) and Schalck and Chenavaz (2015). Figure 3.3 is a graphical representation of the return series.

Regarding the economic variables, we compute and examine the first difference of each non-stationary series. Results placed in Table 3.2a (Panel D) and Table 3.2b (Panel B) confirm the stationarity of the new series. Pairwise correlations among the economic predictors are also reported in Panel E of Table 3.2a. The results indicate that the economic predictors do not exhibit extreme correlation among each other, thus reducing the chances of multicollinearity problems. The evolution of the economic predictors sampled at a monthly frequency is shown in Figure 3.4.

Table 3.2a
Descriptive statistics of the three shortest maturity futures and economic predictors in levels and first differences

	Panel A: Futures Prices (in levels)								
	CL1	CL2	CL3	HO1	HO2	HO3	NG1	NG2	NG3
Observations	168	168	168	168	168	168	168	168	168
Mean	22.106	21.939	21.754	61.302	60.688	60.434	2.682	2.687	2.674
Std. Dev.	5.425	5.168	4.889	15.780	14.873	14.163	1.368	1.290	1.206
Skewness	0.640	0.612	0.574	0.902	0.790	0.728	1.982	1.656	1.302
Kurtosis	2.983	2.863	2.740	4.167	3.482	3.284	8.003	5.890	3.717
Jarque-Bera (JB)	11.480	10.629	9.692	32.308	19.114	15.417	285.158	135.204	51.069
	(0.003)***	(0.005)***	(0.008)***	(0.000)***	(0.000)***	(0.000)***	(0.000)***	(0.000)***	(0.000)***
ADF	-3.061	-2.859	-2.725	-3.340	-2.913	-2.766	-4.483	-3.814	-3.194
	(0.119)	(0.179)	(0.228)	(0.064)*	(0.161)	(0.212)	(0.002)***	(0.018)**	(0.089)*
	Panel B: Futures Returns (prices in logarithmic first differences)								
	R_{CL1}	R_{CL2}	R_{CL3}	R_{HO1}	R_{HO2}	R_{HO3}	R_{NG1}	R_{NG2}	R_{NG3}
Observations	167	167	167	167	167	167	167	167	167
Mean	0.002	0.002	0.002	0.003	0.003	0.003	0.008	0.008	0.007
Std. Dev.	0.095	0.088	0.081	0.105	0.095	0.089	0.164	0.143	0.119
Skewness	0.286	0.238	0.191	0.085	0.203	0.312	-0.309	-0.308	-0.059
Kurtosis	4.768	5.219	5.549	5.744	4.729	4.764	3.690	3.530	2.818
Jarque-Bera (JB)	24.019	35.837	46.241	52.586	21.951	24.360	5.976	4.597	0.328
	(0.000)***	(0.000)***	(0.000)***	(0.000)***	(0.000)***	(0.000)***	(0.050)**	(0.100)*	(0.849)
ADF	-11.503	-11.115	-10.951	-12.872	-11.617	-11.044	-11.214	-10.873	-10.694
	(0.000)***	(0.000)***	(0.000)***	(0.000)***	(0.000)***	(0.000)***	(0.000)***	(0.000)***	(0.000)***

This table reports the descriptive statistics for the NYMEX WTI Crude Oil, Heating Oil, and Natural Gas futures as well as the six economic predictors sampled at monthly frequency over the in-sample period. Panel A and C contain the results of the price levels, while Panel B (futures log-returns) and D reports the results on the stationary series eventually used in the economic model. Pairwise correlations among the economic predictors are also reported in Panel E. The in-sample period is from January 1990 to December 2003. "ADF" refers to the Augmented Dickey Fuller Test where the lag length selection is based upon the Schwarz-Bayesian Information Criterion (SIC). The asterisks *, **, and *** denote a rejection of the null hypothesis at 10%, 5% and 1% significance level, respectively. The null hypothesis for the Jarque-Bera and the ADF tests is that the series is normally distributed and that it contains a unit root, respectively.

Table 3.2a (Cont'd.)

Descriptive statistics of the three shortest maturity futures and economic predictors in levels and first differences

	Panel C: Economic Variables (in levels)					
	oi_gr	rir	ts	dfs	twdi	lwsp
Observations	168	168	168	168	168	168
Mean	0.008	0.014	0.019	0.008	101.115	11.066
Std. Dev.	0.110	0.014	0.011	0.002	19.210	0.095
Skewness	0.276	-0.580	-0.010	0.910	-0.147	0.813
Kurtosis	3.151	2.526	1.825	2.888	1.589	2.813
Jarque-Bera (JB)	2.297	10.994	9.662	23.267	14.537	18.758
	(0.317)	(0.004)***	(0.008)***	(0.000)***	(0.001)***	(0.000)***
ADF	-11.913	-1.431	-1.785	-2.547	-0.475	-2.071
	(0.000)***	(0.849)	(0.708)	(0.305)	(0.984)	(0.558)
DF-GLS	-11.688***	-1.467	-1.591	-2.347	-1.150	-0.843
PP	-13.881	-1.436	-1.836	-2.199	0.446	-1.873
	(0.000)***	(0.847)	(0.683)	(0.487)	(0.999)	(0.664)
	Panel D: Economic Variables (in first differences)					
	oi_gr	d(rir)	d(ts)	d(dfs)	d(twdi)	d(lwsp)
Observations	168	167	167	167	167	167
Mean	0.008	0.000	0.000	0.000	0.266	0.002
Std. Dev.	0.110	0.003	0.002	0.001	1.194	0.012
Skewness	0.276	-0.642	0.472	1.713	-0.031	0.298
Kurtosis	3.151	7.291	3.183	13.227	4.241	3.577
Jarque-Bera (JB)	2.297	139.615	6.425	809.382	10.751	4.781
	(0.317)	(0.000)***	(0.040)**	(0.000)***	(0.004)***	(0.092)*
ADF	-11.913	-6.965	-5.662	-10.178	-8.893	-5.863
	(0.000)***	(0.000)***	(0.000)***	(0.000)***	(0.000)***	(0.000)***
DF-GLS	-11.688***	-6.648***	-10.039***	-5.233***	-8.718***	-5.749***
PP	-13.881	-13.914	-9.161	-9.949	-8.637	-15.106
	(0.000)***	(0.000)***	(0.000)***	(0.000)***	(0.000)***	(0.000)***
	Panel E: Pairwise Correlations of Economic Predictors					
	oi_gr	d(rir)	d(ts)	d(dfs)	d(twdi)	d(lwsp)
oi_gr	1.000					
d(rir)	-0.086	1.000				
d(ts)	0.049	-0.298	1.000			
d(dfs)	0.008	-0.149	0.085	1.000		
d(twdi)	0.009	0.188	0.054	-0.020	1.000	
d(lwsp)	-0.053	-0.160	-0.050	0.048	0.016	1.000

This table reports the descriptive statistics for the NYMEX WTI Crude Oil, Heating Oil, and Natural Gas futures as well as the six economic predictors sampled at monthly frequency over the in-sample period. Panel A and C contain the results of the price levels, while Panel B (futures log-returns) and D reports the results on the stationary series eventually used in the economic model. Pairwise correlations among the economic predictors are also reported in Panel E. The in-sample period is from January 1990 to December 2003. "ADF" refers to the Augmented Dickey Fuller Test where the lag length selection is based upon the Schwarz-Bayesian Information Criterion (SIC). The asterisks *, **, and *** denote a rejection of the null hypothesis at 10%, 5% and 1% significance level, respectively. The null hypothesis for the Jarque-Bera and the ADF tests is that the series is normally distributed and that it contains a unit root, respectively.

Table 3.2b

Robustness tests for the stationarity of the economic predictors.

	Panel A: Economic Variables (in levels)					
	oi_gr	rir	ts	dfs	twdi	lwsp
Z-A	-12.300***	-4.005	-3.703	-4.215	-4.012	-3.376
Z-A Breakpoint	1993m09	2001m01	2001m03	1995m01	2001m09	1998m04
DF-GLS	-11.688***	-1.467	-1.591	-2.347	-1.150	-0.843
PP	-13.881	-1.436	-1.836	-2.199	0.446	-1.873
	(0.000)***	(0.847)	(0.683)	(0.487)	(0.999)	(0.664)
	Panel B: Economic Variables (in first differences)					
		d(rir)	d(ts)	d(dfs)	d(twdi)	d(lwsp)
Z-A		-6.630***	-6.630***	-10.487***	-9.116***	-6.226***
Z-A Breakpoint		2000m12	2001m01	2001m09	2000m07	1998m01
DF-GLS		-6.648***	-10.039***	-5.233***	-8.718***	-5.749***
PP		-13.914	-9.161	-9.949	-8.637	-15.106
		(0.000)***	(0.000)***	(0.000)***	(0.000)***	(0.000)***

This table reports robustness test regarding the stationarity of the economic predictors sampled at monthly frequency over the in-sample period. The in-sample period is from January 1990 to December 2003. Panel A contains the results on levels, while Panel B reports the results on the stationary series eventually used in the economic model. "Z-A", "DF-GLS" and "PP" refer to the Zivot-Andrews Breakpoint Unit Root Test, the Dickey-Fuller GLS Test, and the Phillips-Perron Test, respectively. The asterisks *, **, and *** denote a rejection of the null hypothesis at 10%, 5% and 1% significance level, respectively. The null hypothesis for the DF-GLS and PP tests is that the series contains a unit root, while the null hypothesis for the Z-A test examines for a unit root existence while allowing for a structural break in both the intercept and trend

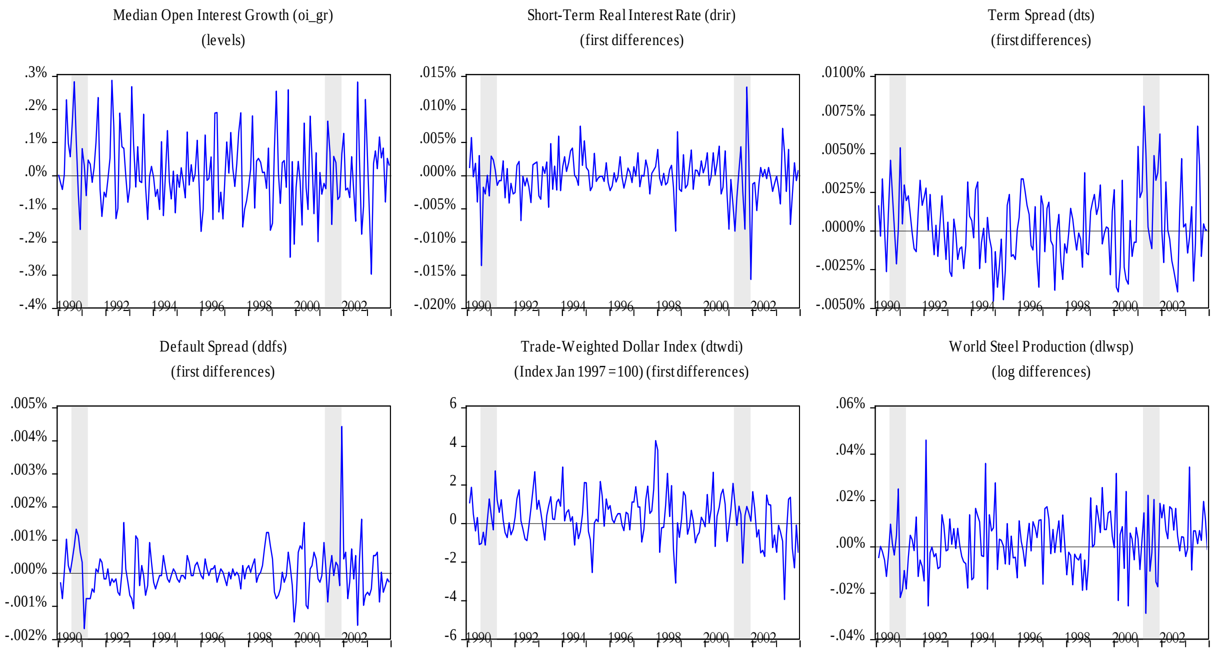

Fig. 3.3 *Stationary series of the six economic predictor from Jan.1990 to Dec.2003.*

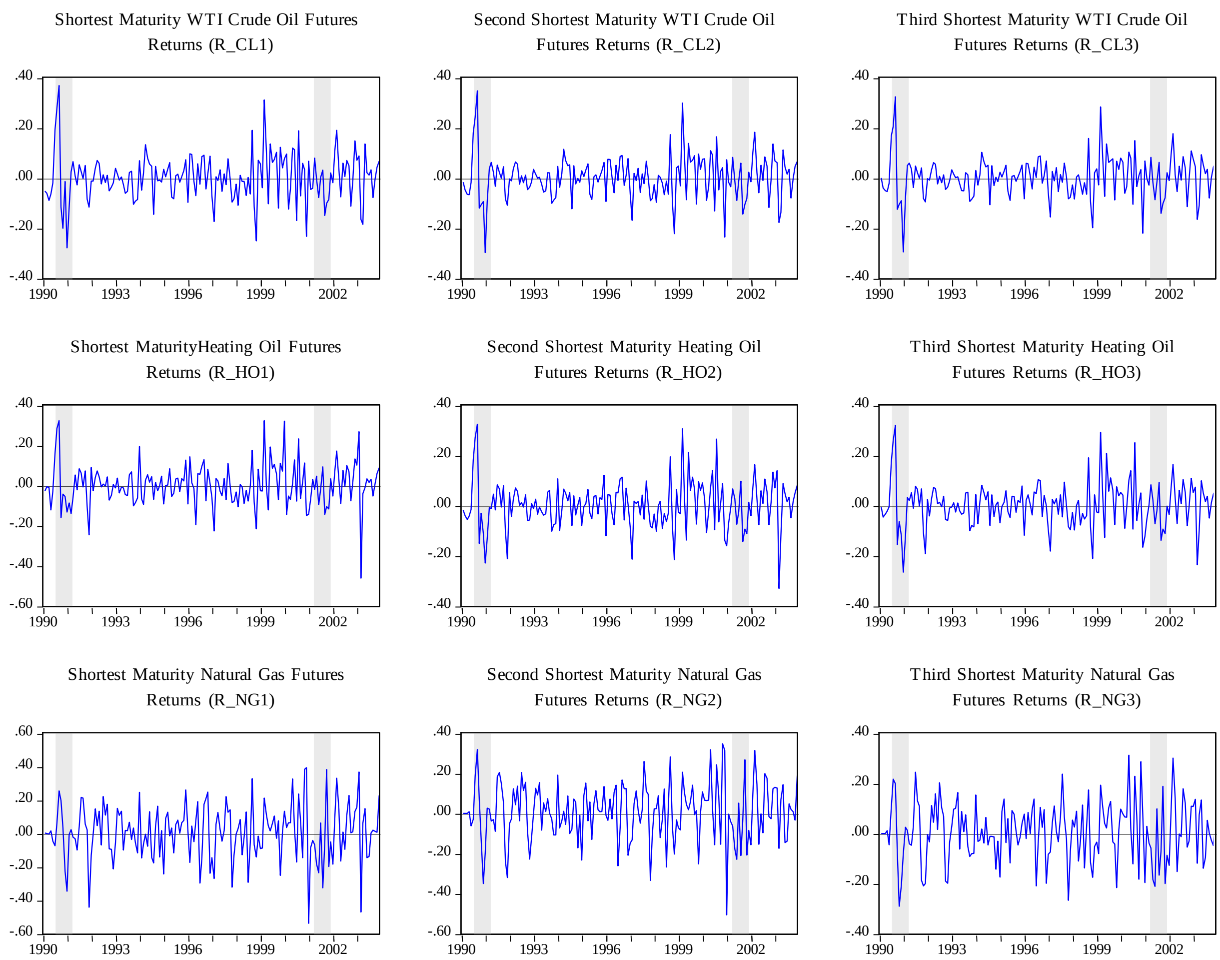

Fig. 3.4 *Returns of the three shortest maturity NYMEX WTI Crude Oil, Heating Oil and Natural Gas futures from Feb.1990 to Dec.2003*.

