Converging in Crisis:
The International Impact of Europe's Energy Crisis on Natural Gas Prices

Georgia Bradley

Presented to the Department of Applied Mathematics

in partial fulfillment of the requirements

for a Bachelor of Arts degree with Honors

Harvard College

Cambridge, Massachusetts

March 24, 2023

Acknowledgements

I would first like to thank my supervisor, Professor James H. Stock, for his guidance and invaluable feedback. I thank him for his direction towards this fascinating topic which I otherwise would not have discovered on my own, as well as for his direction within this field of econometrics.

I would also like to thank my seminar leader, Slavik Sheremirov, for his guidance in the research process and constant feedback. I would also like to thank Professor Matthew Zaragoza-Watkins for being an additional source of direction in the initial stages of this thesis, as well as for his assistance in data collection.

In addition, I am grateful for my family, whose support I am forever thankful for.

Abstract

In the span of seven years, the United States has switched from being a net importer of LNG to the largest LNG exporter in the world. This is a historic shift in energy markets that has received little attention in the literature despite the persistent and immense effects it has had. The effects of the ongoing energy crisis in Europe only add impetus to the study of energy markets, where the constriction of Russian energy exports has seen the historic entrance of the U.S. as a major energy provider to Europe. Accordingly, this paper seeks to analyse the effect of U.S. LNG exports on global market integration, in addition to the effect of the recent European energy crisis on this integration.

This papers thus employs bivariate linear and threshold cointegration techniques to analyse the degree of integration between key European & Asian natural gas prices and oil prices. Results confirm recent findings that European natural gas prices became more integrated during 2016 to 2020 with U.S. natural gas prices whilst oil prices decoupled from U.S. gas prices. However, this paper also finds that when extending the time frame to the present day, these effects appear to reverse. Long-term equilibrium relationships weaken between foreign natural gas prices and U.S. gas prices, while such relationships between oil and U.S. natural gas prices appear stronger than before the U.S. began exporting LNG. Notably, short-term adjustments between U.S. and foreign natural gas prices increase in speed post 2020, indicating heightened price transmission and the potential emergence of a spot LNG market.

Contents

List of Figures

List of Tables

1 Introduction

The Russian invasion of Ukraine has caused an energy crisis without precedent in Europe. The drastic reduction in the availability of Russian natural gas has led to severe price volatility, supply uncertainty, and unprecedented highs in natural gas price benchmarks. Nowhere has this been more pronounced than in Europe itself, where natural gas benchmarks reached more than the equivalent of $600 USD a barrel in March of 2022.

At the same time, the shut-in of Russian natural gas production presents an unparalleled opportunity for the entrance of U.S. liquefied natural gas (LNG) suppliers. While the United States only began exporting LNG in February of 2016, it has since become one of the world's largest exporters of LNG, and became the largest exporter of LNG in the first half of 2022. Notably, Europe and the United Kingdom accounted for 64% of U.S. LNG demand during this latter time period, driven by high price differentials relative to other hubs as seen in Figure 1 [1]. With increasing LNG export capacity and relative contract flexibility compared to other LNG suppliers, US producers were more easily able to divert LNG cargo as to engage in the price arbitrage opportunities created by the supply shock from Russia. The constriction of Russian natural gas imports to Europe has also incentivised Europe to re-activate dormant regasification and import terminal projects. Europe's LNG import capacity is expected to expand by 34% between 2021 and 2024, with 25 new floating storage and regasifcation units (FSRUs)[2] to be installed across the EU in the coming years[3]. This rush to expand LNG importing capabilities presents long term opportunities for U.S. LNG contractors, whom, within 24 hours of the Russian invasion of Ukraine, lobbied President Biden to resume fossil fuel leasing on U.S. federal lands, expedite U.S. LNG export licenses, authorise new U.S. gas infrastructure, and approve $300 million in public funding to build infrastructure in Europe to support LNG import terminals and gas pipelines[4]. Such a conjunction of factors raises the question as to whether the European energy crisis has accelerated LNG market integration, and the ramifications such integration would

[1] Source: EIA (2022)

[2] FSRU: Long-haul LNG carriers whose vaporisation and storage capabilities are on the ship; effectively a floating 'import terminal'

[3] Source: EIA (2022), and S&P Global (2022)

[4] Source: Milman, O, (2022) "How the gas industry capitalized on the Ukraine war to change Biden policy", The Guardian

have on the domestic U.S. energy market.

Historically, contractual terms and fuel switching between natural gas and residual fuel oil in the North American electricity generation market have tied the price of natural gas to petroleum products (Hartley et al., 2008). Further, direct arbitrage opportunities between international gas markets have been limited in the past by liquefaction fees, international logistical, delivery, and legal constraints, regasification costs, as well as the lack of or non-existence of exporting and importing LNG terminals (Stauffer, 1995; Jensen, 2003). Early research thus concludes that while international natural gas prices are regionally integrated, evidence of global integration should be attributed to a common co-movement with oil prices (Siliverstovs et al., 2005; Li et al., 2014; Brown & Yücel, 2009; Neumann, 2009). However, events such as the boom in hydraulic fracking in the US after 2005, the Global Financial Crisis, and the emergence of swing suppliers such as Qatar and the USA have since changed the market dynamics of oil and natural gas. Consequently, recent research examining time periods extending into the late 2010's points toward an increase in the interdependence of North American and Asian & European markets, in addition to evidence of a decoupling with oil prices (Chiappini et al., 2019; Lalenti, 2021). To the best of my knowledge, Lalenti (2021) is the only paper that specifically examines the effect of increased US LNG exports on global market integration. However, it is restricted to the time period prior to 2020 and does not include the recent perturbations to the natural gas market.

As such, the motivation of this paper is twofold: 1) to re-assess how the emergence of the United States as both a flexible and major supplier of LNG in the world market has changed the global integration of natural gas prices using the latest available price data, and 2) provide a novel examination of how the recent energy crisis has affected the degree and pace of any integration present. In tandem with this, this paper also analyses the degree of integration of U.S. gas prices with major oil prices to examine how oil-indexation has changed over tine.

Understanding how relationships between international natural gas benchmarks have changed over the short-term and long-term has many important corollaries. The diversion of natural gas to feedgas[5] deliveries to LNG export terminals as companies seek to profit off of price

[5]Feedgas: natural gas delivered via pipeline to a liquefaction facility for the singular purpose of being converted

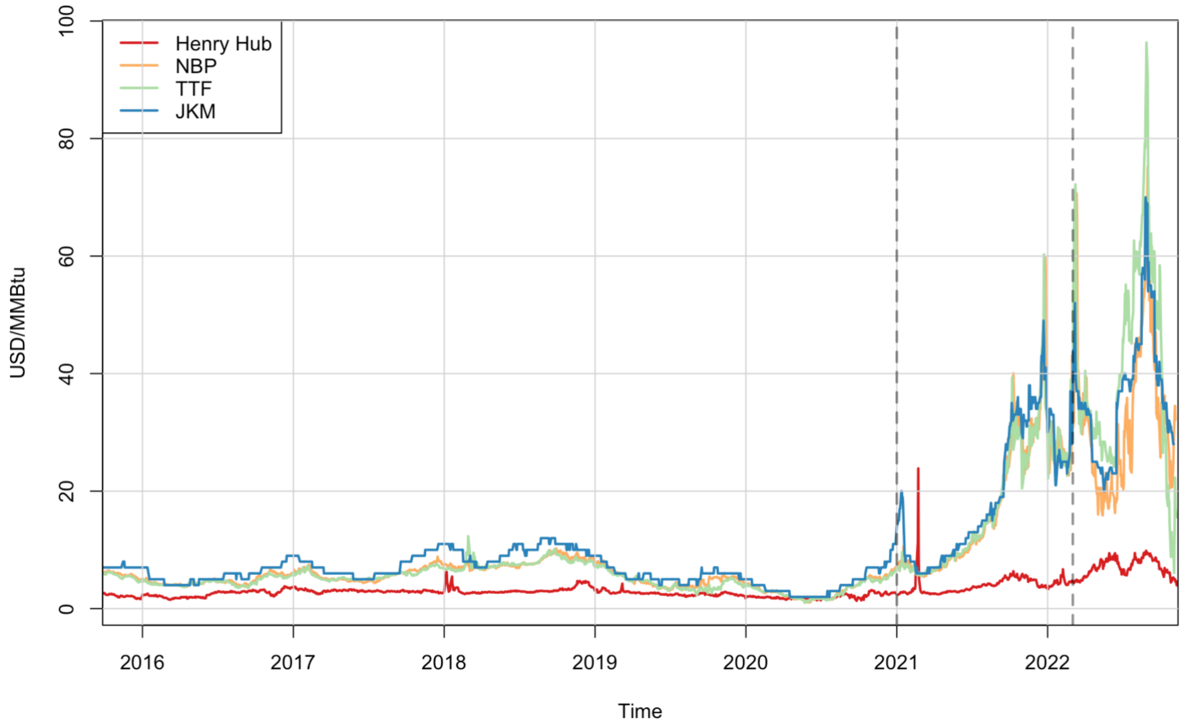
differentials may limit the availability of domestic gas and induce price increases. This affects not only the electricity generation market in the United States, whose largest power source comes from the combustion of natural gas, but also industrial, residential, & commercial sectors who rely on natural gas as a source of heating and an important feedstock in the production of chemicals, fertilizer, and hydrogen. Further, the relative prices of energy can influence investment patterns: if natural gas prices experience a persistent increase, the transition to increasingly low-cost and market shock-robust renewables such as wind, hydro, and solar may be expedited. In addition, waning integration with oil prices may reduce the United State's exposure to oil shocks and the ability to price forecast based off of other energy commodities.