3.4 Principal Component Analysis (PCA) Dataset

3.4.1 McCracken and Ng's Dataset

Despite the fact that our economic model incorporates variables that have some clear economic interpretation, selecting the most suitable predictors out of a wide range of potential driving factors is considered to be a rather subjective process. To this end, the Principal Component Analysis (PCA) employed within the context of this study involves the extraction[19] of few latent factors, i.e. the principal components (PCs), which can be used as predictors in a linear regression equation.

[19] The methodology regarding the extraction of the common latent factors is described in Subsection *4.1.3.1 Description*.

However, the PCA requires the dissemination of information from large macroeconomic datasets, viz. "big data", the collection of which is rather difficult and lengthy. Consisting of 134 monthly macroeconomic US indicators, the McCracken and Ng's (2016) dataset facilitates the data collection process. We thus obtain the McCracken and Ng's macroeconomic dataset which is publicly and readily accessible from FRED[20] and, based on the suggested data transformation that they provide, we proceed to transforming the variables in order to achieve stationarity. We note here that stationarity of the variables included in the large dataset is required in order to obtain valid estimates for the PCs.

3.4.2 Principal Components (PCs) Characteristics

Next, we employ the PCA and extract 134 PCs with the decision upon the number of PCs to retain being left to the discretion of each individual researcher. A common rule of thumb indicates keeping the number of PCs that produce a cumulative variance of 90%. As Table 3.3 shows (Panel B), there are three (3) out of the total number (134) of joint PCs which perform a 90.41% cumulative percentage of the original variance explained. We thus retain the first three standardized PCs in order to employ them in our latent factor model.

Table 3.3 (Panel A) also presents the descriptive statistics of the first three joint standardized PCs derived. The results from the application of the Jarque-Bera test indicate that the first PC is not normally distributed, while the second and the third retained PCs are found to be normally distributed. Moreover, application of the ADF, DF-GLS and PP tests on each of the three retained PCs indicates that the extracted principal components are stationary.

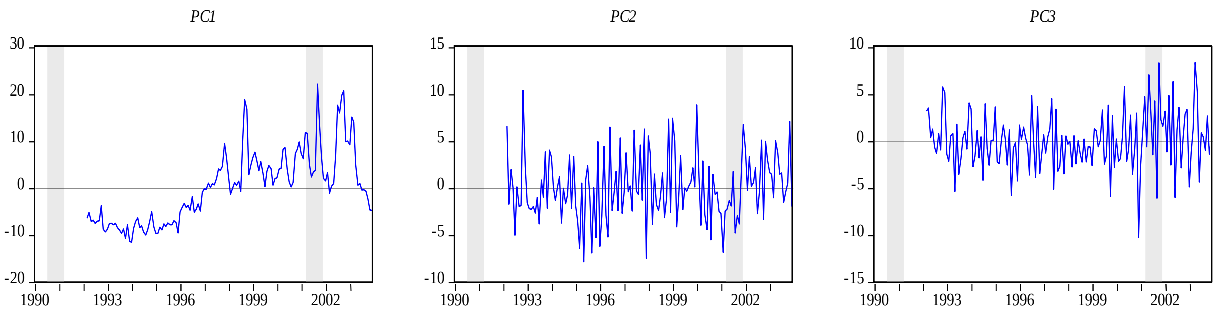

Fig. 3.5 *Graphical representations of the first three extracted PCs from Jan.1990 to Dec.2003*.

[20] The large macroeconomic database is available at https://research.stlouisfed.org/econ/mccracken/fred-databases/ (FRED-MD).

Table 3.3
Descriptive statistics of the three standardized PCs retained from the Joint PCA and cumulative proportion of variance explained.

	Panel A: Descriptive Statistics		
	PC_1	PC_2	PC_3
Missing Observations	26	26	26
Skewness	0.660	0.355	0.079
Kurtosis	2.930	2.940	3.465
Jarque-Bera (JB)	10.346	2.999	1.428
	(0.006)***	(0.223)	(0.489)
	Panel B: Explained Variance		
	PC_1	PC_2	PC_3
Eigenvalue	58.708	12.391	9.382
Cumulative Value	58.708	71.100	80.482
Cumulative Proportion	65.950	79.870	90.410

This table reports the descriptive statistics (Panel A) for the first three standardized principal components (PCs) extracted from the application of Joint PCA on the McCracken and Ng's (2016) large macroeconomic dataset. The 134 macroeconomic indicators incorporated in the dataset are sampled at monthly frequency over the in-sample period, i.e. Jan. 1990 - Dec. 2003. Based on the McCracken and Ng's suggested data transformation, the aforementioned series are transformed into stationary series necessary for obtaining valid estimates for the PCs. The asterisks *, **, and *** denote a rejection of the null hypothesis at 10%, 5% and 1% significance level, respectively. The null hypothesis for the Jarque-Bera test is that the series is normally distributed. Panel B reports the cumulative proportion of the original dataset's variance explained by the retained PCs.

However, it is rather obvious that the PCs extracted from the large dataset do not have any economic interpretation, thus making it rather difficult to us to comprehend the information conveyed. To this end, we are motivated by Zagaglia's (2010) work, who regresses each of his extracted factors on the data series comprising his large dataset and, eventually, identifies which of the series is most correlated with each factor based on the corresponding R^2 values. Similarly, in order to examine the relationship between each individual PC and each series (after transformation) included in the McCracken and Ng's dataset, we proceed to creating three correlation matrices (one for each of the components). We next compute the R^2 values as the square of the pairwise correlation coefficients (this holds for bivariate regressions) in each matrix. Ranking the results in a descending order, Table 3.4 reports the variance explained by the five (McCracken and Ng's transformed) series which are most correlated with the PCs extracted.

The results indicate nearly-perfect correlation between the first extracted principal component (PC_1) and the CBOE S&P 100 Volatility Index (VXO), apparently reflecting the investors' sentiment and expectations regarding the short-term (one-month) volatility in the

US stock market, conveyed by S&P 100 stock index option prices. Increased uncertainty and "fear" in the stock market is typically related to bearish price action. Finally, PC_2 and PC_3 seem to be moderately correlated with two individual indicators, the Consumer Sentiment Index and the Help-Wanted Index. An increased consumer sentiment index is associated with positive feelings and prospects regarding the short-term future economic and financial conditions. This, in turn, leads to a positive expected relationship between the index of Consumer Sentiment and the futures returns. As for the Help-Wanted Index, it captures indicates how many positions need to be filled. An increased index is thought to reflect a shortage of employees, which in turn may lead to higher wages and wage inflation (in order to attract workers) that adversely affect the conventional markets. On the contrary, the increased salaries improve the employees' position as consumers, which might lead to higher commodity prices.

Table 3.4
Proportion of explained variance of highly-correlated series.

R^2 between the McCracken & Ng's Data Series and the Extracted PCs				
			Panel A: PC_1	
FRED mnemonic	Description	Group	ρ	$\rho^2 \approx R^2$
VXOCLSx	CBOE S&P 100 Volatility Index (VXO)	Stock market	0.997	99.391%
L_PERMITW	New Private Housing Permits, West	Housing	0.662	43.785%
L_PERMIT	New Private Housing Permits	Housing	0.658	43.251%
L_PERMITS	New Private Housing Permits, South	Housing	0.656	43.089%
L_HOUST	Housing Starts: Total New Privately Owned		0.580	33.685%
			Panel B: PC_2	
FRED mnemonic	Description	Group	ρ	$\rho^2 \approx R^2$
D_UMCSENTx	Consumer Sentiment Index	Consumption	0.724	52.410%
D_HWI	Help-Wanted Index for US	Labour Market	0.723	52.266%
D_HWIURATIO	Ratio of Help Wanted / No.Unemployed	Labour Market	0.585	34.240%
D_GS1	1-Year Treasury Rate	Interest & Exchange Rates	0.329	10.847%
D_GS5	5-Year Treasury Rate	Interest & Exchange Rates	0.303	9.198%
			Panel C: PC_3	
FRED mnemonic	Description	Group	ρ	$\rho^2 \approx R^2$
D_HWI	Help-Wanted Index for US	Labour Market	-0.656	43.038%
D_UMCSENTx	Consumer Sentiment Index	Consumption, Orders & Inventories	0.606	36.784%
D_HWIURATIO	Ratio of Help Wanted / No.Unemployed	Labour Market	-0.572	32.730%
D_LD_MZMSL	MZM Money Stock	Money & Credit	-0.349	12.163%
D_LD_M2SL	M2: Money Stock	Money & Credit	-0.311	9.690%

This table reports the strongest pairwise correlations between the McCracken and Ng's (2016) data series and the first three principal components (PCs) extracted from their whole dataset. The corresponding FRED mnemonic, description, and group are also reported for each series. The sample covers the period Jan. 1990 - Dec. 2003.

3.4.3 Stability of the PCA results

Furthermore, we check the stability of the PCA results over the examined period, i.e. Jan. 1990 to Dec. 2003. That is because the US economy underwent an eight-month recession period (Mar. 2001 - Nov. 2001), while the McCracken and Ng's (2016) database consists exclusively of US economic indicators.

Therefore, we break up the in-sample period in two sub-periods, i.e. Jan. 1990 - Feb. 2001 and Mar. 2001 - Dec. 2003, the cutoff point reflecting the month of the early-2000s recession break-out, which might have created a structural break in the data. We then apply the PCA on the 134 McCracken and Ng's transformed data series within each sub-period. Table 3.5 reports the results. As Table 3.3 shows (Panel B), there are three (3) out of the total number (134) of joint PCs which perform a 93.22% and 96.09% cumulative proportion of the original variance explained within the first and the second sub-period, respectively. Overall, the results obtained from PCA analysis on the full in-sample period (Jan. 1990 - Dec. 2003) do not demonstrate significant differences from those obtained from the two sub-periods.

Table 3.5
Descriptive statistics of the three standardized PCs retained from the Joint PCA and cumulative proportion of variance explained within the two sub-periods.

	1st Sub-Period: Jan.1990-Feb.2001			*2nd Sub-Period: Mar.1991-Dec.2003*		
	Panel A: Descriptive Statistics					
	PC_1	PC_2	PC_3	PC_1	PC_2	PC_3
Observations	108	108	108	34	34	34
Skewness	0.581	0.442	-0.191	0.567	-0.090	0.440
Kurtosis	2.628	2.976	3.618	2.278	2.442	2.317
Jarque-Bera (JB)	6.708	3.524	2.377	2.564	0.488	1.760
	(0.035)**	(0.172)	(0.305)	(0.277)	(0.784)	(0.415)
	Panel B: Explained Variance					
	PC_1	PC_2	PC_3	PC_1	PC_2	PC_3
Eigenvalue	44.711	13.263	7.304	56.759	15.153	8.039
Cumulative Value	44.711	57.974	65.278	56.759	71.912	79.951
Cumulative Proportion	63.850	82.790	93.220	68.210	86.430	96.090

This table reports the descriptive statistics (Panel A) for the first three standardized principal components (PCs) extracted from the application of Joint PCA on the McCracken and Ng's (2016) large macroeconomic dataset (transformed series) over two sub-periods. The cutoff point corresponds to the month of the early-2000s recession break-out, i.e. March 2001, which might have created a structural break in the data. The asterisks *, **, and *** denote a rejection of the null hypothesis at 10%, 5% and 1% significance level, respectively. The null hypothesis for the Jarque-Bera test is that the series is normally distributed. Panel B reports the cumulative proportion of the original dataset's variance explained by the retained PCs for each sub-period.

Chapter 4: Methodology

In this section, we first provide a brief description of the alternative return prediction models employed within the context of this study in order to individually examine the predictability of the NYMEX energy futures prices. We next describe the methodological approaches followed to estimate the models and produce short-term in-sample predictions (point forecasts) as well as out-of-sample forecasts. Last but not least, the evaluation criteria used to assess the predictive accuracy of the produced forecasts are also reported in this chapter.

4.1 In-Sample Return Prediction Models

Overall, for each futures contract under examination we employ three alternative model specifications; we construct a seven-factor economic model, a univariate autoregressive model *AR(1)*, and a Principal Component Analysis (PCA hereafter) latent factor model.

4.1.1 Economic Variables Model

Concerning our economic model, select a rather small set of variables out of a wide range of factors that have been previously found to possess a predictive power over commodity, bond and/or equity prices and risk premia.

In short, the economic factors employed in this study are the open interest median growth rate (*oi_gr*), the short-term real interest rate (*rir*), the term spread (*ts*), the default spread (*dfs*), the trade-weighted US dollar index (*twdi*) and the world steel production (*wsp*) (see Chapter 2). Since the six aforementioned variables are though to convey information either on the current or the future economic and financial conditions of an economy, we consider them as control variables in our economic model. The cyclical variations of these factors, either they are procyclical or countercyclical, produce cyclical (or time-varying) variations in the assets returns, too.

Given that the futures dataset consists of three alternative energy commodities, each of multiple maturities, we proceed to constructing nine (9) respective forecasting models. In particular, for each commodity i (i=*CL*, *HO*, *NG*) expiring at T periods (T=1, 2, 3 months) we consider the following return prediction model that is commonly employed in the financial literature of risk premium predictability (see among others Baetje and Menkhoff, 2016; Gargano and Timmermann, 2014; Kolev and Karapandza, 2017):

$$R^{iT}_{t+1:t+h} = a_0 + \sum_{k=1}^{K} a_k\, x_{k,t} + \varepsilon^{iT}_{t+1:t+h} \tag{4.1a}$$

where a_0 is a constant, $x_{k,t}$ ($1 \ll k \ll K$) denotes the k-th predictor[21] with its corresponding parameter a_k to be estimated, and $\varepsilon^{iT}_{t+1:t+h}$ denotes the respective random disturbance term usually assumed to be serially uncorrelated and distributed with mean zero and constant variance

$$\varepsilon^{iT}_{t+1:t+h} \sim N(0, {\sigma_\varepsilon}^2) \tag{4.1b}$$

Moreover, $R^{iT}_{t+1:t+h}$ denotes the cumulative[22] return of holding the asset over the interval end-of-month t to end-of-month $t+h$. This is supposed to be the h-step-ahead forecast, where the forecast horizon h is left to the discretion of each author ($1 \ll h < \infty$ and $h \epsilon \mathbb{Z}^+$). Alternatively, the end-of-month t could be expressed as the beginning-of-month $t+1$. The asset corresponds to the i-th futures contract of a given maturity T (T=1,2,3). In addition, t=1,2, …, N=167 (in-sample size). In particular, for the in-sample period (Feb.1990-Dec.2003) considered within the context of this study, the original sample size equals N=167 monthly observations (returns). To this end, $N-h$ observations are available to be used in order to estimate the in-sample predictive model. Alternatively, Equation (4.1a) could be written as:

$$R^{iT}_{t+1-h:t} = a_0 + \sum_{k=1}^{K} a_k\, x_{k,t-h} + \varepsilon^{iT}_{t+1-h:t} \tag{4.1c}$$

Since the set of predictors can incorporate both exogenous economic factors and past values of the dependent variable, we also put the dependent variable lagged h-months:

$$R^{iT}_{t+1-h:t} = a_0 + \sum_{k=1}^{K} a_k\, x_{k,t-h} + a_{K+1}\, R^{iT}_{t-h-1:t-h} + \varepsilon^{iT}_{t+1-h:t} \tag{4.1d}$$

For a forecast horizon of h=1 month ahead Equation (4.1d) can be written as follows:

$$R^{iT}_{t+1-1:t} = a_0 + \sum_{k=1}^{K} a_k\, x_{k,t-1} + a_{K+1}\, R^{iT}_{t-1-1:t-1} + \varepsilon^{iT}_{t+1-1:t} \;\;\Leftrightarrow$$

[21] The set of potential predictors can incorporate exogenous explanatory factors along with lagged values of the dependent variable.

[22] For the cumulative return earned from t to $t+h$ it holds that: $R^{iT}_{t+1:t+h} = \sum_{i=1}^{h} R^{iT}_{t+i}$. For instance, $h=3$ produces a cumulative return of $R^{iT}_{t+1:t+3} = \sum_{i=1}^{h=3} R^{iT}_{t+i} = R^{iT}_{t+1} + R^{iT}_{t+2} + R^{iT}_{t+3}$.

$$R_{t:t}^{iT} = a_0 + \sum_{k=1}^{K} a_k\, x_{k,t-1} + a_{K+1}\, R_{t-2:t-1}^{iT} + \varepsilon_{t:t}^{iT}$$

or

$$R_t^{iT} = a_0 + \sum_{k=1}^{K} a_k\, x_{k,t-1} + a_{K+1}\, R_{t-1}^{iT} + \varepsilon_t^{iT} \qquad (4.1e)$$

To this end, we generate one-step ahead ($h = 1$) in-sample forecasts regressing each R_t^{iT} series on the six (K=6) one-month lagged economic predictors and its own past values (R_{t-1}^{iT}). Note here that R_t^{iT} reflects the cumulative return earned from buying a futures at the end of month $t-1$ and selling it at the end of month t. The nine (9) models are estimated using the Ordinary Least Squares (OLS) method with the *t*-statistics being adjusted for both heteroscedasticity and serial correlation using the Newey-West method (HAC Consistent Covariance for the calculation of the standard errors). The estimation is done using the 166 available observations (N-h, where h=1).

$$R_t^{iT} = \alpha_0 + a_1\, oi\_gr_{t-1} + a_2\, d(rir)_{t-1} + a_3\, d(ts)_{t-1} + a_4\, d(dfs)_{t-1} + a_5\, d(twdi)_{t-1} + \\ + a_6\, d(lwsp)_{t-1} + a_7\, R_{t-1}^{iT} + \varepsilon_t \qquad (4.1e)$$

According to Bessembinder and Chan (1992), such a model specification is assumed to test the semi-strong form efficiency of a market, due to the fact that it considers nearly all the information available up to time t.

4.1.2 Univariate Autoregressive Models AR(1)

Apart from our basic economic model, univariate autoregressive models are also employed in order to investigate the predictability of the NYMEX energy futures from its own past values. This model specification set up is thought to test the weak form efficiency of a market; the weak form market efficiency assumes that assets' market prices incorporate all available information at any point in time.

To this end, we consider the following return prediction model for each commodity i (i=*CL*, *HO*, *NG*) maturing at T months (T=1,2,3):

$$R_{t+1:t+h}^{iT} = a_0 + \theta\, R_{t-1:t}^{iT} + \varepsilon_{t+1:t+h}^{iT} \qquad (4.2a)$$

or

$$R_{t+1-h:t}^{iT} = a_0 + \theta\, R_{t-h-1:t-h}^{iT} + \varepsilon_{t+1-h:t}^{iT} \qquad (4.2b)$$

For a forecast horizon of h=1 month Equation (4.2b) can be written as follows. Therefore, the autoregressive models to be estimated are:

$$R_{t+1-1:t}^{iT} = a_0 + \theta\, R_{t-1-1:t-1}^{iT} + \varepsilon_{t+1-1:t}^{iT}$$
$$R_{t:t}^{iT} = a_0 + \theta\, R_{t-2:t-1}^{iT} + \varepsilon_{t:t}^{iT}$$

or

$$R_t^{iT} = a_0 + \theta\, R_{t-1}^{iT} + \varepsilon_t^{iT} \tag{4.2c}$$

4.1.3 Principal Components Analysis (PCA) Models

In this subsection we provide a brief description of the Principal Components Analysis (PCA hereafter), which constitutes an alternative method of investigating the dynamics of variables under consideration without assuming a predetermined set of control variables. Overall, the PCA involves the extraction of few unobserved (latent) factors out of a large dataset, the so called principal components (PCs hereafter). The extracted factors can then be employed as potential predictors in a linear regression equation set up (Chantziara and Skiadopoulos, 2008).

4.1.3.1 Description

In brief, PCA is a linear statistical technique that contributes to reducing the dimensionality of a large number of potentially correlated stationary variables, transforming them into a set of uncorrelated variables, the common latent factors or PCs, while retaining as much of the original variation as possible (Daskalaki et al., 2014). The PCs are arranged in such a way that the first few retain most of the variation present in the original dataset.

In particular, we start by denoting time by $t = 1, 2, \dots, T$ and assuming a set of n variables under consideration $\boldsymbol{r_1}, \boldsymbol{r_2}, \dots, \boldsymbol{r_n}$, so that each and every original variable j corresponds to a (Tx1) vector $\boldsymbol{r_j}$ and $\boldsymbol{R}$ is a (Txn) matrix of the n original variables:

$$\boldsymbol{R} = [\boldsymbol{r_1}, \boldsymbol{r_2}, \dots, \boldsymbol{r_n}] = \begin{pmatrix} r_{11} & r_{21} & \dots & r_{n1} \\ r_{12} & r_{22} & \dots & r_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ r_{1T} & r_{2T} & \dots & r_{nT} \end{pmatrix} \tag{4.3a}$$

We also assume a variance-covariance matrix, $\boldsymbol{S_r}$.

According to Chantziara and Skiadopoulos (2008) (see also Cavalcante, 2010), employing the PCA, we obtain a new set of n artificial variables (i.e. the PCs) as linear combinations of the vectors $\boldsymbol{r}$, which are orthogonal to each other and capable of replicating the variance-covariance structure ($\boldsymbol{S_r}$) of the n original variables. Therefore, in matrix notation the PCs are calculated as follows:

$$\boldsymbol{P} = \boldsymbol{RL} \tag{4.3b}$$

where $\boldsymbol{P}$ is a (Txn) matrix of PCs due to the fact that there are T observations on each explanatory variable, r_j. Specifically,

$$\boldsymbol{P} = [\boldsymbol{PC_1}, \boldsymbol{PC_2}, \dots, \boldsymbol{PC_n}] = \begin{pmatrix} PC_{11} & PC_{21} & \dots & PC_{n1} \\ PC_{11} & PC_{22} & \dots & PC_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ PC_{1T} & PC_{2T} & \dots & PC_{nT} \end{pmatrix} \tag{4.3c}$$

with the columns $\boldsymbol{PC_j}$ of matrix $\boldsymbol{P}$ being linear combinations of the columns of $\boldsymbol{R}$. For instance, the j^{th} principal component is given as

$$\boldsymbol{PC_j} = a_{1j}\boldsymbol{r_1} + a_{2j}\boldsymbol{r_2} + \dots + a_{nj}\boldsymbol{r_n} \tag{4.3d}$$

where a_{ij} are the coefficients to be estimated (also known as factor loadings), corresponding to the coefficient on the i-th variable in the j-th principal component. Overall, $\boldsymbol{L}$ is a (nxn) matrix of the coefficients used to obtain the PCs (i.e. loadings):

$$\boldsymbol{L} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix} \tag{4.3d}$$

Moreover, the extracted PCs are capable of explaining 100% of the original dataset's variance and are arranged in order of diminishing variance. In this way, the first PC (denoted by PC_1) constitutes a linear combination of variables and accounts for the greatest percentage of variance of the original n variables. The second PC, i.e. the PC_2, explains as much of the remaining variance (the variance left unexplained by PC_1) as possible. The third PC, PC_3, explains the amount of variance not explained by the PC_1 and PC_2, and so on. Subsequently, the sum of the variances of the n PCs equals the total variance of the original n variables.

Furthermore, in cases where a small number of PCs is found to explain a sufficiently large percentage of the total variance of the original variables, we could, for the sake of parsimony, discard (omit) the rest of the PCs that account for only a small amount of the total variation, thus reducing significantly the dimension of the problem. Indicatively, retaining only the first $0 < m < n$ PCs and discarding the remaining $n - m$ PCs, we get

$$\boldsymbol{R} = P_{(m)}L'_{(m)} + \varepsilon_{(m)} \tag{4.3e}$$

where $\boldsymbol{P}$ is a $(T\text{x}m)$ matrix of PCs, $\boldsymbol{R}$ is a $(T\text{x}m)$ matrix of the original variables, $\boldsymbol{L}$ is a $(m\text{x}m)$ matrix of the loadings, and $\varepsilon_{(m)}$ is a $(T\text{x}n)$ matrix of residuals (Chantziara and Skiadopoulos, 2008). Nevertheless, the decision upon the number of PCs to retain is left to the discretion of each individual researcher, since there is not a clearcut criterion. A common rule of thumb, however, indicates keeping the number of PCs that account for the 90% of the total variance (i.e. cumulative explained variance).

4.1.3.2 PCs as Predictors – Joint PCA Models

Having applied the PCA on the McCracken and Ng's (2016) large macroeconomic dataset, the three retained PCs (i.e. $m = 3$ out of the $n = 134$ PCs) can then be used as predictors in a linear regression equation (see Chantziara and Skiadopoulos, 2008).

To this end, we consider the following return prediction model for each commodity i (i=*CL*, *HO*, *NG*) maturing at T months (T=1,2,3):

$$R_{t+1:t+h}^{iT} = a_0 + \sum_{j=1}^{m=3} a_j \, PC_{j,t} + \varepsilon_{t+1:t+h}^{iT} \tag{4.4a}$$

or

$$R_{t+1-h:t}^{iT} = a_0 + \sum_{j=1}^{m=3} a_j \, PC_{j,t-h} + \varepsilon_{t+1-h:t}^{iT} \tag{4.4b}$$

For a forecast horizon of h=1, Equation (4.4b) can be written as:

$$R_t^{iT} = a_0 + \sum_{j=1}^{m=3} a_j \, PC_{j,t-1} + \varepsilon_t^{iT} \tag{4.4c}$$

Consequently, the nine (9) PCA models to be estimated using the Ordinary Least Squares (OLS) method are:

$$R_t^{iT} = a_0 + a_1\, PC_{1,t-1} + a_2\, PC_{2,t-1} + a_3\, PC_{3,t-1} + \varepsilon_t \qquad (4.4d)$$

Not that the t-statistics are again adjusted for both heteroscedasticity and serial correlation using the Newey-West method (HAC Consistent Covariance for the calculation of the standard errors). Equation (4.4d) is estimated using the 166 available observations (N-h, where h=1).

4.1.4 In-Sample Evaluation Criteria

After having estimated the aforementioned model specifications, we proceed to evaluating their in-sample fit employing some commonly used performance measures.

Concerning the in-sample predictive ability of each factor, we examine the t-statistic and the p-value of the corresponding estimated coefficient. We test the null hypothesis of a statistically insignificant parameter; in other words, this indicates non predictability of the futures returns by this factor. In particular, under the model specification of Eq. (4.1a), we assess the predictive ability of each k-th predictor (x_k) by testing the statistical significance of the corresponding a_k parameter as follows:

$$\begin{aligned} &H_0: a_k = 0 \\ &H_1: a_k \neq 0 \end{aligned} \qquad (4.5a)$$

The null hypothesis (H_0) of the test states that there is no predictability. On the contrary, a rejection of the H_0 in favor of the alternative hypothesis (H_1) indicates that information of the k-th predictor available at the end-of-month t is useful for predicting the cumulative return earned in the end of period $t + h$.

Additionally, the R-squared coefficient (R^2) and the adjusted R-squared (adj. R^2) are typically used as criteria to assess the overall goodness of each model's fit. Regarding the R-coefficients (also known as coefficients of determination), they lie between zero (0) and unity (1), i.e. in the interval $[0.0, 1.0]$ or $[0.0, 100.0\%]$, and are computed as follows:

$$R^2 = \frac{\sum_{t=1}^{N-h}(\widehat{y_t}-\overline{y_t})^2}{\sum_{t=1}^{N-h}(y_t-\overline{y_t})^2} \tag{4.5b}$$

$${R_{adj}}^2 = R^2 - \frac{c-1}{N-c-1}\,(1-R^2) \tag{4.5c}$$

where N denotes the in-sample size and c denotes the number of parameters in the model. Note that c does not include the constant (a_0). The R-coefficients reflect the proportion of the total sample variance that is explained (determined) by the explanatory variables included in the selected model specification. Values that are close to unity (or 100%) indicate a rather good specification, while values close to zero do not indicate any explanatory power.