Accordingly, this paper tests for market integration through a combination of linear and threshold cointegration tests, which allow for the size of long-term and short-term equilibrium effects to be estimated. The data includes international natural gas benchmarks such as Henry Hub natural gas prices (the best proxy for U.S. natural gas prices) and key European and Asian benchmarks, such as the Dutch Title Transfer Facility (TTF), the British National Balancing Point (NBP), the Asian S&P Japan Korea Marker (JKM), and the World Bank's Japanese LNG import price. Also included are the European Brent Oil and West Texas Intermediate (WTI) prices, which the aforementioned gas prices have historically been pegged to. The non-stationary properties of each variable are first assessed using unit root tests, where properties such as structural breaks and non-stationary volatility are pre-tested for. For non-stationary variables, linear cointegration tests are employed on the full samples that allow one to five breaks in the cointegrating relationship, all of which are endogenosuly determined. Linear cointegration is also employed on exogenously determined sub-samples of the data, pertaining to the periods of pre-2016, post-2016, 2016 to 2020, and 2020 onwards. These periods pertain to before and after the United States begin first exporting LNG (pre and post February of 2016), as well as before and after the initial constriction of Russian natural gas to Europe beginning at the start of 2020[6]. The long-term and short-term equilibrium effects are then estimated through dynamic ordinary least squares (DOLS) and linear and threshold error correction models.

to LNG

[6]Source: EIA

Figure 1: International Natural Gas Prices[1]



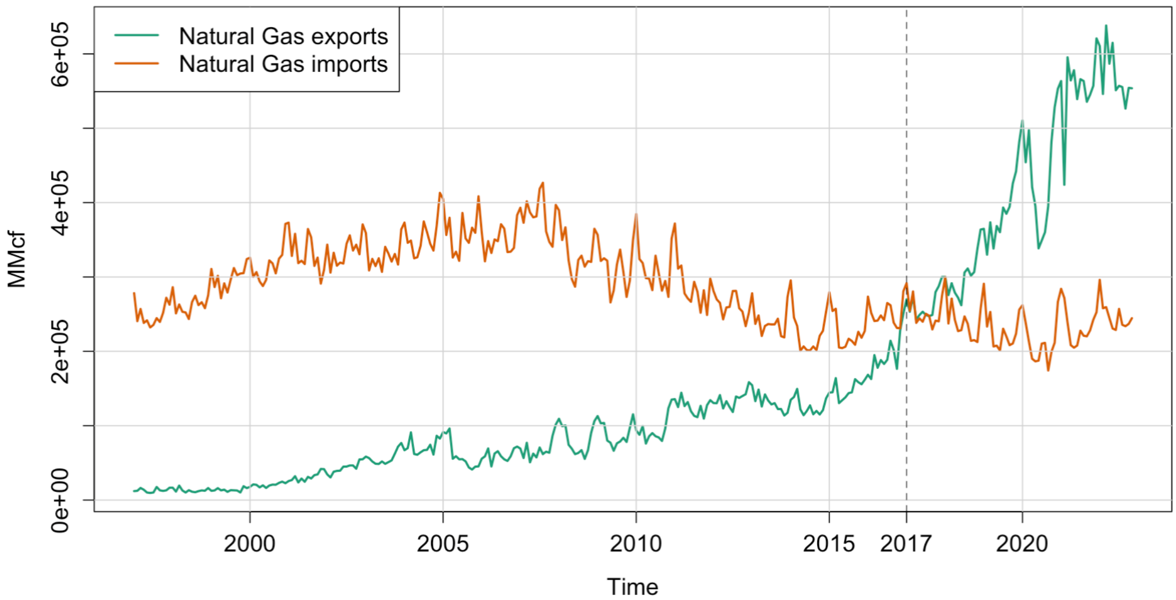
[1]First dashed line is the start of 2021; second dashed line indicates the date of the Russian invasion of Ukraine

This paper finds that European prices and U.S. prices became more integrated in the 2016-2020 period whilst U.S. natural gas prices simultaneously decoupled from oil prices, confirming recent findings. However, extending this analysis to current price data produces mixed results. In particular, the long-run equilibrium relationship between Asian and European prices and U.S. prices is weaker in the wake of the European energy crisis than in prior periods. Further, oil prices appear to have a stronger long-term relationship with U.S. gas prices after 2020 than prior to when the U.S. first began exporting LNG. Nonetheless, short run adjustment processes between Henry Hub prices and international natural gas prices have strengthened after 2020 compared to all other periods. Results thus indicate that while international gas prices do not resemble a law of one price, the speed at which prices are transmitted to foreign markets has increased substantially after the emergence of U.S. LNG exports and the European energy crisis.

1.1 Background

To establish a complete picture of the natural gas market, it serves to first understand the recent phenomena that is the rise of the U.S. LNG export market. In particular, the U.S. LNG market has seen three distinct phases: the period leading up to 2009, where the domestic LNG industry steadily grew through the expansion of horizontal drilling and hydraulic fracturing techniques but large imports of LNG were still needed to satisfy domestic demand; 2010–2016, where fracking heavily increased, low-cost gas drastically increased in supply, and industry committed to expanded domestic LNG production; and 2016 onwards, where as seen in Figure 2, LNG exports rapidly grew and overtook imports.

Figure 2: U.S.Natural Gas Exports vs. U.S. Natural gas Imports over Time



In this first period, natural gas prices in the United States were often linked to the price of oil benchmarks due to their close interchangeability in the electricity generation market. This is often referred to as burner tip parity, as it reflects that prices at gas trading hubs should adjust to residual fuel oil prices as to achieve parity where they are used – the 'burner tip' (Brown, 2005) This relationship was often characterised in a 1:10 ratio, where Henry Hub - denominated in U.S. dollars per million British thermal units - would adjust to be one tenth the price of crude oil, denominated in U.S. dollars per barrel. This relationship between oil and gas naturally extended (and extends) into the LNG market, where contracts are often oil-indexed. This is especially true of Japan, China, and South Korea, whose lack of a domestic natural gas markets means the majority of LNG contracts are indexed to the Japan Crude Cocktail (JCC).

The shale gas revolution of the late 2000's led to a drastic increase in the supply of U.S. natural gas, which has since been attributed to the liberalisation of the North American natural gas market. Particularly, the soaring prices of crude oil during the 2000s commodity boom and the low prices of Henry Hub led to nonviable substitutions between the two fuel sources. Occurring around the same time, the Fukushima nuclear disaster caused a sudden increase in Japanese demand for LNG. The accident not only led to the accelerated development of the LNG spot market as cargoes were suddenly diverted to the world's largest importer of LNG, Asia, but also to worldwide investment in LNG projects as countries diversified away from nuclear power (Hayashi & Hughes, 2013). Unlike Japan and South Korea, which are fully dependent on LNG imports, growth in LNG trade in Europe must compete with the extensive pipeline gas network connecting the European states to Russia. For this reason, Western European states such as the UK, France, Spain, Italy, Greece, and Portugal have a greater history of importing LNG. This confluence of expanded U.S. natural gas supply and the expansion of the global LNG market led to the initial investment in U.S. LNG exporting facilities.

The first liquefaction train[7] in the United States, built at the Sabine Pass terminal in Lousiana by Cheniere, opened in February of 2016. Prior to this, the United States was a net importer of natural gas, largely through pipeline from Canada. Currently, there are seven different operational export terminals in the United States, with three export projects currently under construction with expected finish dates of 2025. As of July 2022, the United States has more LNG export capacity than any other country, and the construction of these three additional projects are expected to further increase U.S. LNG peak export capacity by a combined 5.7 Bcf/dl[8][9] .Bloomberg expects the US to increase its total export capacity to 169 million metric tons (MMt) a year in 2027, which would represent more than a doubling from 2022 capacity[10]. This tremendous investment in and expansion of the LNG export industry places the United

[7] A train refers to the liquefaction and purification facility at an LNG plant, not transportation

[8] Source: EIA (2022)

[9] Source: EIA (2022)

[10] Source: Bloomberg (2022)

States in a prime position to capitalise on future international demand for natural gas.

Most U.S. LNG is currently sold under long-term agreements or contracts unattached to a spot price. This is due to the need to make final investment decisions (FID), which energy companies need in order to sanction and finance the future development of a project. A small percentage of the market is sold on the spot, normally Delivered Ex-Ship (DES) – valued as it is being unloaded at the receiving terminal – or Free-on-Board (FOB) – valued as it is being sold from the liquefaction facility to the owner of the vessel, whom will then go on to sell it overseas. Moreover, the cost of a new LNG carrier is expected to be $200 million to $250 million, with charter rates running at $80,000 - $100,000 per day to support these capital and operating costs (Mokhatab et al., 2014) Spot charter rates are approximately a third to a quarter of these rates, which makes short-term contracts less economically viable.

However, a growing number of long-term U.S. LNG contracts do not include destination clauses, especially relative to other LNG exporters. A destination clause pre-specifies and designates a list of unloading terminals as the destination ports of LNG vessels. Ultimately, the clause prevents buyers from reselling LNG to other terminals or buyers and severely limits contract flexibility. The absence of such a clause in US contracts allows buyers of U.S. LNG to sell cargo anywhere in the world, such that domestic and overseas spot deals can be conducted with U.S. gas purchased under long-term contracts. As such, U.S. LNG cargo may be re-sold multiple times before it is even removed from the carrier and regasified. For example, throughout the first half of 2022 Chinese companies resold surplus LNG cargoes with destination flexibility to Europe due to weak domestic demand for natural gas and high storage levels at home[11]. This allowed Asian traders and LNG companies to profit off of large price differentials. Evidence of this is seen in the decline of U.S. LNG export to Asia by 51% during the first four months of 2022 but the concurrent rise of LNG imports from the United States to Europe and the United Kingdom, which more than triple compared to 2021[12] .

Increases in U.S. LNG exports impact domestic Henry Hub prices in how they divert feedgas deliveries, reduce storage, and tighten demand. Indeed, through November of 2022, feedgas

[11] Source: S &P Global (2022

[12] Source: EIA (2022)

deliveries to the seven major U.S. LNG export facilities averaged 11.7 Bcf/d compared to 10.4 Bcf/d the prior year[13]. The effect of such an increase in feedgas deliveries can be seen in the reaction of gas prices to the Freeport LNG plant explosion: the explosion cut export capacity by approximately 15%, and pushed in 2 Bcf/d of feedgas demand back into the domestic market. The price of Henry Hub dropped by 12% when the plant was taken offline.

The strengthening of the U.S. export industry, its relative flexibility, and growing global demand for LNG present a unique confluence of factors that have the potential to drastically change the U.S. domestic natural gas market. In particular, this paper seeks to establish how recent events have changed the ability of foreign natural gas prices to influence domestic prices through the price levelling mechanism of the LNG market.