Finally, the *F*-Statistic *p*-value (*F*-stat prob.) is also used; the overall goodness of fit is evaluated in a statistically significance sense.

$$\begin{aligned} &H_0: a_1 = a_2 = \dots a_k \dots = a_K = 0 \\ &H_1: at\ least\ one\ a_k \neq 0 \end{aligned} \tag{4.5d}$$

H_0 of the test states that the current model does not perform well in-sample, while evidence in favor of the alternative hypothesis (H_1) indicate good fit and explanatory power of the regressors (explanatory variables) on the regressand variable.

$$F = \frac{R^2}{1-R^2}\,\frac{N-K-1}{K} \tag{4.5e}$$

where K, N denote the number of the regressors and the number of the in-sample observations, respectively.

4.2 Out-of-Sample (OoS) Forecasting Models

4.2.1 Forecasting Models

We next examine the Out-of-Sample (OoS hereafter) performance of the three predictive models described in Subsections *4.1.1*, *4.1.2,* and *4.1.3.2*. The OoS analysis is commonly used because of an empirical finding; in-sample inference of predictability does not always entail out-of-sample predictability. Predictors found to possess a consistent predictive power over the in-sample sub-period might not constitute ex-ante factors of expected futures returns (Gargano and Timmermann, 2014).

Therefore point forecasts of returns are generated for any given energy product i (i=*CL*, *HO*, *NG*) and maturity T (T=1,2,3 months) over the interval Jan.2004-Dec.2016. Specifically, we first split the original full sample (i.e. Feb.1990-Dec.2016) into two sub-periods. The initial in-sample period covers the interval from Feb.1990 to Dec.2003, spanning the first N=167 observations, while the out-of-sample period spans Jan.2004 to Dec.2016 (i.e. P=156 observations). The OoS sub-period constitute the forecast evaluation period.

For each of the predictive models, i.e. the Eq. (4.1e), (4.2c), and (4.4d), we generate point forecasts in a recursive rolling scheme. In other words, the starting date is fixed (anchored at start) and the window size grows as we move forward in time; each observation becoming available (known) is added to the in-sample sub-period and taken into consideration during the re-estimation of the model. This expanding window for estimating the coefficients is supposed to replicate the situation when the futures returns become known in real time and are used to generate ex-ante forecasts of the next period's futures returns (expected returns). To this end, we first estimate each model using the N=167 observations of the in-sample period (Feb.1990-Dec.2003) and obtain the OLS estimates of the coefficients. The fitted models are then used to produce h-step-ahead (h=1,3, or 12 months) OoS forecasts. For h=1 the first point forecast refers to Jan.2004.

$$[1, \dots, t, \dots, t+h]$$

Regarding our economic models, the initial 1-step-ahead OoS forecast is given as:

$$R^{iT}_{t:t+1} = \widehat{a_{0,t}} + \sum_{k=1}^{K=6} \widehat{a_{k,t}}\, x_{k,t} + \widehat{a_{7,t}}\, R^{iT}_{t-1:t} \tag{4.6}$$

where $\widehat{a_{0,t}}$, $\widehat{a_{k,t}}$, and $\widehat{a_{7,t}}$ are the OLS estimates generated by regressing the dependent variable $\left\{R^{iT}_{t:t+1}\right\}_{t=2}^{N}$ on a constant term $a_{0,t}$, $\left\{R^{iT}_{t}\right\}_{t=1}^{N-1}$ and $\left\{\sum_{k=1}^{K=6} \widehat{a_{k,t}}\, x_{k,t}\right\}_{t=1}^{N-1}$. Similarly, for the univariate autoregressive models *AR(1)*, the respective forecast is computes as:

$$R^{iT}_{t:t+1} = \widehat{a_{0,t}} + \widehat{\theta}_t\, R^{iT}_{t-1:t} \tag{4.7}$$

where $\widehat{\theta_t}$ is the OLS estimate from regressing $\left\{R^{iT}_{t:t+1}\right\}_{t=2}^{N}$ on a constant term $a_{0,t}$ and $\left\{R^{iT}_{t}\right\}_{t=1}^{N-1}$. As for the latent factor models:

$$R^{iT}_{t:t+1} = \widehat{a_{0,t}} + \sum_{j=1}^{m=3} \widehat{a_{j,t}}\, PC_{j,t} \quad (4.8)$$

where $\widehat{a_{0,t}}$ and $\widehat{a_{j,t}}$ are the OLS estimates generated by regressing $\left\{R^{iT}_{t:t+1}\right\}_{t=2}^{N}$ on a constant term $a_{0,t}$, and $\left\{\sum_{j=1}^{m=3} \widehat{a_{j,t}}\, PC_{j,t}\right\}_{t=1}^{N-1}$.

Next, the sample expands by one month, thus spanning the first 168 observations (t+1=N+1=167+1). The models are re-estimated based on the 168 observations and new coefficient values are obtained in order to generate the second 1-step-ahead OoS forecasts which correspond to Feb.2004 (i.e. the 169th value).

$$[1, \dots, t+1, \dots, t+1+h]$$

Concerning our economic models, the second 1-step-ahead OoS forecast is given as:

$$R^{iT}_{t+1:t+2} = \widehat{a_{0,t+1}} + \sum_{k=1}^{K=6} \widehat{a_{k,t+1}}\, x_{k,t+1} + \widehat{a_{7,t}}\, R^{iT}_{t:t+1} \quad (4.9)$$

where $\widehat{a_{0,t+1}}$, $\widehat{a_{k,t+1}}$, and $\widehat{a_{7,t+1}}$ are the OLS estimates generated by regressing $\left\{R^{iT}_{t+1:t+2}\right\}_{t=2}^{N+1}$ on a constant term term $a_{0,t+1}$, $\left\{R^{iT}_{t:t+1}\right\}_{t=1}^{N-1+1}$ and $\left\{\sum_{k=1}^{K=6} \widehat{a_{k,t}}\, x_{k,t}\right\}_{t=1}^{N-1+1}$. Similarly, we generate the second 1-step-ahead OoS point forecasts using the two alternative predictive models. Next, we re-estimate the model expanding our sample by one-month and so forth.

$$\begin{gathered}[1, \dots, t+2, \dots, t+2+h] \\ [1, \dots, t+3, \dots, t+3+h] \\ \vdots \\ [1, \dots, t+P, \dots, t+P+h]\end{gathered}$$

Repeating this process through the end of the OoS sub-period, we eventually create a series including P=156 OoS point forecasts of futures returns. Let $\left\{\hat{R}^{iT,j}_{t:t+1|t}\right\}_{t=1}^{P}$ denote the sequence of the one-month ahead (h=1) return forecasts generated by the j^{th} (j=*Econ., PCA, AR(1)*) model for a given commodity i (i=*CL, HO, NG*) maturing at T (T=1,2,3 months) .

4.2.2 Out-of-Sample Evaluation Criteria

4.2.2.1 Standard Performance Measures

In order to evaluate the OoS performance of each predictive model employed relative to a benchmark model, we first consider three forecast evaluation criteria (also known as performance measures or metrics) that have been predominantly used in the literature. Hence, for each model we employ the Root Mean Squared Prediction Error (RMSE), the Mean Absolute Prediction Error (MAE), and the Theil's inequality coefficient *U*.

Suppose using the j^{th} (j=*Econ., PCA, AR(1)*) model for a given commodity i (i=*CL, HO, NG*) maturing at T (T=1,2,3 months). Based on the j^{th} model, at any point in time an one-month ahead (h=1) point forecast using the information set up to period t (i.e., $\hat{R}^{iT}_{t+1|t}$ or $\hat{R}^{iT}_{t:t+1}$) produces the following prediction error, $\varepsilon^{j}_{t+1|t}$:

$$\varepsilon^{iT,j}_{t+1|t} = \hat{R}^{iT,j}_{t:t+1} - R^{iT}_{t:t+1} \tag{4.10}$$

where $\mathrm{R}^{\mathrm{iT}}_{\mathrm{t:t+1}}$ denotes the actual futures return realized in period t+1. Let $\left\{\hat{\mathrm{R}}^{\mathrm{iT,j}}_{\mathrm{t:t+1}}\right\}^{\mathrm{P}}_{\mathrm{t=1}}$ and $\left\{R^{iT}_{t:t+1}\right\}^{P}_{t=1}$ denote the sequence of the one-step ahead point forecasts and the sequence of the corresponding actual returns over the whole OoS sub-period, respectively. Consequently, the corresponding vector of errors is $\left\{\varepsilon^{iT,j}_{t+1|t}\right\}^{P}_{t=1}$, where t*=1(Jan.2004),2(Feb.2004),…, P(Dec.2016).*

Based on this prediction error that occurs at any point in time t *(t=1,2,…,P)* by the j^{th} model, the corresponding metrics are calculated as follows. In particular, RMSE is the most commonly used metric when evaluating the OoS fit of a model. It is calculated as the square root of the average squared deviations of the model based returns forecasts from the actual futures returns (Chantziara and Skiadopoulos, 2008). MAE constitutes the average of the absolute differences between the model based returns forecasts and the actual futures returns. Finally, the Theil's U is computed as follows.

$$RMSE^{iT,j} = \sqrt{\frac{1}{P}\sum_{t=1}^{P}(\varepsilon_{t+h|t}^{iT,j})^2} \tag{4.11a}$$

$$MAE^{iT,j} = \frac{1}{P}\sum_{t=1}^{P}\left|\varepsilon_{t+h|t}^{iT,j}\right| \tag{4.11b}$$

$$Theil's\ U^{\ iT,j} = \frac{RMSE^{iT,j}}{\sqrt{\frac{1}{P}\sum_{t=1}^{P}\hat{R}_t^{iT,j\ 2}} + \sqrt{\frac{1}{P}\sum_{t=1}^{P}R_t^{iT\ 2}}} \tag{4.11c}$$

Note that the Theil's *U* obtains values in the interval $[0.0, 1.0]$; values close to zero (0) indicate a rather good model specification, because the forecasts produced are very close to the actual data. On the contrary, values that are close to unity (1) reflect a bad model. As for the RMSE and MAE metrics, the smaller their values, the better predictive accuracy.

We next proceed to selecting the optimum specification model according to its forecasting accuracy. For any two specification models (*j=Econ., PCA, AR(1)*), we compute two ratios, the RMSE and the MAE ratios and conduct pairwise comparisons between the models. In particular, we compare the two ratios as follows: *AR(1)* versus *Econ.*, *AR(1)* versus *PCA*, and *Econ.* versus *PCA*. A ratio that is lower than one (<1) indicates a superior forecasting performance in favor of the model mentioned first. On the contrary, ratios that exceed unity (>1) indicate superior forecasting performance for the model mentioned second.

4.2.2.2 Test of Equal Predictive Accuracy

Apart from ratios based on the standard performance measures, we proceed to selecting the optimum specification model according to statistical tests of their predictive accuracy. This is achieved by determining whether the OoS fit of a benchmark model is (on average) better or worse than that of an alternative/competing specification in a statistically significant sense. In our case, the univariate autoregressive model (*j=AR(1))* serves as the benchmark model and is compared to its competing models, i.e. the economic variables (*j=Econ.)* and the PCA

(j=*PCA)* model. Univariate autoregressive models have commonly been used as baseline models in the related literature.

Such an evaluation is made in terms of loss functions, which are associated with the prediction errors. Specifically, under the RMSE and MAE metrics let the respective loss functions for the benchmark *AR(1)* model and the j^{th} model (j=*Econ., PCA*) be

$$g\left(\hat{R}_{t:t+1}^{iT,\ AR(1)}, R_{t:t+1}^{iT} \right) = g(\ \varepsilon_{t+1|t}^{iT,AR(1)}) \qquad (4.12a)$$

and

$$g(\ \hat{R}_{t:t+1}^{iT,j}, R_{t:t+1}^{iT}) = g(\ \varepsilon_{t+1|t}^{iT,j}) \qquad (4.12b)$$

where $\left\{\varepsilon_{t+1|t}^{iT,AR(1)}\right\}_{t=1}^{P}$ and $\left\{\varepsilon_{t+1|t}^{iT,j}\right\}_{t=1}^{P}$. A loss differential is also defined based on the (4.12a) and (4.12b) loss functions:

$$d_{t+1|t}^{iT,j} = g\left(\ \varepsilon_{t+1|t}^{iT,j}\right) - g(\ \varepsilon_{t+1|t}^{iT,AR(1)}) \qquad (4.12c)$$

The corresponding vector is $\left\{d_{t+1|t}^{iT,j}\right\}_{t=1}^{P}$. Then the statistical significance of the difference between the benchmark and the competing j^{th} model's loss function is assessed based on the modified Diebold-Mariano (MDM hereafter) test proposed by Harvey et al. (1997) [23].

The MDM test based on the average loss differential (Sq. or Abs. Errors) examines the null hypothesis (H_0) of equal predictive accuracy (on average) against two alternative hypotheses (one-sided tests), both indicating that the two models examined do not perform equally well. For $H_1\colon E\left(d_{t+1|t}^{iT,j}\right) > 0$, the benchmark *AR(1)* model outperforms the j^{th} model, while for $H_2\colon E\left(d_{t+1|t}^{iT,j}\right) < 0$, the respective j^{th} model outperforms the benchmark *AR(1)*. That is because positive values (on average) reflect greater prediction errors. In particular,

$$\begin{aligned}
&H_0\colon E\left(d_{t+1|t}^{iT,j}\right) = 0 \quad \text{or} \quad H_0\colon E\left[g\left(\ \varepsilon_{t+1|t}^{iT,j}\right)\right] = E\left[g\left(\ \varepsilon_{t+1|t}^{iT,AR(1)}\right)\right] \\
&H_1\colon E\left(d_{t+1|t}^{iT,j}\right) > 0 \quad \text{or} \quad H_1\colon E\left[g\left(\ \varepsilon_{t+1|t}^{iT,j}\right)\right] > E\left[g\left(\ \varepsilon_{t+1|t}^{iT,AR(1)}\right)\right] \\
&H_2\colon E\left(d_{t+1|t}^{iT,j}\right) < 0 \quad \text{or} \quad H_2\colon E\left[g\left(\ \varepsilon_{t+1|t}^{iT,j}\right)\right] < E\left[g\left(\ \varepsilon_{t+1|t}^{iT,AR(1)}\right)\right] \qquad (4.13)
\end{aligned}$$

[23] An alternative method to evaluate the statistical significance of the OoS results of the aforementioned metrics would be the Clark and West's (2007) t-statistic (see Gargano and Timmermann, 20014).