1.2 Literature Review

Early research on the integration between American and international markets before 2016 presents mixed results, with findings generally supporting regional over global integration. Where global integration is found, it is normally attributable to a common linkage to crude oil prices – this is especially true of U.S. markets which did not have LNG export terminals prior to 2016.

Early evidence of this behaviour comes from Siliverstovs et al. (2005), who utilise PCA and Johansen likelihood based cointegration tests to analyse the degree of natural gas market integration across American, European, and Japanese markets from the early 1990s to 2004. They find evidence of greater cointegration and price co-movement within the regional markets of Europe/Japan and America, but find limited integration between them. Similarly, Li et al. (2014) find evidence of convergence between Asian prices and UK prices from 1997 to 2011, but distinct behaviour from North American markets. In addition, they find the integration between European and Asian regions attributable to contractual mechanisms linking the price of natural gas to oil benchmarks as opposed to market forces of supply and demand.

In support of global integration, Nick and Tischler, 2014 find a cointegrating relationship during

[13] Source: S&P Global (2022)

2000 and 2008 between U.S. and British natural gas markets using a threshold error-correction model. While they find a decoupling between 2009 and 2012, they attribute this to evidence of high threshold estimates in the period that indicate non-transportation related impediments to arbitrage. Brown and Yücel (2009) similarly analyse U.S. and UK natural gas prices during 1997 to 2008, finding bivariate causality between the two prices. They also find evidence of short-run and long-term effects between natural gas and oil prices. Thus, while they posit that gas price co-movement could be due to the diversion of LNG cargoes & free market forces, it is also possible that natural gas prices co-movements are facilitated by crude oil prices. Barnes and Bosworth (2015) analyse trade volumes, instead of price movements, to determine the extent to which LNG is a regional commodity. Using a gravity model, they find that trade in LNG from 1998 to 2011 has de-regionalised the total natural gas market. Neumann (2009) use the Kalman filter technique to identify price convergence between US and European prices in the period from 1999-2008. While similar to Nick and Tischler, 2014 they find evidence of transatlantic price convergence, they also find oil prices strongly influence the convergence process as noted by Li et al. (2014).

The shifting relationship between gas and oil prices prior to the beginning of U.S. LNG exports has too been analysed. Prior to 2006, gas prices tended to shift upwards to the price of oil, but from 2005 through to 2009 gas prices were predicted to be lower given crude oil prices (Ramberg & Parsons, 2012). Lower price dependency from 2006 through to 2014 was also noted by Batten et al. (2017), who attributed it to the result of demand and supply side shocks; these include Hurricane Katrina, the Tohoku earthquake, the Global Financial Crisis, and technological and infrastructure improvements, particularly the increased use of hydraulic fracking.

More recently, Lalenti (2021) finds greater co-movement between global benchmark prices for natural gas since the United States became a mass exporter of LNG in 2016. Further, they find that prices in Europe and Asia respond negatively to increased exports of U.S. LNG but do not yet adhere to a law of one price. Chiappini et al. (2019) employ linear and threshold cointegration techniques to find integration between North American and European regions has increased in recent years while integration between gas and oil prices have decreased. Further, they also note the presence of asymmetric convergence between American, Asian, and European prices could indicate market arbitrage strategies being employed by exporting countries with excess supplies. This finding is supported by Kim et al. (2020), who finds indirect evidence that the emergence of swing suppliers, such as Qatar and Russia, has helped increase market integration between Europe and Asia after 2009. Similarly, Barbe and Riker (2015) employ a counterfactual framework to estimate that the trade volume of LNG would more than double in the absence of technical, logistical, and legal limitations. Further, Zhang and Ji (2018) employ a long-memory approach to find that the U.S. gas market is significantly decoupled from oil prices, similar to Chiappini et al. (2019). However, they note that oil-indexation is still widely prevalent in the European and Japanese markets.

The idea of expanded LNG exports serving as a price levelling mechanism has not been unchallenged. Loureiro et al. (2022) employ a relative convergence test on Asian, Russian, American, and European natural gas benchmarks and find no evidence of gas price convergence outside Europe from 2010 to 2020. While they find evidence of partial convergence in previous sub-periods, they attribute such periods to oil-indexation and coincidental regional demand and supply conditions.

In this way, this paper extends the literature by re-assessing cointegration between U.S. gas prices, international gas prices, and oil prices using the latest price data. Further, it pays a particular attention to how foreign prices affect domestic U.S. prices and the speed at which such effects are transmitted. It also provides a novel examination of how the recent shock to energy markets induced by the energy crisis in Europe has affected, and potentially accelerated, its integration.

2 Data

To study the international natural gas market, the collected data comprises of major natural gas benchmarks and oil prices. In particular, data is gathered on the natural gas benchmarks for the United States, Europe, and Asia, as well as the two major price markers for oil purchases, Brent Crude and West Texas Intermediate (WTI) crude oil. All collected data sources are summarised in Table 1.

Table 1: Data Sources

	Units*	Frequency	Range	Source
Henry Hub	$/MMBtu	Daily	1997 – 2022	EIA
NBP	$/MMBtu	Daily	1997 – 2022	Refinitiv
TTF	$/MWh	Daily	2010 – 2022	Refinitiv
JKM	$/MMBtu	Monthly	2012 – 2022	Refinitiv
LNG, Japan	$/MMBtu	Daily	1977 – 2022	World Bank
EU Brent	$/Barrel*	Daily	1987 – 2022	EIA
WTI Crude	$/Barrel	Daily	1986 – 2022	EIA

$/mmBtu = U.S. dollars per million british thermal unit; Symbols refer to those referenced prices symbols on Refinitiv workspace; All $ represent USD; Price symbols from Refinitv, respectively as mentioned, are NGLNMc1, TRNLTTFD1, JKMc1

For the United States, spot and future natural gas prices set at Henry Hub are widely seen to be the U.S. benchmark for American natural gas. Henry Hub is a distribution hub of pipelines, located in Erath, Louisiana, that serves as the official delivery location for futures contracts on the New York Mercantile Exchange (to which it lends its name too as the pricing point). While there are many regional spot prices in the U.S.A, Henry Hub's strategic location and interconnectedness has made it the leading indicator of natural gas supply and demand. The Energy Information Administration (EIA) collects daily data on Henry Hub Natural Gas Spot prices, denominated in U.S. Dollars per Million British Thermal units (MMBtu). This data ranges from 1997 to the present, and sets the time frame for which we collect all other data time frames.

For Europe, there are two major natural gas benchmarks - the Title Transfer Facility (denoted as TTF), and the National Balancing Point (NBP). The TTF, first operational in 2003, is a virtual trading point for natural gas located in the Netherlands. The NBP, first operational in 1996, is similarly a virtual trading point for natural gas, though for the purposes of the trade of UK natural gas (Heather, 2010). The TTF is Europe's most liquid natural gas hub, having overtaken the NBP as the most most liquid (and important) gas trading point in Europe in 2016 due to its closer location to markets and the rise of LNG trade throughout the 2010's (Heather, 2019). Regardless of this difference in liquidity, both are important data sources in both their historic and current impact on European natural gas prices. In addition, the World Bank publishes a European Natural gas index in its pink book; however, it is just a composite index of the TTF and NBP and as such is not used.

For Asia, the major benchmark is the LNG Japan/Korea Marker generated by S&P Platts (denoted as JKM). The Platts JKM reflects the spot price index for LNG delivered ex-ship (DES)[14] into Japan, South Korea, China, and Taiwan. Historically, deliveries to this Northeast Asian region have comprised the majority of international LNG import demand, making the JKM marker important within and outside of Asia for pricing LNG cargo. Further, prior to 2022 the Asian region served as the primary destination for U.S. LNG exports. The Japan/Korea marker was launched in February of 2009, and is reported in a daily frequency by S&P Global.

The World Bank also reports a monthly Japanese LNG import price (CIF)[15] in its pink book, which extends farther back to 1977. There is noticeable deviation between the World Bank monthly reported price and the monthly JKM marker. While these two time series somewhat track each other in trend, they differ significantly in magnitude. This could be attributed to the tracked metrics – the World Bank is purely tracking Japanese import prices at the border, while the S&P Platts indicator captures the entirety of the Asian region. Thus, while the World Bank marker is not a substitute for the JKM marker, it can be used to represent the Asian region for time spans that go further back than 2010.

Further, data is also collected on the European Brent Spot Price (FOB)[16], and on WTI Crude Oil prices at Cushing, Oklahoma. As aforementioned, natural gas and refined petroleum products

[14]DES: Deliverd Ex-Ship. The price of LNG at its final location, including shipping costs. The seller retains title and risk until the LNG is unloaded

[15]CIF: Cost, insurance, and freight. It is the price of LNG cargo as well as the price of insurance and transport to deliver to its final destination

[16]FOB: Free on Board: the price is calculated at the producing country's port of loading, and includes the price of the cargo plus the cost of loading

have historically been close substitutes in U.S. industrial and electric power generation, leading to significant price co-movement (Brown, 2005). While this price cointegration broke down in the late 2000's due to declining fuel switching capabilities and rising spot gas price indexation, oil indexation is still highly prevalent in LNG contract indexation and an important confounding variable (De Bock & José, 2011).

2.1 Data Transformation

All data manipulation, cleaning, and transformation is recorded in the public Git repository. The summary statistics of the transformed data are shown in Table 2.