For one-step ahead (h=1month) point forecasts the MDM t-statistic is calculated as

$$MDM^{iT,j} = \frac{\overline{d_{t+1|t}^{iT,j}}}{\widehat{se}(\overline{d_{t+1|t}^{iT,j}})} \tag{4.14a}$$

and follows a Student's t-distribution with P-1 degrees of freedom (*d.f.*) rather than the standard normal distribution. Therefore, the $MDM^{iT,j}$ t-statistic is compared to the critical values. Note that $\overline{d_{t+1|t}^{iT,J}}$ and $\widehat{se}(\overline{d_{t+1|t}^{iT,J}})$ are the mean of the loss differential (i.e. $\overline{d_{t+1|t}^{iT,J}} = \frac{1}{P}\sum_{t=1}^{P} d_{t+1|t}^{iT,j}$) and a Newey-West estimator of the standard deviation of $\overline{d_{t+1|t}^{iT,J}}$, respectively Harvey et al. (1997.

Chapter 5: Empirical Results and Discussion

This section presents and discusses the empirical findings. Moreover, the results are also compared with those of the existing literature. First, we present the empirical results of the in-sample analysis conducted. We then show and evaluate those of the out-of-sample analysis.

5.1 In-Sample Evidence

Table 5.1 presents in-sample least squares estimates for all the alternative predictive one-step-ahead regression models employed within the context of this study. The estimation sample used to produce point forecasts covers the period Feb.1990-Dec.2003; one observation was lost due to calculating the futures returns. In particular, Panel A, Panel B, and Panel C show the performance of the economic variables, PCA, and AR(1) models, respectively. Estimates of each parameter are reported for each individual regression. The estimated OLS coefficients are followed by the corresponding Newey-West *t*-statistics in parentheses, the adjusted R^2, and the *F*-statistics's *p*-values.

Regarding our economic model (Equation 4.1e), the WTI crude oil futures seem to have the lowest values of adjusted R^2 (see Panel A). On the contrary, the adjusted R^2 associated with the natural gas futures, and especially this of the second shortest maturity contract, seem to be the greatest (i.e.11.79%). With the exception of R_{NG3}, we note that the adjusted R^2 rises as the corresponding futures' maturity lengthens. These findings are also confirmed by the *F*-statistics. These findings are in line with Killian and Vega (2011); in short, macroeconomic news of the US economy affect mostly the longer maturity contracts. We can also see that the statistically significant factors for R_{NG2} are the lagged open interest growth rate and the lagged changes in default spread (at 1%), as well as the lagged world steel production growth rate (at 5%).

There is also evidence that the latter possesses predictive power to all of the energy futures under examination (at 5% significance level for crude oil and natural gas, while at 1% for heating oil). This is in line with Ravazzolo and Vespignani's (2017) findings that the World Steel Production possesses a significant predictive power over crude oil prices. Ravazzolo and Vespignani also provide evidence of natural gas predictability by two alternative measures of global real economic activity, the OECD IP and the Killian's rea. However, Gargano and Timmermann (2014) find limited predictive ability of Killian's rea on various commodity returns, Alquist et al. (2013) reveal a negative impact on oil futures returns.

On the other hand, the results indicate that the first two factors account only for the natural gas futures. Regarding the median open interest growth, our findings is in line with those of Cummins et al. (2016); employing a five-factor model to study the predictability of oil futures returns over the interval 2007-2016 they find that speculation is not statistically significant. This is also in line with Bastianin et al. (2012), while in contrast to Hong and Yogo's (2012) findings. As for the default spread, our findings are in line with Bessembinder and Chan (1992), who find limited predictive power of the junk bond premium in 12 futures markets. Gargano and Timmermann (2014) also report its modest predictability over the spot prices of various commodity indices (i.e. textiles, metals, livestock, foods, etc.), however, with a positive sign.

This also seems to be the case for the term spread; our conclusion of positive, however, insignificant term spread conforms to the findings of Alquist et al. (2013) regarding the NYMEX crude oil futures of 3, 6, and 12 maturity over 1989-2012. On the other hand, we find negative, however, insignificant results in the case of the natural gas futures. According to Alquist et al. (2013), accumulation of oil as inventories tends to be high during recessions. As a result, an upward sloping term structure of interest rates and a positive change of its slope create expectations of economic expansion in the future, during which time energy commodities are expected to be relatively more expensive. However, the findings reveal a negative relationship between the term spread and the natural gas returns. Concerning this finding, we guess that this is a consequence of the natural gas storage costs. As far as we are concerned, the cost of keeping natural gas reserves is significantly higher than the cost of crude and heating oil storage, which probably discourages traders from holding inventories during periods of economic contraction. We form this hypothesis with respect to consumers as well as producers, who generally have direct interests in the natural gas.

Furthermore, our findings regarding the Trade-Weighted Dollar Index are in contrast to Cummins et al. (2016); they provide evidence of strong predictability and positive effect on the shortest maturity oil futures contracts. Reboredo and Revara-Castro (2013) and Gargano and Timmermann (2014) have also confirmed the significant impact of the US exchange rates on oil prices and commodity returns, respectively.

Table 5.1 (Panel B) also shows the results from the PCA model (Equation 4.4.d). Even though the lagged values of the second principal component (PC_2) are found to be statistically significant for the WTI crude oil and heating oil futures, the results indicate an overall poor fitness of the PCA model; the *F*-statistics indicate that the PCs are not capable of predicting the expected energy futures returns. The highest adjusted R^2 value (1.87%) appears for R_{CL1}.

This is also in line with Killian and Vega (2011). Regarding the PC_1, which was found to represent the CBOE S&P 100 Volatility Index in Subsection *3.4.2*, the expected negative - however insignificant - impact is evident. As for PC_2, which was found to be moderately correlated with the Consumer Sentiment Index, the expected positive relationship is also confirmed. We highlight again that the three extracted common factors are exclusively associated with the US economic conditions, rather than capturing both US and global macroeconomic aggregates.

Moreover, Panel C of Table 5.1 reports the results from the AR(1) models (Equation 4.2c). The adjusted R^2 for nearly all the energy futures examined seems to take rather low values. The greatest value for the adjusted R^2 it obtained for R_{NG3} (2.47%). Based on the *F*-statistics, the findings are mixed. It seems that the *AR(1)* models hold for the longer maturity energy futures (R_{CL2}, R_{CL3}, R_{HO3}, R_{NG3}). In these cases we reject the weak-form efficiency of the futures market. This is in line with Chantziara and Skiadopoulos (2008) who, studying the NYMEX crude oil and heating oil futures over the period Jan.1993-Dec.2003, find statistical insignificant coefficients and zero R^2 values.

Overall, concerning the in-sample one-step ahead (h=1 month) futures return predictability (individually), we can infer that the most appropriate models to use are the economic variables models. The economic variables models employed appear to fit satisfactorily almost all the energy products examined (seven out of the nine futures). Predictability is found to be strongest for the natural gas futures (1%). Alternatively, the second best performing models would be the univariate autoregressive *AR(1)*, which accounts for almost half of the futures (four (4) out of the nine (9) futures examined). As for the employed predictors, those factors commonly found to predict equities and bonds premia (real short-term interest rate, term spread, and default spread) are the least effective in the case of commodities. This is also in line with Gargano and Timmermann (2014) who find modest predictability over commodity prices.

Table 5.1

In-sample one-step ahead (h=1) individual predictive regressions.

	Dependent Variable								
	Panel A: Economic Model								
	$R_{CL1,\,t}$	$R_{CL2,\,t}$	$R_{CL3,\,t}$	$R_{HO1,\,t}$	$R_{HO2,\,t}$	$R_{HO3,\,t}$	$R_{NG1,\,t}$	$R_{NG2,\,t}$	$R_{NG3,\,t}$
c_1	0.000	0.000	0.000	0.000	0.001	0.001	0.001	0.000	0.000
	(-0.034)	(0.034)	(0.048)	(-0.015)	(0.093)	(0.092)	(0.070)	(-0.334)	(-0.334)
oi_gr $_{t-1}$	0.055	0.048	0.465	0.100	0.121	0.102	**0.494*****	**0.356*****	0.109
	(0.711)	(0.704)	(0.756)	(1.031)	(1.514)	(1.509)	(4.270)	(3.385)	(1.379)
d(rir) $_{t-1}$	1.957	1.923	1.847	2.350	2.737	2.650	-1.517	-1.977	-2.352
	(0.631)	(0.653)	(0.663)	(0.845)	(0.970)	(0.980)	(-0.417)	(-0.566)	(-0.738)
d(ts) $_{t-1}$	2.060	1.904	1.702	0.598	1.219	1.619	-5.835	-4.331	-3.627
	(0.687)	(0.696)	(0.691)	(0.191)	(0.425)	(0.641)	(-1.138)	(-0.898)	(-0.868)
d(dfs) $_{t-1}$	-16.547	-15.425	-14.834	-15.096	-13.247	**-15.279***	**-36.306*****	**-44.778*****	**-39.558*****
	(-1.596)	(-1.556)	(-1.532)	(-1.584)	(-1.473)	(-1.614)	(-2.625)	(-3.225)	(-3.534)
d(twdi) $_{t-1}$	0.001	0.000	0.000	-0.002	-0.003	-0.002	0.006	0.009	**0.011***
	(0.137)	(-0.024)	(-0.032)	(-0.308)	(-0.410)	(-0.272)	(0.566)	(1.026)	(1.696)
d(lwsp) $_{t-1}$	**1.371****	**1.257****	**1.174****	**2.106*****	**1.619*****	**1.447*****	**1.756****	**1.790****	**1.474****
	(2.270)	(2.346)	(2.407)	(3.977)	(2.899)	(2.638)	(1.983)	(2.407)	(2.360)
$R_{i,\,t-1}$	0.075	0.108	0.120	-0.016	0.058	0.105	-0.015	0.045	**0.131***
	(0.850)	(1.329)	(1.165)	(-0.195)	(0.736)	(1.334)	(-0.223)	(0.618)	(1.737)
F-stat	1.422	1.610	**1.778***	**1.878***	**1.878***	**2.146****	**4.058*****	**4.149*****	**3.346*****
Adj. R^2 (%)	*1.760*	*2.522*	*3.196*	*3.593*	*3.592*	*4.637*	*11.485*	*11.785*	*9.052*

The current table's entries report results from individually regressing the NYMEX WTI crude oil, heating oil, and natural gas first three shortest maturity futures returns on a set of lagged economic and financial variables over the in-sample period Feb.1990-Dec.2003. In particular, the following specification is estimated by Ordinary Least Squares (OLS): $R_{i,t} = \alpha_0 + a_1\, oi\_gr_{t-1} + a_2\, d(rir)_{t-1} + a_3\, d(ts)_{t-1} + a_4\, d(dfs)_{t-1} + a_5\, d(twdi)_{t-1} + a_6\, d(lwsp)_{t-1} + a_7\, R_{i,t-1} + \varepsilon_t$, where R_i: the i-th commodity futures monthly prices, log-differenced, oi_gr: the median open interest growth rate of all the underlying assets examined, rir: the changes of the short- term real interest rate, calculated as the difference between the nominal interest rate (i.e. the 3-month Treasury Bill rate) and the expected inflation (i.e. median expected change of prices during the following 12 months), ts: the changes of the yield spread calculated as the difference between the yield of the 10-year Government Bond and the 3-month Treasury Bill rate, dfs: the changes of the default spread calculated as the difference between the yield of the Moody's Seasoned BAA Corporate Bond and the Moody's Seasoned AAA Corporate Bond, $twdi$: the changes of the Trade Weighted U.S. Dollar Index (with index being Jan 1997=100), wsp: the Ravazzolo and Vespignani's (2016) world steel production in log-differences. The above-mentioned specification's OLS estimates of the parameters along with their corresponding t-statistics within parentheses, the adjusted R^2, and the F-statistics are also reported in the table. The t-statistics have been adjusted for both heteroscedasticity and serial correlation of unknown form using the Newey-West method (HAC Consistent Covariance). The asterisks *, **, and *** denote a rejection of the null hypothesis (i.e. a zero coefficient for the individual t-statistics or all coefficients are zero for the F-statistic) at 10%, 5% and 1% significance level, respectively.

Table 5.1 (Cont'd.)

In-sample one-step ahead (h=1) individual predictive regressions.

	Panel B: PCA Model								
	$R_{CL1,t}$	$R_{CL2,t}$	$R_{CL3,t}$	$R_{HO1,t}$	$R_{HO2,t}$	$R_{HO3,t}$	$R_{NG1,t}$	$R_{NG2,t}$	$R_{NG3,t}$
c_1	0.004	0.004	0.004	0.004	0.004	0.004	0.010	0.010	0.009
	(0.645)	(0.666)	(0.676)	(0.683)	(0.653)	(0.644)	(0.904)	(1.002)	(1.063)
$PC_{1,t-1}$	-0.001	0.000	0.000	-0.001	-0.001	0.000	0.000	0.000	0.000
	(-0.557)	(-0.509)	(-0.462)	(-0.735)	(-0.532)	(-0.458)	(0.242)	(0.127)	(-0.011)
$PC_{2,t-1}$	**0.004****	**0.004****	**0.004****	0.003	**0.004****	**0.004****	0.002	0.004	0.002
	(1.992)	(1.927)	(1.982)	(1.160)	(2.108)	(2.120)	(0.402)	(0.784)	(0.812)
$PC_{3,t-1}$	0.001	0.001	0.001	0.001	0.000	0.000	-0.004	-0.004	-0.002
	(0.536)	(0.581)	(0.604)	(0.403)	(0.126)	(0.116)	(-0.636)	(-0.886)	(-0.562)
F-stat	1.889	1.771	1.798	0.714	1.598	1.631	0.282	0.680	0.294
Adj. R^2	*1.870*	*1.625*	*1.681*	*-0.616*	*1.266*	*1.335*	*-1.563*	*-0.691*	*-1.537*

The current table's entries report results from individually regressing the NYMEX WTI crude oil, heating oil, and natural gas first three shortest maturity futures returns on a set of lagged latent factors extracted from the application of Joint PCA on McCracken and Ng's (2016) large dataset. In order to obtain validate estimates for the PCs, the 134 monthly macroeconomic US indicators included in the dataset are first transformed into stationary series based on McCracken and Ng's suggested data transformation. The latent factors retained and used as control variables in the following specification are the first three principal components, i.e. PC_1, PC_2, and PC_3. In particular, the specification is estimated by Ordinary Least Squares (OLS) is: $R_{i,t} = \alpha_0 + a_1 PC_{1,t-1} + a_2 PC_{2,t-1} + a_3 PC_{3,t-1} + \varepsilon_t$, , where R_i: the i-th commodity futures monthly prices, log-differenced. The above-mentioned specification's OLS estimates of the parameters along with their corresponding t-statistics within parentheses, the adjusted R^2, and the F-statistics are also reported in the table. The t-statistics have been adjusted for both heteroscedasticity and serial correlation of unknown form using the Newey-West method

(HAC Consistent Covariance). The asterisks *, **, and *** denote a rejection of the null hypothesis (i.e. a zero coefficient for the individual t-statistics or all coefficients are zero for the F-statistic) at 10%, 5% and 1% significance level, respectively. The in-sample estimation period corresponds to Feb.1990-Dec.2003.