Table 2: Summary statistics of transformed data

	Mean	Min	Max	Std. Dev.	N	Units	Start	End
Henry Hub	5.91	1.5	23.06	3.27	9425	$/MMBtu	01-Feb-1997	27-Nov-2022
TTF	11.22	1.12	96.34	9.73	4691	$/MMBtu	03-Jan-2010	09-Nov-2022
NBP	8.94	1.16	75.22	6.75	9404	$/MMBtu	01-Feb-1997	06-Nov-2022
JKM	12.31	2.26	70	10.27	3014	$/MMBtu	01-Aug-2014	02-Nov-2022
LNG, Japan	11.73	4.865	23.08	4.84	308	$/MMBtu	Feb-1997	Sep-2022
EU Brent	78.34	10.31	195.68	37.87	9431	$/Barrel	01-Feb-1997	03-Dec-2022
WTI Crude	75.71	10.07	197.53	33.69	9425	$/Barrel	01-Feb-1997	27-Nov-2022

N = Number of observations; LNG Japan is the only data source reported at a monthly frequency while all other sources are reported at the daily frequency; LNG, Japan is abbreviated to JLNG int the results

In examining the daily frequency time series, most exhibit significant outliers. Where outliers correspond to energy spikes arising from the interaction of extraordinary market and weather conditions, they are imputed with linear interpolation. This is because these outliers should not represent the true stochastic processes of the variable and can interfere with the power of unit root and cointegration tests (Franses & Haldrup, 1994). These imputed outliers and their justification are included in Appendix A. Further, surrounding values are sometimes also linearly interpolated around these spikes (especially in the case where the spike lingers for more than one day). Linear interpolation is also used to impute values missing on Saturday and Sunday. The JKM variable is the only variable with a sufficient number of missing data entries such that it would be spurious to impute it. As such, the range of JKM is narrowed down to August 2014 onwards.

In terms of the transformations of the time range of all these time series, WTI crude, EU Brent, NBP, and the World Bank Japan LNG index data have been subsetted to pertain to the period of February 1997 onwards. This is to align it to the start date of the first reported Henry Hub entry. Further, TTF is transformed from its original quote of USD per Megawatt hour (MWh) to USD/MMBtu using the appropriate conversion factor (≈ 3.412). In addition, all prices are realised to 2022 levels using annual level data on the CPI from the FRED. Notably, prices are not converted to their natural logarithms. As noted by Nick and Tischler (2014), converting to logarithms creates an isoelastic relationship between prices that does not accurately represent the economic intuition underlying price convergence or a law of one price. Accordingly, all interpretation of regression coefficients are as the effect of one unit changes.

2.2 Choice & Effect of Frequency

Previous analysis in this field has often been done at the monthly and annual frequency, with few studies focusing on higher frequencies. In the past, such sampling choices have not been thought to significantly affect the power of unit root tests; indeed, it has been well established that the power of unit root tests depends much more on time span than frequency (Perron, 1989b; Campbell & Perron, 1991). Similarly, Pierse and Snell (1995) find that temporal aggregation does not affect the asymptotic local power of unit root tests – only small increases in the time span are needed to compensate for a lower sampling frequency. However, in the absence of being able to increase the time span of the variables under study, some authors have found evidence that higher frequencies can in fact compensate for smaller time spans. Choi (1992) note that in finite samples that increasing the frequency from annual to monthly or quarterly may increase the power of PP and ADF tests. Similarly, In Choi and Bhum Suk Chung (1995) note that the finite sample powers of the ADF are increased when using a high sampling frequency; specifically, there is a significant gain in empirical power by using monthly data when ρ is below 0.99 (where $\rho = \delta + 1$ as in 1 and $\rho = 1$ represents the presence of a unit root).

Cointegration tests have also been subject to similar debate. Hakkio and Rush (1991) emphasise that since cointegration is a long-run property of data, that the sample size to be tested on must too be sufficiently long. They note that adding data by increasing the frequency does not yield more observations on the long-run fluctuations – the best way to increase the power of cointegration tests is to increase the time span of the sample. With regards to what constitutes a long-run equilibrium, they note that this depends on the relativity of the length of the long-run equilibrium to the total length of the sample period. Zhou (2001) confirms Hakkio and Rush (1991)'s statement that increasing sample time spans is more powerful than increasing observations; however, they also note that when studies are restricted to short time spans, power loss can be compensated by increasing data frequency. In particular, they note for shorter spans of 30–50 years, increasing the frequency from annual to monthly resulted in the same gain in power as doubling the time span of the data.

The period in which our data range - the lowest being 8 years for JKM and the highest being 25 for Henry Hub, WTI, and Brent - are relatively short sample periods in the context of cointegration and unit root tests. Even if the long-run equilibrium between variables is relatively short compared to the span of the total sample period, it would still serve to increase the power of our tests and reduce size distortion by increasing the frequency of the analysed data. As such, all tests are conducted on data observed at the daily level, despite the data from the World Bank only being available at a monthly frequency. For comparison with the World Bank data, the transformed Henry Hub price data is aggregated to the monthly frequency.

3 Methodology

To identify cointegration and bivariate regime shifts in international natural gas benchmarks multiple econometric techniques are employed. These include unit root and structural break tests to identify the order of integration of each variable; linear cointegration tests with and without structural breaks; and the estimation of long-term and short-term effects using regression and error-correction-models (ECM).

3.1 Unit root and structural break testing

Before being able to employ cointegration techniques, the order of integration of each time series must be known. The order of integration of a time series refers to the minimum amount of times a time series must be differenced before it becomes covariance-stationary such as to longer exhibit time-dependent structure. Order of integration is denoted $I(d)$ throughout this paper, where $I(1)$ denotes a time series that requires exactly one differencing before it is stationary.

A widely applied unit-root test is that of the Augmented Dickey Fuller test (ADF) (Dickey & Fuller, 1979) The ADF test compares the null hypothesis H_0 that a unit root is present in the time series against the alternative H_A that the time series is stationary, assuming that the underlying data is well modelled by an auto-regressive moving average (ARMA) structure. The ADF test as applied in this study is modelled in 1:

$$\Delta p_t = \alpha D + \delta p_{t-1} + \sum_{i=1}^{k} \theta_i \Delta p_{t-i} + \varepsilon_t \tag{1}$$

Throughout this paper, we denote Δ as the first difference operator and p_t as the natural logarithm of a price series (where hh_t is a special case singularly specifying the Henry Hub price series). D is the deterministic matrix which, for the three potential linear models of the ADF, respectively contains an intercept term, a linear trend term, and then both. The inclusion of a linear trend term accounts for time dependent features in the series while an intercept term accounts for the presence of drift.

However, the presence of structural breaks is known to have adverse behavior on standard unit root tests, including the ADF. Specifically, the ADF test is known to be biased towards the non-rejection of the null hypothesis of a unit root (Perron, 1989a; Ketenci, 2016; J. Lee et al., 1997). As C.-C. Lee and Lee (2009) and other authors posit, energy prices are subject to significant structural shifts due to the scope of macroeconomic, technological, and political shocks they are exposed to, in addition to their long time span. Structural shifts refer to when time series exhibit persistently different behaviours at different periods in time, such as how real GDP changes around the business cycle, shifts in inflation, and the interest rate, or how international trade temporally adjusts to exogenous shocks such as wars and global recessions. It can formally be expressed as an idiosyncratic change in the parameters of a time series model, such as the mean or variance.

As such, tests for structural breaks are tested for via the framework established in Bai and Perron (1998). Their methodology includes utilising a dynamic algorithm that computes the triangular Residual Sum of Squares (RSS) matrix of the linear model 2, which then seeks to extract an optimal segmentation of the form:

$$\begin{aligned} p_t &= 1_t^T \beta^{(j)} + \varepsilon_t \\ t &= n_{j-1}+1,\ldots,n_j \\ j &= 1,\ldots,m+1 \end{aligned} \tag{2}$$

Where $j = 1,\ldots,m$ is the segment index, $B^{(j)} \in \mathbb{R}^{m\times 1}$ is the segment-specific set of regression coefficients, and $\{n_1 \ldots, n_m\}$ is the set of unknown breakpoints. Given this setup, breakpoints are calculated using a dynamic programming algorithm based on the Bellman principle, where m breakpoints are chosen as to minimize the RSS of the model with $m+1$ segments.

The linear model we ingest into the algorithm is a regression of $p_t \sim 1_N^T$. This assumes the variables p_t are generated by the process:

$$p_t = \begin{cases} \mu_1 + \varepsilon_t & ,0 < t < \tau_1 \\ \mu_2 + \varepsilon_t & ,\tau_1 < t < \tau_2 \\ \vdots & \\ \mu_k + \varepsilon_t & ,\tau_{k-1} < t < n \end{cases}$$

where τ represents the breakpoints in each series. As in Bai and Perron (2003), this papers lets serial correlation be accounted through the inclusion of non-parametric adjustments as opposed to the inclusion of lagged regressors.

The decision of selecting breakpoints – their number and position – is made according to two main criteria: the Bayesian Information Criteria (BIC) score and the sequential SupF test. The BIC works well when the underlying variable has at least one break; however, the BIC accounts for potential heterogeneity across different periods and will choose a much higher number of breakpoints than the true amount in the presence of such serial correlation in the errors. As such, Bai and Perron (2003) recommend a sequential examination of the SupF$(l+1|l)$ statistics if there is at least one break in the underlying variable. The SupF$(l+1|l)$ statistics test the alternate hypothesis of $l+1$ breaks against a null of l breaks by using estimates of the break dates obtained from a global minimization of the sum of squared residuals. Bai and Perron (2003) also suggest that if a UDmax test reports at least one structural break, then the number of breakpoints (m) should be chosen according to when the SupF$(l+1|l)$ statistic is insignificant for $l \geq m$ even if SupF$(1|0)$ is insignificant. As such, a UDmax test, which tests the null of no structural breaks against the alternative of one structural break, is also utilised. Further, for the purposes of cross-validation, the number of breakpoints chosen by the BIC is compared against the results of the sequential SupF tests.

In addition, the presence of non stationary volatility in each price series is also tested. While unit roots often allow for the presence of breaks in the mean or intercept, they do not often allow for breaks or time-varying behaviour in the unconditional volatility. This is an important consideration, since it has been demonstrated that conventional unit root tests suffer large size distortions in the presence of non-stationary unconditional volatility (Busetti & Taylor, 2003; Cavaliere & Taylor, 2006). Further, Ghoshray (2018) provides an empirical review of unit-root testing in energy prices and notes that energy prices can be misspecified as stationary when excluding the consideration of structural breaks and non-stationary volatility. Accordingly, breaks in the variance of each time series are tested using the cumulative sum of square methodology developed by Inclan and Tiao (1994), hereon denoted as the IT test. However, Sanso et al. (2003) note that the IT test has large size distortions when the data is generated by a heteroskedastic conditional variance process or possesses leptokurtic and playkurtic innovations. As such, Sanso et al. (2003) propose a modified version of the IT test, hereon denoted the SAC test, that allows for conditional heteroscedasticity in addition to allowing for the consideration of fourth order moment processes. Accordingly, both the IT and SAC test are employed; of the latter, only the κ^2 test which corrects for non-mesokurtosis and persistent conditional variance is employed due to its lower computational complexity and greater power.