Table 5.1 (Cont'd.)

In-sample one-step ahead (h=1) individual predictive regressions.

	Panel C: Univariate Autoregressive Model AR(1)								
	$R_{CL1,\,t}$	$R_{CL2,\,t}$	$R_{CL3,\,t}$	$R_{HO1,\,t}$	$R_{HO2,\,t}$	$R_{HO3,\,t}$	$R_{NG1,\,t}$	$R_{NG2,\,t}$	$R_{NG3,\,t}$
c_1	0.002	0.002	0.002	0.003	0.003	0.003	0.008	0.007	0.006
	(0.345)	(0.353)	(0.364)	(0.382)	(0.391)	(0.395)	(0.787)	(0.780)	(0.729)
$R_{i,\,t-1}$	0.105	0.138	**0.153***	-0.008	0.094	**0.144***	0.028	0.109	**0.175****
	(1.091)	(1.521)	(1.869)	(-0.103)	(1.105)	(1.657)	(0.411)	(1.417)	(2.303)
F-stat	1.841	**3.197***	**3.921****	0.012	1.453	**3.468***	0.129	1.949	**5.179****
Adj. R^2	*0.507*	*1.314*	*1.739*	*-0.603*	*0.274*	*1.474*	*-0.530*	*0.572*	*2.470*

The current table's entries report results from individually regressing the NYMEX WTI crude oil, heating oil, and natural gas first three shortest maturity futures returns on their one-period lagged values. In particular, a univariate autoregressive *AR(1)* model is estimated by Ordinary Least Squares (OLS), i.e. $R_{i,t} = \alpha_0 + \rho_1 R_{i,t-1} + \varepsilon_t$, where R_i: the i-th commodity futures monthly prices, log-differenced. The above-mentioned specification's OLS estimates of the parameters along with their corresponding t-statistics within parentheses, and the adjusted R^2, and the F-statistics are also reported in the table. The t-statistics have been adjusted for both heteroscedasticity and serial correlation of unknown form using the Newey-West method (HAC Consistent Covariance). The asterisks *, **, and *** denote a rejection of the null hypothesis (i.e. of a zero coefficient) at 10%, 5% and 1% significance level, respectively. The in-sample estimation period corresponds to Feb.1990-Dec.2003.

5.2 Out-of-Sample Forecasting Performance

We next estimate the three alternative predictive models described in Subsection *4.2.1*. For monthly, quarterly, and annual horizons (i.e. for h=1,3, and 12 months ahead, respectively), point forecasts of the futures returns are generated in a recursive scheme over the OoS sub-period, Jan.2004-Dec.2016. To this end, in this subsection we evaluate the OoS performance of the three model specifications employed to model each i-th (i=*CL, HO, NG*) energy futures maturing at T (T=1,2,3 months). We first calculate the standard OoS evaluation criteria described in Subsection *4.2.2.1* and then perform pairwise comparisons computing their ratios (RMSE and MAE ratios) as well as the corresponding Modified Diebold-Mariano (MDM) tests described in Subsection *4.2.2.2*.

In particular, Table 5.2 reports the results for the RMSE and MAE ratios obtained from each model specification. The two ratios are reported in the case of the three shortest maturity energy futures for each forecast horizon. We first consider the RMSE ratios. Admittedly, for each forecast horizon h and individual energy futures the benchmark *AR(1)* models appear rather superior to the PCA models. In addition, the *AR(1)* models seem to outperform the corresponding Economic models in all but one case; for each forecast horizon h, the Economic models perform better for *CL1* and *CL2,* as well as for *CL3* for quarterly horizon (h=3). This is somehow in line with the findings of the MAE ratios; for each forecast horizon h and energy futures the benchmark *AR(1)* models outperform the respective Economic and PCA models. However, PCA models perform better for *CL2* and *CL3* for annual horizons (h=12). Overall, the two ratios indicate that the benchmark *AR(1)* models perform better OoS.

Finally, we examine whether the proposed models by the standard evaluation criteria are also optimum under the terms of certain statistical tests of their predictive accuracy. Based on the modified Diebold-Mariano test (MDM), we examine the statistical significance of the average differencial loss function between the benchmark *AR(1)* model and the j^{th} competing model (j=Econ.,PCA). Table 5.3a, 5.3b, and 5.3c report the MDM t-statistics values of the relative predictive accuracy for monthly, quarterly, and annual horizons (h=1,3,12), respectively. One, two, and three asterisks (crosses) denote a rejection of the null in favor of the first (second) alternative hypothesis at 10%, 5%, and 1% significance level, respectively.

Specifically, for each forecast horizon h Panel A and B provide evidence for the rejection of the null hypothesis (of equal predictive accuracy) in favor of the first alternative hypothesis (H_1). As a result, the benchmark *AR(1)* model outperforms both the Economic and the PCA models for each individual commodity (at 1% and 5% significance level) and forecast horizon. In other words, the MDM findings confirm those obtained from the two standard OoS performance measures (i.e. RMSE and MAE metrics). In a statistically significant sense, the *AR(1)* model is the optimum model specification in order to model the NYMEX energy futures returns. Nevertheless, these findings stand in sharp contrast to the results of the Theil's U regarding the OoS performance of each predictive model. Theil's U (see Panel C of Table 5.3a, 5.3b, and 5.3c) obtains the lowest values in the case of the economic models. However, comparing the RMSE and MAE ratios along with the Theil's U of the *AR(1)* models for the forecast horizon of one month, we conclude that the out-of-sample predictability becomes stronger as maturity of the respective futures contract lengthens. We also note that this predictive ability does not become stronger as the forecast horizon extends.

These results are also in contrast to the finding of the in-sample analysis; the in-sample analysis proposed the economic variables models as the optimum specifications, while the *AR(1)* as the second best performing models. This is in line with Gargano and Timmermann (2014); in-sample inference of predictability of the employed economic predictors does not entail out-of-sample predictability.

Table 5.2
Out-of-sample performance of the various models.

h=1	R_{CL1}	R_{CL2}	R_{CL3}	R_{HO1}	R_{HO2}	R_{HO3}	R_{NG1}	R_{NG2}	R_{NG3}
					RMSE ratios				
AR(1)/Econ.	1.009	1.001	0.998	0.996	0.996	0.992	0.959	0.951	0.974
AR(1)/PCA	0.958	0.961	0.962	0.981	0.969	0.963	0.993	0.986	0.974
Econ./PCA	0.949	0.960	0.964	0.985	0.973	0.971	1.036	1.037	1.000
					MAE ratios				
AR(1)/Econ.	0.990	0.978	0.973	0.955	0.972	0.966	0.944	0.948	0.975
AR(1)/PCA	0.992	0.999	1.000	0.976	0.971	0.968	0.994	0.985	0.969
Econ./PCA	1.002	1.022	1.028	1.022	0.999	1.002	1.053	1.039	0.993
h=3	R_{CL1}	R_{CL2}	R_{CL3}	R_{HO1}	R_{HO2}	R_{HO3}	R_{NG1}	R_{NG2}	R_{NG3}
					RMSE ratios				
AR(1)/Econ.	1.010	1.004	1.002	0.996	0.997	0.994	0.959	0.948	0.971
AR(1)/PCA	0.951	0.955	0.956	0.981	0.966	0.958	0.990	0.981	0.970
Econ./PCA	0.941	0.951	0.954	0.985	0.969	0.964	1.032	1.035	0.998
					MAE ratios				
AR(1)/Econ.	0.993	0.984	0.981	0.955	0.972	0.969	0.942	0.943	0.974
AR(1)/PCA	0.986	0.993	0.993	0.976	0.968	0.964	0.988	0.981	0.965
Econ./PCA	0.993	1.009	1.012	1.023	0.996	0.995	1.050	1.039	0.992

For the forecast horizon of one, three, and twelve months (h=1,3,12) the table reports two ratios based on the Out-of-Sample (OoS) performance measures of the three predictive models employed within the context of this study, i.e. the Economic model, the joint PCA model, and the Univariate Autoregressive model AR(1). The performance measures used to compute the ratios are the Root Mean Squared Prediction Error (RMSE) and the Mean Absolute Prediction Error (MAE) for each of the three shortest maturity energy futures. Specifically, for any two specification models we compute the ratio of their RMSE values as well as the ratio of their MAE values. Consequently, the entries report the results as follows: *AR(1)* versus *Econ.*, *AR(1)* versus *PCA*, and *Econ.* versus *PCA*. A ratio that is lower than one (<1) indicates a superior forecasting performance in favor of the model mentioned first. On the contrary, ratios that exceed unity (>1) indicate superior forecasting performance for the model mentioned second. For each of the recursively estimated predictive models, the metrics have been calculated over the OoS sub-period (i.e. Jan.2004-Dec.2016).

Table 5.2 (Cont'd.)
Out-of-sample performance of the various models.

h=12	R_{CL1}	R_{CL2}	R_{CL3}	R_{HO1}	R_{HO2}	R_{HO3}	R_{NG1}	R_{NG2}	R_{NG3}
					RMSE ratios				
AR(1)/Econ.	1.010	1.001	0.998	0.988	0.992	0.989	0.937	0.919	0.956
AR(1)/PCA	0.964	0.965	0.965	0.982	0.972	0.965	0.991	0.985	0.975
Econ./PCA	0.954	0.964	0.966	0.993	0.980	0.976	1.057	1.072	1.020
					MAE ratios				
AR(1)/Econ.	0.998	0.987	0.985	0.944	0.967	0.966	0.918	0.914	0.954
AR(1)/PCA	0.999	1.004	1.003	0.977	0.973	0.972	0.993	0.982	0.975
Econ./PCA	1.001	1.017	1.018	1.036	1.006	1.006	1.082	1.074	1.022

For the forecast horizon of one, three, and twelve months (h=1,3,12) the table reports two ratios based on the Out-of-Sample (OoS) performance measures of the three predictive models employed within the context of this study, i.e. the Economic model, the joint PCA model, and the Univariate Autoregressive model AR(1). The performance measures used to compute the ratios are the Root Mean Squared Prediction Error (RMSE) and the Mean Absolute Prediction Error (MAE) for each of the three shortest maturity energy futures. Specifically, for any two specification models we compute the ratio of their RMSE values as well as the ratio of their MAE values. Consequently, the entries report the results as follows: *AR(1)* versus *Econ.*, *AR(1)* versus *PCA*, and *Econ.* versus *PCA*. A ratio that is lower than one (<1) indicates a superior forecasting performance in favor of the model mentioned first. On the contrary, ratios that exceed unity (>1) indicate superior forecasting performance for the model mentioned second. For each of the recursively estimated predictive models, the metrics have been calculated over the OoS sub-period (i.e. Jan.2004-Dec.2016).

Table 5.3a

One-step ahead (h=1 month) out-of-sample forecast accuracy: Modified Diebold-Mariano Test and Theil's U.

j^{th} Model vs. *AR(1)*	*Panel A: MDM Test Sq. Error (RMSE)*								
	R_{CL1}	R_{CL2}	R_{CL3}	R_{HO1}	R_{HO2}	R_{HO3}	R_{NG1}	R_{NG2}	R_{NG3}
Econ.	-2.753***	-2.548**	-2.476**	-2.848***	-3.005***	-2.814***	-3.769***	-7.014***	-7.587***
PCA	-9.184***	-9.761***	-9.826***	-6.221***	-9.708***	-9.781***	-6.435***	-8.452***	-6.665***
j^{th} Model vs. *AR(1)*	*Panel B: MDM Test Abs. Error (MAE)*								
	R_{CL1}	R_{CL2}	R_{CL3}	R_{HO1}	R_{HO2}	R_{HO3}	R_{NG1}	R_{NG2}	R_{NG3}
Econ.	-8.713***	-8.576***	-8.543***	-9.826***	-10.073***	-9.560***	-13.389***	-15.486***	-14.910***
PCA	-15.601***	-16.121***	-16.197***	-13.244***	-16.448***	-16.943***	-15.808***	-14.823***	-14.232***
	Panel C: Theil's U								
	R_{CL1}	R_{CL2}	R_{CL3}	R_{HO1}	R_{HO2}	R_{HO3}	R_{NG1}	R_{NG2}	R_{NG3}
AR(1)	0.848	0.833	0.825	0.942	0.875	0.840	0.960	0.897	0.852
Econ.	0.652	0.664	0.669	0.692	0.677	0.671	0.708	0.698	0.713
PCA	0.813	0.820	0.824	0.875	0.886	0.885	0.893	0.866	0.875

The entries of this table report the results regarding the Modified Diebold-Mariano Test (MDM) of Harvey et al. (1997). Panel A and Panel B contain the MDM t-statistic values of the relative predictive accuracy of the benchmark *AR(1)* model against the j^{th} competing model (j=Econ., PCA). The MDM test based on the average loss differential (i.e. Squared Errors and Absolute Errors in Panel A and B, respectively) and a Newey-West estimator of the standard deviation of $\overline{d_{t+1|t}^{iT,j}}$ examines the null hypothesis (H_0) against two alternative hypotheses, H_1 and H_2. In particular, H_0 indicates equal predictive accuracy between the benchmark *AR(1)* model and the j^{th} competing model, H_0:$E\left(d_{t+1|t}^{iT,j}\right) = 0$. On the contrary, H_1:$E\left(d_{t+1|t}^{iT,j}\right) > 0$ indicates that the benchmark outperforms the j^{th} model, while for H_2: $E\left(d_{t+1|t}^{iT,j}\right) < 0$, the competing j^{th} model outperforms the benchmark *AR(1)*. One (*), two (**) and three (***) asterisks denote a rejection of the null hypothesis (H_0) in favor of the first alternative (H_1) at 10%, 5%, and 1% significance level, respectively. Similarly, one (+), two (++) and three (+++) crosses denote a rejection of H_0 in favor of the second alternative hypothesis (H_2) at 10%, 5%, and 1% significance level, respectively. Panel C also reports the results regarding the third standard performance measure employed, the Theil's U. Results are reported for the three shortest maturity energy futures. For each of the recursively estimated predictive models, the metrics have been calculated over the OoS sub-period (i.e. Jan.2004-Dec.2016) for a forecast horizon of one month (h=1).

Table 5.3b

Three-step ahead (h=3 months) out-of-sample forecast accuracy: Modified Diebold-Mariano Test and Theil's U.

j^{th} *Model vs. AR(1)*	*Panel A: MDM Test Sq. Error (RMSE)*								
	R_{CL1}	R_{CL2}	R_{CL3}	R_{HO1}	R_{HO2}	R_{HO3}	R_{NG1}	R_{NG2}	R_{NG3}
Econ.	-4.070***	-3.909***	-3.863***	-4.240***	-4.614***	-4.300***	-3.967***	-3.368***	-3.418**
PCA	-8.745***	-9.536***	-9.673***	-7.187***	-9.665***	-9.657***	-5.596***	-6.786***	-6.914***
j^{th} *Model vs. AR(1)*	*Panel B: MDM Test Abs. Error (MAE)*								
	R_{CL1}	R_{CL2}	R_{CL3}	R_{HO1}	R_{HO2}	R_{HO3}	R_{NG1}	R_{NG2}	R_{NG3}
Econ.	-9.572***	-9.520***	-9.535***	-10.789***	-11.200***	-10.626***	-12.891***	-11.119***	-9.935***
PCA	-15.239***	-15.862***	-16.130***	-14.408***	-16.378***	-16.778***	-11.935***	-13.355***	-13.611***
	Panel C: Theil's U								
	R_{CL1}	R_{CL2}	R_{CL3}	R_{HO1}	R_{HO2}	R_{HO3}	R_{NG1}	R_{NG2}	R_{NG3}
AR(1)	0.861	0.846	0.837	0.945	0.884	0.849	0.961	0.901	0.856
Econ.	0.675	0.686	0.691	0.709	0.695	0.689	0.694	0.686	0.706
PCA	0.832	0.837	0.841	0.885	0.900	0.901	0.900	0.874	0.883

The entries of this table report the results regarding the Modified Diebold-Mariano Test (MDM) of Harvey et al. (1997). Panel A and Panel B contain the MDM t-statistic values of the relative predictive accuracy of the benchmark *AR(1)* model against the j^{th} competing model (j=Econ., PCA). The MDM test based on the average loss differential (i.e. Squared Errors and Absolute Errors in Panel A and B, respectively) and a Newey-West estimator of the standard deviation of $\overline{d_{t+1|t}^{iT,j}}$ examines the null hypothesis (H_0) against two alternative hypotheses, H_1 and H_2. In particular, H_0 indicates equal predictive accuracy between the benchmark *AR(1)* model and the j^{th} competing model, H_0:$E\left(d_{t+1|t}^{iT,j}\right) = 0$. On the contrary, H_1:$E\left(d_{t+1|t}^{iT,j}\right) > 0$ indicates that the benchmark outperforms the j^{th} model, while for H_2: $E\left(d_{t+1|t}^{iT,j}\right) < 0$, the competing j^{th} model outperforms the benchmark *AR(1)*. One (*), two (**) and three (***) asterisks denote a rejection of the null hypothesis (H_0) in favor of the first alternative (H_1) at 10%, 5%, and 1% significance level, respectively. Similarly, one ($^{+}$), two ($^{++}$) and three ($^{+++}$) crosses denote a rejection of H_0 in favor of the second alternative hypothesis (H_2) at 10%, 5%, and 1% significance level, respectively. Panel C also reports the results regarding the third standard performance measure employed, the Theil's U. Results are reported for the three shortest maturity energy futures. For each of the recursively estimated predictive models, the metrics have been calculated over the OoS sub-period (i.e. Jan.2004-Dec.2016) for a forecast horizon of three months (h=3).

Table 5.3c

Twelve-step ahead (h=12 months) out-of-sample forecast accuracy: Modified Diebold-Mariano Test and Theil's U.

j^{th} Model vs. AR(1)	*Panel A: MDM Test Sq. Error (RMSE)*								
	R_{CL1}	R_{CL2}	R_{CL3}	R_{HO1}	R_{HO2}	R_{HO3}	R_{NG1}	R_{NG2}	R_{NG3}
Econ.	-4.117***	-3.992***	-3.951***	-4.324***	-4.554***	-4.235***	-4.187***	-3.617***	-3.690**
PCA	-9.647***	-9.828***	-9.810***	-6.001***	-9.556***	-9.779***	-6.360***	-7.097***	-7.389***
j^{th} Model vs. AR(1)	*Panel B: MDM Test Abs. Error (MAE)*								
	R_{CL1}	R_{CL2}	R_{CL3}	R_{HO1}	R_{HO2}	R_{HO3}	R_{NG1}	R_{NG2}	R_{NG3}
Econ.	-9.978***	-9.809***	-9.775***	-10.830***	-11.406***	-10.741***	-12.385***	-10.630***	-9.656***
PCA	-15.809***	-16.199***	-16.388***	-14.635***	-16.374***	-16.819***	-13.514***	-13.187***	-13.687***
	Panel C: Theil's U								
	R_{CL1}	R_{CL2}	R_{CL3}	R_{HO1}	R_{HO2}	R_{HO3}	R_{NG1}	R_{NG2}	R_{NG3}
AR(1)	0.865	0.847	0.838	0.944	0.885	0.850	0.957	0.898	0.853
Econ.	0.679	0.687	0.690	0.702	0.693	0.686	0.682	0.671	0.688
PCA	0.824	0.831	0.836	0.884	0.903	0.904	0.906	0.875	0.882

The entries of this table report the results regarding the Modified Diebold-Mariano Test (MDM) of Harvey et al. (1997). Panel A and Panel B contain the MDM t-statistic values of the relative predictive accuracy of the benchmark *AR(1)* model against the j^{th} competing model (j=Econ., PCA). The MDM test based on the average loss differential (i.e. Squared Errors and Absolute Errors in Panel A and B, respectively) and a Newey-West estimator of the standard deviation of $\overline{d_{t+1|t}^{iT,j}}$ examines the null hypothesis (H_0) against two alternative hypotheses, H_1 and H_2. In particular, H_0 indicates equal predictive accuracy between the benchmark *AR(1)* model and the j^{th} competing model, H_0:$E\left(d_{t+1|t}^{iT,j}\right) = 0$. On the contrary, H_1:$E\left(d_{t+1|t}^{iT,j}\right) > 0$ indicates that the benchmark outperforms the j^{th} model, while for $H_2: E\left(d_{t+1|t}^{iT,j}\right) < 0$, the competing j^{th} model outperforms the benchmark *AR(1)*. One (*), two (**) and three (***) asterisks denote a rejection of the null hypothesis (H_0) in favor of the first alternative (H_1) at 10%, 5%, and 1% significance level, respectively. Similarly, one ($^{+}$), two ($^{++}$) and three ($^{+++}$) crosses denote a rejection of H_0 in favor of the second alternative hypothesis (H_2) at 10%, 5%, and 1% significance level, respectively. Panel C also reports the results regarding the third standard performance measure employed, the Theil's U. Results are reported for the three shortest maturity energy futures. For each of the recursively estimated predictive models, the metrics have been calculated over the OoS sub-period (i.e. Jan.2004-Dec.2016) for a forecast horizon of twelve months (h=12).

Chapter 6: Conclusions and Implications

Admittedly, petroleum products, such as crude and heating oil, along with natural gas constitute the primary non-renewable energy sources around the world; they have a fundamental role in determining macroeconomic aggregates and economic more broadly. Consequently, it is of great importance for any individual or organization/institute, who in one way or another participate in the energy markets, to know whether the future evolution of energy prices can be predicted and, if so, by which factors.

With few exceptions, however, there is a paucity of literature concerning the construction of well-established models capable of reliably describing and predicting individual energy futures contracts. There has been found only limited and inconclusive evidence concerning the predictability of energy futures dynamics. The controversy regarding the proper theoretical approach and the corresponding optimum predictive model lies in the 'hybrid role' of these products. Indicatively, they constitute both consumption and production assets, and alternative investment instruments; it is, therefore, reasonable for the energy futures market to attract both traditional commercial traders (i.e. hedgers such as direct consumers and producers) and the non-commercial traders (i.e. speculators and financial intermediaries), respectively.

Commodity prices are widely believed to be driven by time-varying storage costs and convenience yields, both influenced by the underlying state of the economy. To this end, focusing on the second strand of the relevant literature, we employ three competing linear model specifications and evaluate their forecasting performance both in-sample and out-of-sample on the energy futures returns. We examine the predictability of the three shortest maturity NYMEX crude oil, heating oil, and natural gas futures returns over the period Jan.1990-Dec.2016. Hence, the predictability of the expected futures returns is conducted in terms of common underlying risk-factors.

In particular, we first examine whether returns exhibit a predictable pattern by means of prominent risk-factors that are found to price equities and bonds as well as alternative investment instruments (i.e. commodities) rather successfully. Six macroeconomic, financial and commodity-specific factors have been examined as potential predictors in an economic model set up. In short, the economic predictors employed in this study are the open interest median growth rate, the short-term real interest rate, the term spread, the default spread, the trade-weighted US dollar index, the world steel production. We also consider incorporate the previous realized returns in the economic model.

In order to avoid a possible omitted variable or irrelevant variable bias and account for possible energy market's segmentation, we construct two competing predictive models; for each individual futures, we examine a univariate first-order autoregressive model as well as a latent factor model, respectively. The latter employs as potential predictors the first three principal components (PCs) extracted by a large macroeconomic database of 134 US economic indicators created by McCracken and Ng (2016) through the Principal Component Analysis (PCA). Subsequently, the principal components are used as state variables in a linear predictive model. We find that the first PC is almost perfectly correlated to the CBOE S&P 100 Volatility Index (VXO), apparently reflecting the investors' sentiment and expectations regarding the following month's volatility in the US stock market. Moreover PC2 and PC3 seem to be moderately correlated with two individual indicators, the Consumer Sentiment Index and the Help-Wanted Index.

Our findings reveal significant in-sample predictability of futures returns. For seven out of the nine futures under examination, the economic models are found to satisfactorily fit the corresponding returns. Overall, the risk-factors that, apparently, possess a statistically significant predictive power over the expected futures returns are the lagged median open interest growth rate, the lagged changes in default spread as well as the lagged world steel production growth rate. This predictability is found to be strongest for the natural gas futures (i.e. at 1% significance level). With the exception of R_{NG3}, we also note that the adjusted R^2 rises as the corresponding futures' maturity lengthens.

There is also evidence that the second best performing models in-sample are the univariate autoregressive *AR(1)*; they account for almost half of the futures (four (4) out of the nine (9) futures examined). However, in the out-of-sample (OoS) analysis this doe not seem to be the case; for each individual futures the OoS results indicate that the benchmark *AR(1)* models outperform both the Economic and the PCA models. Finally, in a statistically significant sense, we provide evidence that the *AR(1)* is the optimum model specification in order to model the NYMEX energy futures returns.

In light of these findings, we conclude that the NYMEX energy market is not efficient even in its weak-form; the current prices of the examined energy futures do not incorporate all available information at any point in time, thus creating a possible predictive pattern through the past realized returns. Therefore, technical analysis, relevant indicators as well as neural networks could be used to predict future returns. However, based on our findings, there is evidence that the NYMEX energy market does not violate the semi-strong form efficiency. That is because the informational content of the remainder economic risk-factors incorporated in the economic variables models is not beneficial enough to predict the futures expected returns. As a consequence, the *AR(1)* arises as the optimal model specification OoS. Comparing the RMSE and MAE ratios along with the Theil's U yielded form the *AR(1)* models for the forecast horizon of h=1 month, we conclude that the out-of-sample predictability becomes stronger as maturity of the respective futures contract lengthens. This, however, does not seem to hold for longer forecast horizons.

Eventually, the only thing that we could state with full conviction is that the predictability or not of the expected futures returns depends largely on the chosen predictive models. We thus provide a few suggestions regarding possible future research extensions of the current analysis. First and foremost, in line with Le Pen and Sévi (2011), Ludvigson and Ng (2009), and Gargano et al. (2016), it would be more appropriate to extract common latent factors form a database comprising of both US and global macroeconomic aggregates. The reasons lying behind this suggestion have already been extensively discussed. The McCracken and Ng's big data account only for US economic conditions, while the energy futures under consideration are traded globally. Furthermore, predictability should also be analyzed for longer maturity futures contracts, i.e. maturing at 1-12 months (e.g. *CL1-CL12)*. Moreover, regarding our economic models, commodity specific and technical indicators would be most suitable in case of energy market segmentation; segmentation has been confirmed by Daskalaki et al. (2014). As for the econometric approaches, more sophisticated variants of the current models should be employed (e.g. GARCH-type model specifications, both symmetric and asymmetric). Last but not least, as mentioned by Chantziara and Skiadopoulos (2008), such studies postulate that market participants hold a portfolio exclusively composed of the i-th (i=*CL*, *HO*, *NG*) energy product expiring at T months (T=1,2,3) rather than a well-diversified portfolio of energy futures; various speculative strategies could be constructed and analyzed, instead.