As such, in addition to the ADF, this paper utilises a novel unit root test developed by Cavaliere et al. (2011), denoted CHLT, that permits a possible break in the trend in addition to nonstationarity volatility. The CHLT test is an extension of the the ADF based test developed by Harris et al. (2009) and the corresponding $\mathscr{M}$-type tests; the CHLT improves upon the test by utilising a wild bootstrap-based implementation which allows for nonstationary volatility. To implement the procedure developed by the CHLT test, the following model is adopted:

$$
\begin{aligned}
p_t &= \alpha + \beta t + \gamma DT(\tau) + u_t \qquad (3) \\
u_t &= \rho_T u_{t-1} + \varepsilon_t \\
\varepsilon_t &= C(L)e_t = \sum_{j=0}^{\infty} c_j e_{t-j}
\end{aligned}
$$

where $t = 1, \ldots T$, $e_t = \sigma_t z_t$, and $z_t \sim IID(0,1)$. τ denotes the unknown break fraction, such that the break point is given by τT and the magnitude of the trend break is given by γ. CHLT ingests this linear model 3 and then computes the three $\mathscr{M}$ statistics, MZ_a, MSB and MZ_t.

3.2 Cointegration Tests

Once having established the order of integration of each respective time series, cointegration tests are conducted. As characterised by Engle and Granger, 1987, cointegration tests examine whether a a linear combination of $I(1)$ variables have a stationary distribution. Cointegration aligns with equilibrium theories that posit if two variables are in a long run equilibrium relationship, deviations from that equilibrium will be temporaneous and adjusted for quickly. The Engle-Granger test is performed by fitting the linear regression $hh_t = \mu + \beta p_t + u_t$ (4) and then testing u_t for stationarity using the ADF; it may be modified by including a trend term to $hh_t = \mu + \alpha t + \beta p_t + u_t$ (5). The Engle-Granger test is carried out on both the entire sample as well as the sub-samples of pre–2016, post–2016, 2016–2020, and 2020 onwards to test for linear cointegration in these periods.

Standard cointegration tests normally assume that the cointegration vector is time invariant under the alternative hypothesis (Campos et al., 1996; Noriega & Ventosa-Santaulària, 2012). However, there may exist some common stochastic trend between two variables that exists for some segment of time, which is then exposed to a shock that induces an abrupt and persistent change in its behaviour. We hypothesise that the geopolitical effects of the shut-in of Russian Gas has induced a severe enough behavioural change in those actors participating in the international natural gas markets that a structural shift has recently occurred in any equilibrating relationship between them. As such, for the full-sample data, tests that account for such breaks in the cointegrating vector are employed.

Modifying the model proposed by Engle and Granger, A. Gregory and Hansen (1996) propose a test allowing for a single structural break in the cointegrating vector. They extend the ADF, Z_a, and Z_t tests for cointegration by allowing a regime shift in either the intercept or the entire coefficient vector. Gregory and Hansen (1996a) originally introduce three linear models to expand upon the standard model of cointegration – one with a level shift, a second with a level shift with trend, and a third with a regime shift. In A. W. Gregory and Hansen (1996), they introduce a fourth model that permits a regime and a trend shift in addition to changes in the intercept and slope coefficients . We use this later model with regime and trend shifts (7) in addition to the model with just a regime shift (6):

$$hh_t = \mu_0 + \mu_1 \phi_{t\tau} + \beta_0^T p_t + \beta_1^T p_t \phi_{t\tau} + u_t \tag{6}$$

$$hh_t = \mu_0 + \mu_1 \phi_{t\tau} + \alpha_0 t + \alpha_1 t \phi_{t\tau} + \beta_0^T p_t + \beta_1^T p_t \phi_{t\tau} + u_t \tag{7}$$

where the structural change is modelled by the dummy variable ϕ:

$$\phi_{t\tau} = \begin{cases} 0, & t \leq [n\tau] \\ 1, & t > [n\tau] \end{cases} \tag{8}$$

and τ denotes the relative timing of the regime change point. Subscripts of 0 and 1 respectively represent their values in the period before and after the regime shift.

The Gregory and Hansen (1996b) and Gregory and Hansen (1996a) tests only allows for one endogenous break. Given the time span of our time series, as well as the multitude of macroeconomic, geopolitical, and institutional changes the data is subjected to, it is plausible that any cointegrating vectors have been exposed to multiple structural breaks. Two recent tests have been created that allow for a greater number of breaks in the cointegration relationship. Generalising the Engle-Granger model, Hatemi-J (2008) (2008) proposes a cointegration test with two unknown regime shifts, while Maki (2012) proposes a cointegration test allowing for up to five unknown breaks. Like the aforementioned tests, both the Maki and Hatemi-J tests are constructed from residual-based models; however, the Maki test not only takes into account more breaks but is significantly less computationally expensive. Maki considers four linear models to which cointegration testing may be applied:

$$hh_t = \mu_0 + \sum_{i=1}^{m} \mu_{1i}\phi_{it} + \beta_0^T p_t + u_t \tag{9}$$

$$hh_t = \mu_0 + \sum_{i=1}^{m} \mu_{1i}\phi_{it} + \beta_0^T p_t + \sum_{i=1}^{m} \beta_{1i}^T p_t \phi_{it} + u_t \tag{10}$$

$$hh_t = \mu_0 + \sum_{i=1}^{m} \mu_{1i}\phi_{it} + \alpha_0 t + \beta_0^T p_t + \sum_{i=1}^{m} \beta_{1i}^T p_t \phi_{it} + u_t \tag{11}$$

$$hh_t = \mu_0 + \sum_{i=1}^{m} \mu_{1i}\phi_{it} + \alpha_0 t + \sum_{i=1}^{m} \alpha_{1i} t \phi_{i,t} + \beta_0^T p_t + \sum_{i=1}^{m} \beta_{1i}^T p_t \phi_{it} + u_t \tag{12}$$

where m represents the number of breaks. Each model allows for different features of a time series to be tested: Equation 9 allows for a model with level shifts; Equation 10 (denoted the regime-shifts model) allows for structural breaks in β in addition to μ; Equation 11 allows for a linear trend αt in the time series; and Equation 12 allows structural breaks in the levels μ, trends αt, and regressors. The null hypothesis of the cointegration test introduced in the Maki tests is that of no cointegration, while the alternative hypothesis is cointegration with i breaks (where $i \leq m$). Maki provides evidence that the Gregory and Hansen test is equivalent to the Maki test when $m = 1$, and equivalent to the Hatemi-J test when $m = 2$.

Of importance is the trimming parameter η, which concerns the percentage of observations to trim from the sample size when finding a breakpoint. This means that the first break point, b_1 is estimated to be within $[\eta T, (1-\eta)T]$, where T is the sample size, the second break point is estimated to be within $\{[\eta T, b_1 - \eta] \cup [b_1 + \eta, (1-\eta)T]\}$, and so on. This is an important consideration to make, since Buberkoku, 2017 finds that Maki critical values may be sensitive to the choice of trimming value. In particular, they find a 5% trimming parameter obtains more similar results to the Gregory and Hansen and Hatemi-J tests than the default 10% parameter. The choice of 5% is also used in Maki's paper. However, reducing the trimming parameter severely reduces the maximum time interval a segment may span. This leads to scenarios where the cointegrating vector may have multiple breaks in the span of a year, which not only limits the number of observations in each segment but does align with economic-intuition about the nature of cointegrating vectors. Further, since the trimming rate for the Gregory Hansen and Hatemi-J tests cannot be changed from 10%, maintaining a trimming rate of 10% for the

Maki test allows for better internal comparison between tests.

3.3 Long Term Estimates: DOLS

The cointegrating vector can then be estimated between cointegrated $I(1)$ variables using the dynamic OLS (DOLS) regressions proposed by (Stock & Watson, 1993). They develop an asymptotically efficient estimator for the normalised cointegrating vector $\beta = [1, -\beta_2^T]^T$, as well as a formula for computing its asymptotic variance. By including leads and lags of the first difference of the regressor and utilising a generalised least squares (GLS) procedure, both regressor endogeneity and serial correlation in the error term can be corrected for. The DOLS model can be represented in our bivariate case as:

$$hh_t = \delta' D + \beta_2 p_t + \sum_{j=-n_2}^{n_1} \phi_j' \Delta p_{t+j} + \varepsilon_t \tag{13}$$

Where D is the deterministic matrix which includes an intercept term for the purposes of our analysis, while n_1 and n_2 are the respective maximum number of lead and lag terms. The number of leads and lags is dynamically chosen according to the methodology of Choi and Kurozumi (2012). In particular, we allows the number of leads and lags for any model to differ, compared to having a fixed selection rule where their number must be equal. Choi and Kurozumi (2012) find that removing the restriction of having an equalised number of leads and lags leads to a lower MSE in the long-run coefficients.

3.4 Threshold Cointegration Tests

The aforementioned cointegration models assume a linear cointegrating vector – that is, deviations away from the long run equilibrium are corrected for constantly and instantaneously in each period. However, as Balke and Fomby (1997) note, this is not necessarily representative of real world phenomena where adjustment to a long run equilibrium may occur discontinuously. In particular, the presence of transaction costs may prevent actors from engaging in the arbitrage opportunities that enable price realignment. Actors may only partake in those purchases that sufficiently deviate from the long run equilibrium relationship as to overcome their associated transaction costs. This is an important consideration for liquefied natural gas markets, where companies face significant fixed costs in the form of liquefaction, shipping, and regasification fees.

Further, linear cointegration techniques assume a symmetrical adjustment towards equilibrium – that is, deviation from the equilibrium is corrected in the same way regardless of direction. This is not necessarily representative of real world behaviour, where Stigler (2020) notes in a survey on threshold cointegration that there may be price rigidity, menu costs, or such market power present such that price increases are reflected more quickly than price decreases, or vice versa. Accordingly, there may be an asymmetric adjustment towards the long run equilibrium. The model introduced by Balke and Fomby allows for such a consideration asymmetric adjustment towards the long run equilibrium. While the work of Balke and Fomby (1997) largely focused on establishing the threshold long-run relationship representation, it has since been extended to a threshold VECM (TVECM) (Hansen & Seo, 2002; Lo & Zivot, 2001). Thus, this paper adopts a three regime threshold error correction model (ECM) for the modelling of threshold cointegration:

$$H_a : \begin{bmatrix} \Delta hh_t \\ \Delta p_t \end{bmatrix} = \begin{cases} \begin{bmatrix} \mu_1^{(1)} \\ \mu_2^{(1)} \end{bmatrix} + \begin{bmatrix} \alpha_1^{(1)} \\ \alpha_2^{(1)} \end{bmatrix} ECT_{t-1}^{(1)} + \sum_{i=1}^{k} \gamma_i^{(1)} \begin{bmatrix} \Delta hh_{t-i} \\ \Delta p_{t-i} \end{bmatrix} + \begin{bmatrix} \mu_{0,t} \\ \mu_{1,t} \end{bmatrix}, & ECT_{t-1} \leq \theta_L \\ \begin{bmatrix} \mu_1^{(2)} \\ \mu_2^{(2)} \end{bmatrix} + \begin{bmatrix} \alpha_1^{(2)} \\ \alpha_2^{(2)} \end{bmatrix} ECT_{t-1}^{(2)} + \sum_{i=1}^{k} \gamma_i^{(2)} \begin{bmatrix} \Delta hh_{t-i} \\ \Delta p_{t-i} \end{bmatrix} + \begin{bmatrix} \mu_{0,t} \\ \mu_{1,t} \end{bmatrix}, & \theta_L \leq ECT_{t-1} \leq \theta_U \\ \begin{bmatrix} \mu_1^{(3)} \\ \mu_2^{(3)} \end{bmatrix} + \begin{bmatrix} \alpha_1^{(3)} \\ \alpha_2^{(3)} \end{bmatrix} ECT_{t-1}^{(3)} + \sum_{i=1}^{k} \gamma_i^{(3)} \begin{bmatrix} \Delta hh_{t-i} \\ \Delta p_{t-i} \end{bmatrix} + \begin{bmatrix} \mu_{0,t} \\ \mu_{1,t} \end{bmatrix}, & ECT_{t-1} > \theta_U \end{cases} \tag{14}$$

The error-correction term, denoted by ECT_{t-1},s is split into three regimes under the TVECM. These regimes are dependent on the two threhold values θ_L and θ_U, which sort the observations of each price series into each regime according to a minimum trimming parameter η.

Balke and Famby (1997) suggest a two step approach for testing for threshold cointegration. While cointegration is tested as a global behaviour of a time series, threshold cointegration is a local behaviour that does not imply linear cointegration nor is implied by it. As such, for those time series that we reject the null of no linear cointegration, we can test for threshold cointegration against a null of linear cointegration. This can be done by utilising the methodology introduced by Hansen and Seo (2002). They propose a SupLM test of a threshold bivariate ECM with two regimes against a bivariate ECM, as shown in 15:

$$\text{SupLM} = \sup_{\theta_L \leq \theta \leq \theta_U} \text{LM}\left(\tilde{\beta}, \theta\right) \tag{15}$$

The LM statistic is a function of $\tilde{\beta}$, the null hypothesis estimate of the linear cointegrating vector, β, and the threshold value θ. The linear cointergating vector is taken from the results of the DOLS estimates. However, there is no point estimate of the threshold value under H_0. As such, the threshold value is chosen from the search region $[\theta_L, \theta_U]$ as to maximise the *LM* function. Since the LM function is non-differentiable, grid evaluation is used to determine the choice of θ. The LM statistics are computed with heteroskedasticity-consistent covariance matrix estimates.

In the absence of threshold effects, a linear ECM is used to model short-term effects. The linear model takes the form of any of the cases presented in 14.

4 Results

4.1 Unit Root and Structural Break Tests

Results for the ADF on full samples of all variables observed at daily and weekly frequencies are shown in Table 3, where weekly results have been shown as a point of comparision. ADF results for full samples observed at the monthly frequency are shown in Table B.1. In the following tables, certain variable names have been abbreviated, including Henry Hub to HH and Japanese LNG import prices to JLNG. Across the daily and weekly observations, only WTI, Brent Crude, and the JKM are definitively non-stationary at levels; when utilising a model with drift but no trend, the presence of a unit root is rejected at at least the 5% level for Henry Hub, TTF, and NBP data across daily and weekly observations. Conversely, at the monthly frequency, all variables are effectively non-stationary at levels (Table B.1). While we do not read into these results substantially given the aforementioned issues with the ADF test, a robust discussion of the properties of unit roots in natural gas prices is provided in the discussion to clarify these peculiar results.

Table 3: Daily & Weekly ADF Tests: Levels

	Daily			Weekly		
	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
HH	$-1.68(10)^{*}$	$-3.64(10)^{***}$	$-3.92(10)^{**}$	$-1.32(7)$	$-3.00(7)^{**}$	$-3.33(7)^{*}$
TTF	$-2.33(9)^{**}$	$-3.72(9)^{***}$	$-3.85(9)^{**}$	$-2.74(6)^{***}$	$-4.25(6)^{***}$	$-4.31(6)^{***}$
JKM	$-0.986(8)$	$-1.93(8)$	$-2.56(8)$	$-1.062(5)$	$-2.04(5)$	$-2.78(5)$
NBP	$-2.20(10)^{**}$	$-4.28(10)^{***}$	$-4.95(10)^{***}$	$-0.563(8)$	$-2.075(8)$	$-2.694(8)$
WTI	$-0.642(10)$	$-2.24(10)$	$-2.24(10)$	$-0.806(7)$	$-2.62^{*}(7)$	$-2.68(7)$
Brent	$-0.628(10)$	$-2.08(10)$	$-2.09(10)$	$-0.766(7)$	$-2.44(7)$	$-2.53(7)$

* Significant at the 10% level; ** Significant at the 5% level; ; *** Significant at the 1% level; () denote lag length. Model 1: No drift nor trend ; Model 2: With drift, no trend ; Model 3: With drift and trend

Structural break tests are reported in Table 4, which show the results for the UDMax test and the sequential SupF tests. The UDMax test examines the alternative of one structural break against a null of zero breaks, while the sequential SupF test examines the alternative of m breaks against $m-1$ breaks, sequentially. The number of breakpoints selected by the sequential SupF tests is reported in the second last column, $m^{[1]}$, while the number of breakpoints chosen through cross-examination with the BIC scores is reported in the last column, $m^{[2]}$. The results of the BIC scores are shown in Figure B.1. The BIC scores generally conform to the breakpoints chosen by the sequential SupF test, though a minimum is reached for the TTF data at two breakpoints instead of the three breakpoints chosen by the SupF test. Similarly, neither the BIC nor the RSS significantly decrease after three breakpoints for the Brent and WTI data, while the sequential tests choose four breakpoints for the Brent data. Bai and Perron, 2003 note the BIC is biased towards overestimating, as opposed to underestimating, the number of breakpoints. Given that the minimum of the BIC is reached at a lower breakpoint number than the SupF tests for the Brent Crude (at three instead of four) and TTF (at two instead of three) data, the BIC determined number of breakpoints is chosen for these two variables. From an economic standpoint, as oil prices Brent Crude and WTI should both be exposed to the same number of structural breaks; choosing the BIC determined breakpoints for Brent equalises the number of breakpoints for both oil prices. Similarly, it is unlikely TTF has been exposed to three structural breaks given its relatively shorter time frame of approximately 12 years. Further, cross-examination against the BIC provides evidence that the true number of breaks for NBP should be three; this choice aligns it with the TTF data, which closely tracks NBP but is of a smaller time span, insinuating it should have less than or equal to the number of breakpoints of NBP.

Table 4: UDMax Test & Sequential SupF Tests

	UDMax	Sup$F(1\|0)$	Sup$F(2\|1)$	Sup$F(3\|2)$	Sup$F(4\|3)$	Sup$F(5\|4)$	$m^{[1]}$	$m^{[2]}$
HH	30.536***	7.104	27.484***	6.500	6.500	0.520	2	2
NBP	146.917***	17.812***	0.340	25.606***	3.021	0.0	1	3
TTF	61.650***	5.643	33.268***	20.076***	4.501	0.588	3	2
JKM	83.104***	83.104***	3.110	15.539***	2.823	15.539**	1	1
Brent	385.919***	21.299***	15.082***	42.167***	12.128**	0.942	4	3
WTI	679.790***	17.220***	34.293***	20.978***	0.715	0.751	3	3
JLNG	163.173***	45.02**	23.579***	65.296***	17.258***	0.0	4	4