References

Akram, Q. F. (2009) ‘Commodity prices, interest rates and the dollar’. *Energy Economics* 31 (6), pp. 838-851. Available at: https://doi.org/10.1016/j.eneco.2009.05.016

Alquist, R., Bauer, G. H., and de los Rios, A. D. (2013) ‘Macroeconomic drivers of crude oil futures risk premia’. *Working Paper*, Bank of Canada. Available at: https://www.business.unsw.edu.au/About-Site/Schools-Site/banking-finance-site/Documents/Macroeconomic-Drivers-Of-Cruse-Oil-Futures-Risk-Premia.pdf

Anzuini, A., Lombardi, M. J., and Pagano, P. (2010) ‘The impact of monetary policy shocks on commodity prices’. *Working Paper No. 1232*, European Central Bank, pp.5-27. Available at: https://ssrn.com/abstract=1650595

Apergis, N., and Miller, S. M. (2009) ‘Do structural oil-market shocks affect stock prices?’. *Energy Economics* 31 (4), pp. 569-575. Available at: https://doi.org/10.1016/j.eneco.2009.03.001

Apergis, N., and Payne, J. E. (2010) ‘Natural gas consumption and economic growth: A panel investigation of 67 countries’. *Applied Energy* 87 (8), pp. 2759-2763. Available at: https://doi.org/10.1016/j.apenergy.2010.01.002

Baetje, F., and Menkhoff, L. (2016) ‘Equity premium prediction: Are economic and technical indicators unstable?’. *International Journal of Forecasting* 32 (4), pp. 1193-1207. Available at: https://doi.org/10.1016/j.ijforecast.2016.02.006

Baker, S. D. (2012) ‘The financialization of storable commodities’. *Working Paper*, Carnegie Mellon University. Available at: http://haskayne.ucalgary.ca/files/haskayne/bakeroilfinancialization_0.pdf

Barsky, R.B., and Kilian, L. (2002) ‘Do we really know that oil caused the great stagflation? A monetary alternative’. *NBER Working Paper No. 8389.* Available at: http://www.nber.org/papers/w8389

Barsky, R.B., and Kilian, L. (2004), ‘Oil and the macroeconomy since the 1970s’. *Journal of Economic Perspectives* 18 (4), pp. 115-134. Available at: https://wwz.unibas.ch/fileadmin/wwz/redaktion/makrooekonomie/intermediate_macro/reader/9/02_R_Oil_and_the_Macroeconomy.pdf

Bashiri, A., and Lawryshyn, Y. (2017) ‘A stochastic supply/demand model for storable commodity prices’. Available at: http://www.realoptions.org/openconf2017/data/papers/36.pdf

Bastianin, A., Manera, M., Nicolini, M., and Vignati, I. (2012) ‘Speculation, Returns, Volume and Volatility in Commodities Futures Markets’. *Review of Environment, Energy and Economics*. Available at: https://econpapers.repec.org/article/femfemre3/2012.01-07.htm

Bastourre, D., Carrera, J., and Ibarlucia, J. (2010) ‘Commodity prices: Structural factors, financial markets and non-linear dynamics’. *BCRA Working Paper No. 201050*, Bank of Argentina. Available at: http://www.bcra.gov.ar/pdfs/investigaciones/WP_50_2010i.pdf

Basu, D., and Miffre, J. (2013) ‘Capturing the risk premium of commodity futures: The role of hedging pressure’. *Journal of Banking & Finance* 37 (7), pp. 2652-2664. Available at: https://doi.org/10.1016/j.jbankfin.2013.02.031

Bessembinder, H., and Chan, K. (1992) ‘Time-varying risk premia and forecastable returns in futures markets’. *Journal of Financial Economics* 32 (2), pp. 169-193. Available at: https://doi.org/10.1016/0304-405X(92)90017-R

Campbell, J. Y. (1987) ‘Stock returns and the term structure’. *Journal of Financial Economics 18* (2), pp. 373-399. Available at: https://doi.org/10.1016/0304-405X(87)90045-6

Campbell, J. Y., and Thompson, S. B. (2008) ‘Predicting excess stock returns out of sample: Can anything beat the historical average?’. *The Review of Financial Studies* 21 (4), pp. 1509-1531. Available at: https://doi.org/10.1093/rfs/hhm055

Cavalcante, M. T. (2010) ‘An Analysis of the relationship between WTI term structure and oil market fundamentals in 2002-2009’. *MPRA Paper No. 24263*. Available at: https://mpra.ub.uni-muenchen.de/24263/

Chan, K., Tse, Y., and Williams, M. (2011) ‘The relationship between commodity price and currency exchange rate: Evidence from the futures markets’. *Commodity Prices and Markets, East Asia Seminar on Economics* 20, pp. 47-71. Available at: http://www.nber.org/chapters/c11859.pdf

Chantziara, T., and Skiadopoulos, G. (2008) ‘Can the dynamics of the term structure of petroleum futures be forecasted? Evidence from major markets’. *Energy Economics* 30 (3), pp. 962-985. Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1125623

Chen, Y.-C., Rogoff, K. S., and Rossi, B. (2010) ‘Can exchange rates forecast commodity prices?’. ERID *Working Paper No.1*, Duke University. Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1183164

Cummins, M., Dowling, M., and Kearney, F. (2016) ‘Oil market modelling: A comparative analysis of fundamental and latent factor approaches’. *International Review of Financial Analysis* 46, pp. 211-218. Available at: https://pure.qub.ac.uk/portal/files/43048158/oil_market_modelling.pdf

Daskalaki, C., Kostakis, A., and Skiadopoulos, G. (2014) ‘Are there common factors in commodity futures returns?’. *Journal of Banking & Finance* 40, pp. 346-363 Available at: https://doi.org/10.1016/j.jbankfin.2013.11.034

Diebold, F. X., and Mariano, R. S. (1995) ‘Comparing predictive accuracy’. *Journal of Business & Economic Statistics* 13 (3), pp. 256-263. Available at: http://eurequa.univ-paris1.fr/membres/Ahamada/cours/article1.pdf

Erb, C. B., and Harvey, C. R. (2006) 'The strategic and tactical value of commodity futures', *Financial Analysts Journal* 62 (2), pp. 69-97. Available at: https://faculty.fuqua.duke.edu/~charvey/Research/Published_Papers/P91_The_strategic_and.pdf

Estrella, A., and Hardouvelis, G. A. (1991) 'The Term Structure as a Predictor of Real Economic Activity'. *The Journal of Finance* 46 (2), pp. 555-576. Available at: http://www.hardouvelis.gr/FILES/PUBLICATIONS/JF_June_1991_46(2)_555-76_HQ.pdf

Fama, E. F., and French, K. R. (1989) 'Business conditions and expected returns on stocks and bonds'. *Journal of Financial Economics* 25 (1), pp. 23-49. Available at: https://doi.org/10.1016/0304-405X(89)90095-0

Frankel, J.A. (1984), 'Commodity prices and money: lessons from international finance'. *American Journal of Agricultural Economics* 66 (5), pp. 560-566. Available at: https://sites.hks.harvard.edu/fs/jfrankel/commodityprices.pdf

Frankel, J. A. (2006) 'The effect of monetary policy on real commodity prices'. *NBER Working Paper No. 12713*, pp. 1-40. Available at: http://www.nber.org/papers/w12713

Frankel, J.A. (2014), 'Effects of speculation and interest rates in a "carry trade" model of commodity prices'. *Journal of International Money and Finance* 42, pp. 88-112. Available at: https://doi.org/10.1016/j.jimonfin.2013.08.006

Frankel, J.A., and Hardouvelis, G. A. (1983), 'Commodity Prices, Overshooting, Money Surprises, and Fed Credibility'. *NBER Working Paper No. 1121*. Available at: http://www.nber.org/papers/w1121.pdf

Gargano, A. (2013) 'Return predictability in recessions: An asset pricing perspective'. *Job Market Paper*. Available at: https://pdfs.semanticscholar.org/2dc4/607f5df31d214de7f4581489c102c109ff52.pdf

Gargano, A., and Timmermann, A. G. (2014) 'Forecasting commodity price indexes using macroeconomic and financial predictors'. *International Journal of Forecasting* 30 (3), pp. 825-843. Available at: https://doi.org/10.1016/j.ijforecast.2013.09.003

Gargano, A., Pettenuzzo, D., and Timmermann, A. G. (2016) 'Bond return predictability: Economic value and links to the macroeconomy'. *Working Paper No. 75R*, Brandeis University. Available at: http://dx.doi.org/10.2139/ssrn.2471273

Gorton, G., and Rouwenhorst, K. G. (2004) 'Facts and fantasies about commodity futures'. *NBER Working Paper No. 10595*. Available at: http://www.nber.org/papers/w10595

Gorton, G. B., Hayashi, F., and Rouwenhorst, K. G. (2012) 'The fundamentals of commodity futures returns'. *Yale ICF Working Paper No. 07-08*. Available at: https://ssrn.com/abstract=996930

Goyal, A., and Welch, I. (2008) ‘A comprhensive look at the empirical performance of equity premium prediction’. *Review of Financial Studies* 21 (4), pp. 1455-1508. Available at: https://ssrn.com/abstract=517667

Groen, J. J. (2014) ‘Forecasting commodity price indexes using macroeconomic and financial predictors’. *International Journal of Forecasting* 30, pp. 844-846. Available at: 10.1016/j.ijforecast.2013.09.002

Gross, C. (2017) ‘Examining the Common Dynamics of Commodity Futures Prices’. *CQE Working Papers No. 6317*, University of Muenster. Available at: https://www.wiwi.uni-muenster.de/cqe/sites/cqe/files/CQE_Paper/cqe_wp_63_2017.pdf

Haase, M., and Zimmermann, H. (2013) ‘Scarcity, risk premiums, and the pricing of commodity futures: The case of crude oil contracts’. *The Journal of Alternative Investments* 16 (1), pp. 43-71. Available at: https://doi.org/10.3905/jai.2013.16.1.043

Hamilton, J. D. (2009) ‘Causes and consequences of the oil shock of 2007-2008’. *NBER Working Paper No. 15002*. Available at: http://www.nber.org/papers/w15002

Hamilton, J. D., and Wu, J. C. (2013) ‘Risk Premia in Crude Oil Futures Prices’. *NBER Working Paper No. 19056*. Available at: http://www.nber.org/papers/w19056

Harvey, D., Leybourne, S., and Newbold, P. (1997) ‘Testing the equality of prediction mean squared errors’. *International Journal of Finance* 13 (2), pp. 281-291. Available at: https://doi.org/10.1016/S0169-2070(96)00719-4

Hong, H., and Yogo, M. (2012) ‘What does futures market interest tell us about the macroeconomy and asser prices?’. *Journal of Financial Economics* 105 (3), pp. 473-490. Available at: https://doi.org/10.1016/j.jfineco.2012.04.005

Killian, L. (2009) ‘Not all oil price shocks are alike: Disentangling demand and supply shocks in the crude oil market’. *Americal Economic Review* 99 (3), pp. 1053-1069. Available at: http://www-personal.umich.edu/~lkilian/aer061308final.pdf

Killian, L. (2017) ‘How the tight oil boom has changed oil and gasoline markets’. *CESifo Working Paper No. 6380*. Available at: http://www.cesifo-group.de/DocDL/cesifo1_wp6380.pdf

Killian, L., and Vega, C. (2011) ‘Do energy prices respond to U.S. macroeconomic mews? A test of the hypothesis of predetermined energy prices’. *The Review of Economics and Statistics* 93 (2), pp. 660-671. Available at: https://doi.org/10.1162/REST_a_00086

Kolev, G. I., and Karapandza, R. (2017) ‘Out-of-sample equity premium predictability and sample split–invariant inference’. *Journal of Banking & Finance* 84, pp. 188-201. Available at: https://doi.org/10.1016/j.jbankfin.2016.07.017

Kostakis, A., Magdalinos, T., and Stamatogiannis, M. (2015) ‘Robust Econometric Inference for Stock Return Predictability’. *The Review of Financial Studies* 28 (5), pp. 1506-1553. Available at: https://doi.org/10.1093/rfs/hhu139

Le Pen, Y., and Sévi, B. (2012) ‘Macro factors in oil futures returns’. *International Economics* 126-127 (2), pp. 13-38. Available at: https://doi.org/10.1016/S2110-7017(13)60034-7

Lübbers, J., and Posch, P. N. (2016) ‘Commodities' common factor: An empirical assessment of the markets' drivers’. *Journal of Commodity Markets* 4 (1), pp. 28-40. Available at: https://doi.org/10.1016/j.jcomm.2016.10.002

Ludvigson, S., and Ng, S. (2009) ‘A Factor Analysis of Bond Risk Premia’. N*BER Working Paper No. 15188.* Available at: http://www.nber.org/papers/w15188

McCracken, M. W., and Ng, S. (2016) ‘FRED-MD: A monthly database for macroeconomic research’. *Journal of Business and Economic Statistics* 34 (4), pp. 574-589. Available at: https://files.stlouisfed.org/files/htdocs/wp/2015/2015-012.pdf

Medlock III, K. B., and Jaffe, A. M. (2009) ‘Who is in the oil futures market and how has it changed?’. *James A. Baker III Institute for Public Policy,* Rice University. . Available at: https://www.bakerinstitute.org/files/512/

Melolinna, M. (2011) ‘What explains risk premia in crude oil futures?’. *Bank of Finland Research Discussion Paper No.2/2011*. Available at: http://dx.doi.org/10.2139/ssrn.1763231

Mileva, E., and Siegfried, N. (2012) ‘Oil market structure, network effects and the choice of currency for oil invoicing’. *Energy Policy* 44, pp. 358-394. Available at: https://doi.org/10.1016/j.enpol.2012.02.002

Mishra, V., and Smyth R. (2016) ‘Are natural gas spot and futures prices predictable’. *Economic Modelling* 54, pp. 178-186. Available at: https://doi.org/10.1016/j.econmod.2015.12.034

Radetzki, M. (2006) ‘The anatomy of three commodity booms’. *Resources Policy* 31, pp. 56-64. Available at: http://www.radetzki.biz/rapporter/ThreeBooms_71.pdf

Pradhananga, M. (2015) ‘Financialization and the rise in co-movement of commodity prices’. *PERI Working Paper No. 376*, University of Massachusetts Amherst. Available at: http://www.peri.umass.edu/fileadmin/pdf/working_papers/working_papers_351-400/WP376.pdf

Rapach, D. E., and Wohar, M. E. (2006) ‘In-sample vs. out-of-sample tests of stock return predictability in the context of data mining’. *Journal of Empirical Finance* 13 (2), pp. 231-247. Available at: https://doi.org/10.1016/j.jempfin.2005.08.001

Rapach, D. E., Strauss, J. K., and Zhou, G. (2013) ‘International stock return predictability: What is the role of the United States?’. *Journal of Finance* 68 (4), pp. 1633-1662. Available at: http://www.jstor.org/stable/42002633

Ravazzolo, F., and Vespignani, J. L. (2016) ‘A new monthly indicator of global real economic activity’. Available at: https://www.ecb.europa.eu/pub/conferences/shared/pdf/20160603_forecasting/Paper_for_poster_5_Ravazzolo.pdf

Ravazzolo, F., and Vespignani, J. L. (2017) 'World steel production: A new monthly indicator of global real economic activity'. *CAMA Working Paper No. 42*. Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2992983

Reboredo, J. C., and Revera-Castro, M. A. (2013) 'A wavelet decomposition approach to crude oil price and exchange rate dependence'. *Economic Modelling* 32, pp. 42-57. Available at: https://doi.org/10.1016/j.econmod.2012.12.028

Sadorsky, P. (2002) 'Time-varying risk premiums in petroleum futures prices'. *Energy Economics* 24 (6), pp. 539-556. Available at: https://doi.org/10.1016/S0140-9883(02)00062-2

Schalck, C., and Chenavaz, R. (2015) 'Oil commodity returns and macroeconomic factors: A time-varying approach'. *Research in International Business and Finance* 33, pp. 290-303. Available at: https://doi.org/10.1016/j.ribaf.2014.05.002

Schofield, N. C. (2007) *Commodity derivatives: Markets and applications*. West Sussex, UK: John Wiley & Sons

Zagaglia, P. (2010) 'Macroeconomic factors and oil futures prices: A data-rich model'. *Energy Economics* 32 (2), pp. 409-417. Available at: https://doi.org/10.1016/j.eneco.2009.11.003