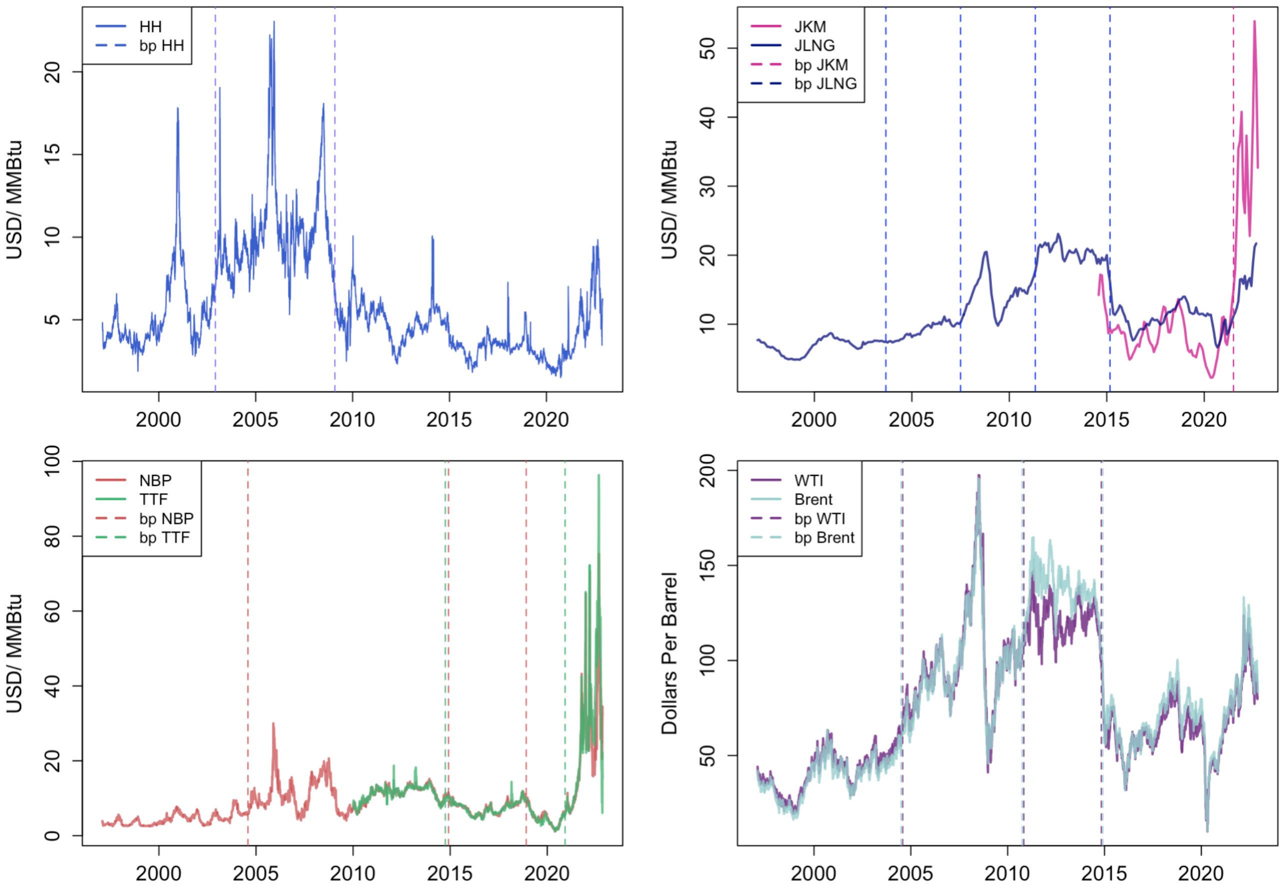
* Significant at the 10% level; ** Significant at the 5% level; *** Significant at the 1% level
$m^{[1]}$, the number of breakpoints, is chosen by the last significant SupF test
$m^{[2]}$, the number of breakpoints, chosen in combination with BIC scores

A visualisation of the dates chosen by the sequential SupF tests and the BIC scores is given in Figure 3. Graphs have been sorted by region, such that European prices are reported in the bottom left, Asian prices are reported in the top right, North American prices are in the top left, and oil prices are in the bottom right. The break-dates completely overlap for Brent and WTI, which is not unexpected given how closely the two prices track each other and their exposure to the same commodity market. TTF and NBP experience very similar breaks in the period for which they both co-exist: they experience overlapping breaks at the end of 2014 and similar breaks surrounding 2020. There are no similar structural breaks experienced in the Asian region: this could be attributable to how JLNG tracks the LNG import price for Japan while JKM tracks the entire Asian region, or that there is a relatively small time frame for which these two prices overlap. It could also be a result of the Japanese LNG import data being sampled at the monthly frequency: there may be too few observations to distinguish a regime shift from a random walk. Aligning with noted observations in the literature, Henry Hub and the two oil prices both experience breaks around 2009-2010 in the wake of the GFC.

Table 5 reports the test statistics for the CHLT test for unit roots, in addition to the results of the IT and SAC tests. The IT and SAC test reports the number of breaks in the variance in each time series. Both tests find evidence of nonstationary volatility in every variable under scrutiny. Accordingly, the CHLT test, which accounts for non-stationary volatility and a potential trend break, is utilised. For the three test statistics reported by the CHLT test, MZ_a, MSB and MZ_t, the null hypothesis of a unit root is rejected if the test statistic is higher than the bootstrapped critical values. Examining the table, it is found that all three tests fail to reject the null hypothesis of a unit root present for every variable at the 5% level. However, the MZ_a and MZ_t tests reject at the 10% level for Henry Hub. While this could be a type I error, as aforementioned a robust treatment of this issue is provided in the discussion. Nonetheless, it is concluded here that all variables under scrutiny are non-stationary at levels.

.

Figure 3: Structural Breaks in Regional Gas Markets & Oil Markets



The number and dates of structural breaks for each variable is determined from Table 4 in conjunction with the BIC scores. The dates for Henry Hub are: 11/2002, 01/2009; JKM: 06/2021; JLNG: 08/2003, 06/2007, 04/2011, 02/2015; TTF: 09/2014, 11/2020; NBP: 07/2004, 11/2014, 11/2018, WTI & Brent: 07/2004, 10/2010, 10/2014

Table 5: Results of IT, SAC, and CHLT test

			MZ_a			MSB			MZ_t		
	IT	SAC		Bootstrap c.v			Bootstrap c.v			Bootstrapped c.v	
		κ^2	Stat	5%	10%	Stat	5%	10%	Stat	5%	10%
HH	23	26	-18.857*	-21.421	-17.522	0.163	0.149	0.162	-3.070*	-3.238	-2.912
TTF	11	22	-24.756	-73.689	-60.838	0.142	0.081	0.089	-3.518	-5.984	-5.482
NBP	36	23	-43.844	-54.926	-44.527	0.104	0.095	0.103	-4.549	-5.143	-4.608
WTI	17	39	-14.273	-25.781	-20.042	0.187	0.136	0.155	-2.670	-3.536	-3.082
Brent	17	37	-11.178	-19.998	-15.420	0.210	0.157	0.176	-2.345	-3.094	-2.709
JKM	11	14	-25.155	-40.754	-31.927	0.141	0.108	0.123	-3.142	-4.654	-4.031
JLNG	2	1	-13.661	-25.556	-21.784	0.188	0.139	0.149	-2.563	-3.558	-3.277

IT refers to the Inclan and Tiao (1994) test, while SAC denotes the Sanso et al. (2003) test for nonstationarity volatility in the data. The SAC has two different tests, but only κ^2 is employed: κ^2 corrects for non-mesokurtosis and persistence in conditional variance. The CHLT test from Cavaliere et al. (2011)is run with 500 bootstraps, with tests run in gauss using code provided by The Granger Center for Econometrics. The CHLT test reports the three statistics, MZ_a, MSB and MZ_t, which are significant if they are higher than the boostrapped c.v.. All tests carried out on daily frequency par JLNG, which is only available at the monthly frequency.

4.2 Full Sample Results

4.2.1 Cointegration Results

Table 6 reports the findings for the Engle-Granger (EG) and Gregoy-Hansen (GH) cointegration tests. The EG test is reported as two models, neither of which allows for breaks in the cointegrating relationship: the first model allows for a constant (Model 1), while the second allows for a constant and trend (Model 2). The GH test allows for one break in the cointegrating relationship, and is modelled by two processes: one allowing for a regime shift (breaks in β, as in equation 6) and one allowing for regime and trend shifts (breaks in μ and β, as in equation 7). As can be seen, the null of no cointegration is rejected at the 1% level for all bivariate tests of TTF and JKM against Henry Hub for the EG tests, while it is rejected with at least 5% significance for those tests of WTI, Brent Crude, and NBP against Henry Hub. However, there is no evidence of cointegration between Japanese LNG prices and Henry Hub. This could be due to the sampled frequency, which at the monthly level produces a less powerful test. Conversely, the GH test reports cointegration at the 1% level for all variables under scrutiny when tested against Henry Hub, including Japanse LNG prices. These two tests provide very strong evidence of cointegration between international natural gas benchmarks and oil prices against Henry Hub, especially when including a break in the cointegrating vector.

Endogenously determined breakpoints from the GH test reported in Table 6 show breaks for the Japan/Korea Marker and Dutch prices occurring around the 2015-2016 mark, pertaining roughly to the date when the U.S. first began exporting LNG. Further, across NBP, Japanese LNG import prices, and both oil prices, breakpoints are found to occur 2008-2009, the time of the GFC.

Table 6: Engle-Granger cointegration test & Gregory Hansen cointegration test

	Engle-Granger		Gregory-Hansen	
vs. HH	Model 1	Model 2	Model 3	Model 4
NBP	-3.934^{**}	-5.412^{***}	-7.675^{***}	-8.253^{***}
Breakpoint			13-Sep-2008	01-Dec-2005
TTF	-5.483^{***}	-6.502^{***}	-6.729^{***}	-6.965^{***}
Breakpoint			08-Feb-2015	17-Oct-2012
JKM	-7.118^{***}	-7.265^{***}	-5.864^{***}	-6.048^{***}
Breakpoint			28-Dec-2017	24-May-2016
JLNG	-2.947	-3.261	-5.817^{***}	-5.924^{**}
Breakpoint			01-Dec-2007	01-Aug-2009
WTI	-3.908^{**}	-4.579^{***}	-6.659^{***}	-6.786^{***}
Breakpoint			26-Jul-2008	05-Feb-2009
Brent	-4.783^{**}	-4.326^{**}	-6.444^{***}	-6.595^{***}
Breakpoint			05-Sep-2008	05-Feb-2009

* Significant at the 10% level; ** Significant at the 5% level; ; *** Significant at the 1% level; Engle-Granger – Model 1: with constant, no trend ; Model 2: constant and trend; SIC lag selection (max of 8 lags) Gregory-Hansen – Model 3 incorporates a regime shift (6) , while Model 4 incorporates regime and trend shifts (7) ; Default trimming rate of $\eta = 0.10$; ADF statistic reported. Tests conducted using the tspdlib library in Gauss

Table 7 reports the findings of the Maki cointegration test, which tests the null of no cointegration against the alternative of cointegration with up to m breaks. Two models are reported: one allowing for regime shifts (as in equation 11) and one allowing for regime and trend shifts (as in equation 12). Table B.3 and Table B.4 respectively report the results of the Maki test when only up to three and up to four breaks are allowed. From Table 7, it can be seen that the null of no cointegration is rejected at the 1% level for Henry Hub and its bivariate relationships with NBP, the TTF, and WTI, and at least the 5% level for Brent. However, only when considering up to three breaks is the relationship between Henry Hub and JKM considered significant at the 5% level; further, this is only for a model allowing regime shifts. No number of breaks (from 2 to 5) finds a significant cointegrating relationship between Japanese LNG and Henry Hub prices.

Table 7 also reports the endogenously determined breakponts in the cointegrating relationship, when one is found. Across variables that span from 1997 onwards, there are noted breakpoints occurring around 2009 – around the GFC – while, notably, for all variables there are breaks occurring around the 2020 and 2021 mark. Breakpoints across Brent and WTI are generally consistent – that is, there are similar breaks around 2000, 2003/4, 2008, 2014, 2017/19 and 2020. The breakpoints chosen for NBP in model 3, past 2010, are also consistent with the chosen breakpoints for TTF – occurring around 2014 and 2020/2021. Accordingly, regional and market similarities seem well reflected in the choosing of breakpoints.

Table 7: Results of Maki Cointegration Test: $m = 5$

Vs. HH	Model	Statistic	Breakpoints
JKM	2	-5.895^{**}	04-Jun-2015, 20-Jan-2018, 19-Oct-2021
	3	-6.199	
JLNG [1]	2	-5.398	
	3	-6.211	
NBP	2	-8.760^{***}	30-May-2000, 30-Dec-2002, 02-Aug-2005, 28-Mar-2009, 14-Feb-2014
	3	-8.924^{***}	11-Dec-2000, 30-Aug-2005, 07-Mar-2009, 07-Feb-2014, 06-Apr-2020
TTF [2]	2	-7.706^{***}	06-Sep-2011, 27-Mar-2013, 18-Oct-2015, 8-Mar-2019, 24-Apr-2021
	3	-7.573^{**}	17-Sep-2011, 07-Feb-2014, 03-Jun-2016, 08-Mar-2019, 15-Feb-2021
Brent	2	-8.259^{***}	05-Jun-2000, 26-Feb-2003, 18-Nov-2005, 06-Dec-2008, 20-Jan-2018
	3	-7.422^{**}	28-Dec-2000, 06-Feb-2007, 09-Jan-2010, 07-Feb-2014, 14-Mar-2020
WTI	2	-7.906^{***}	27-Mar-2000, 5-Nov-2004, 23-Dec-2008, 14-Feb-2014, 6-Apr-2020
	3	-8.809^{***}	28-Dec-2000, 16-Nov-2004, 06-Dec-2008, 08-Aug-2017, 09-Mar-2020

*** Significant at the 1% level; ** Significant at the 5% level, * Significant at the 10% level. Model 0 = Level shift, Model 1 = Level shift with trend, Model 2 = regime Shift, Model 3= regime & trend shift. Up to five breaks allowed in the cointegrating vector. Trimming parameter $\eta = 0.10$. Tests conducted in Gauss (Nazlioglu, 2021); [1] Use $\eta = 0.11$ for model 3 as unable to compute with $\eta = 0.10$; [2] Use $\eta = 0.12$ for model 2 as unable to compute with $\eta = 0.10$

4.2.2 Long-term estimates

Using the cointegration results from Table 6, the full sample is split into two sub-periods according to the endogenously determined breaks. The second column, δ_0 reports the intercept value, while the third column reports the long-term coefficient, β_2, from the linear model 13. Strikingly, it can immediately be seen there is a decrease in the long-run effect of WTI and

Brent prices on Henry Hub prices after 2009. That is, prior to 2009, a one unit increase dollars per barrel of oil leads to an increase of ~ 0.077 \$/MMBtu in the price of Henry Hub, but post-2019 has nearly a quarter of the effect, leading only to an increase of ~ 0.0025 \$/MMBtu in the price of Henry Hub. Results for natural gas benchmarks also show a decrease in the long-term coefficient when moving from the first sub-sample to the second sub-sample, par TTF which moves from having a depressive effect on Henry Hub to a positive effect. Both NBP and Japanese LNG prices have high estimates in their first periods (1997 – 2005 and 1997 – 2009), with NBP nearly approximating a 1-1 relationship with Henry Hub. However, after these first periods, both long term estimates on Japanese LNG prices and NBP decrease more than six-fold.

In addition, using the cointegration results from Table 7, the full sample is split into 3 to 5 sub-periods according to the endogenously determined breaks chosen by the Maki test. Table 9 reports the long-term estimates for natural gas prices, while Table 10 reports the estimates for oil prices. The estimates for TTF tend to increase from 2010 to 2016, moving from -0.112 to 0.584; they then decrease onwards to 2019 before once again increasing, though these last estimates are relatively small. JKM appears to have a constant but weak relationship with Henry Hub between 2014 and 2021, which then decreases moving into the post-2021 perio. Just like in the Gregory-Hansen sub-samples, NBP appears to have a relatively strong relationship with Henry Hub prior to 2005, which then weakens in latter time periods. All three variables display small long-term estimates post 2021.

Table 10 finds a general downwards trend in the magnitude of the long-term estimates for WTI and Brent. Long term estimates up until 2007 for Brent and up to 2004 tend to revolve around 0.1, which aligns with the 10-1 burner tip parity rule between oil prices and Henry Hub. This relationship breaks down earlier for WTI, which shows decreased estimates beginning in 2004. Notably, both WTI and Brent experienced an increase in their coefficient estimates post 2020, with the estimate for WTI of 0.07 from 2020 onwwards being twice as high as the estimate from 2004 to 2008. Conversely, both Brent Crude and WTI experience their lowest estimates (in magnitude) in the segments containing 2016 to 2020, though the coefficient on WTI during this time is not significant. These are interesting results that imply a reversion in the long run relationship between oil prices and Henry Hub around the 2020 mark.

A visualisation of these DOLS estimates and the respective prices underlying them are shown in the Discussion section below, where Figure 9 shows the long term estimates for natural gas benchmarks, and Figure 10 shows the estimates for oil prices.

Table 8: DOLS: Long-term Estimates Based on Gregory Hansen Sub-samples

Vs. HH	δ_0	β_2	(n_1, n_2)	Segment
TTF	8.106***	−0.298**	(0, 7)	03-Jan-2010 – 17-Oct-2012
	(1.034)	(0.094)		
	2.837***	0.096***	(0, 9)	17-Oct-2012 – 09-Nov-2022
	(0.137)	(0.009)		
NBP	1.367*	1.072***	(9, 9)	01-Feb-1997 – 01-Dec-2005
	(0.56)	(0.102)		
	3.777***	0.148***	(0, 11)	01-Dec-2005 – 06-Nov-2022
	(0.479)	(0.037)		
JKM	1.141***	0.232***	(0, 0)	01-Aug-2014 – 24-May-2016
	(0.19)	(0.018)		
	2.33**	0.109***	(8, 0)	24-May-2016 – 02-Nov-2022
	(0.141)	(0.008)		
JLNG	1.508	0.714***	(4, 0)	Feb-1997 – Aug-2009
	(1.641)	(0.177)		
	1.887**	0.145***	(3,4)	Aug-2009 – Sep-2022
	(0.630)	(0.042)		
WTI	2.906***	0.078***	(0, 0)	01-Feb-1997 – 5-Feb-2009
	(0.52)	(0.007)		
	1.718***	0.029***	(0, 1)	23-Mar-2009 – 27-Nov-2022
	(0.35)	(0.004)		
Brent	3.197***	0.076***	(0, 0)	01-Feb-1997 – 23-Mar-2009
	(0.509)	(0.007)		
	2.177***	0.022***	(0, 1)	5-Feb-2009 – 27-Nov-2022
	(0.35)	(0.004)		

* Significant at the 10% level; ** Significant at the 5% level; ; *** Significant at the 1% level. Breakpoints selected according to results of the Gregory-Hansen ADF statistic (model 4); number of leads (n_1) and number of lags (n_2) chosen according to BIC; Newey-West standard errors; standard errors are given beneath coefficient values.

Table 9: DOLS: Maki subsamples of natural gas benchmarks

vs. HH	δ_0	β_2	(n_1,n_2)	Segment
TTF	6.868***	−0.112*	(1, 6)	03-Jan-2010 – 17-Sep-2011
	(0.534)	(0.051)		
	−0.608	0.383***	(6, 6)	17-Sep-2011– 07-Feb-2014
	(1.197)	(0.095)		
	−0.866	0.584***	(0, 6)	07-Feb-2014 – 03-Jun-2016
	(0.482)	(0.058)		
	3.319***	0.037	(1, 0)	03-Jun-2016 – 08-Mar-2019
	(0.309)	(0.04)		
	1.656***	0.214***	(0, 2)	08-Mar-2019 – 15-Feb-2021
	(0.132)	(0.029)		
	3.07***	0.08***	(0, 6)	15-Feb-2021 – 09-Nov-2022
	(0.41)	(0.012)		
JKM	1.963***	0.177***	(0, 0)	01-Aug-2014 – 04-Jun-2015
	(0.291)	(0.023)		
	2.092***	0.147***	(3,0)	04-Jun-2015– 20-Jan-2018
	(0.308)	(0.037)		
	1.842***	0.147***	(0,7)	20-Jan-2018 – 19-Oct-2021
	(0.118)	(0.012)		
	4.5***	0.047	(0, 0)	19-Oct-2021 – 02-Nov-2022
	(1.2)	(0.033)		
NBP	0.615	1.136***	(7, 9)	01-Feb-1997 – 11-Dec-2000
	(0.51)	(0.138)		
	3.878***	0.683***	(0, 8)	11-Dec-2000 – 30-Aug-2005
	(0.875)	(0.139)		
	6.663***	0.355***	(0, 2)	30-Aug-2005 – 07-Mar-2009
	(1.247)	(0.094)		
	6.001***	−0.106*	(0, 0)	07-Mar-2009 –07-Feb-2014
	(0.521)	(0.048)		
	1.139***	0.324***	(0, 8)	07-Feb-2014 – 06-Apr-2020
	(0.28)	(0.036)		
	2.421***	0.109***	(3, 0)	06-Apr-2020 – 06-Nov-2022
	(0.311)	(0.013)		

* Significant at the 10% level; ** Significant at the 5% level; ; *** Significant at the 1% level. Breakpoints selected according to results of Maki test; number of leads (n_1) and number of lags (n_2) chosen according to BIC; Newey-West standard errors; Bartlett kernel used.

Table 10: DOLS: Maki subsamples of oil benchmarks

vs. HH	δ_0	β_2	(n_1, n_2)	Segment
WTI	-0.02	0.13^{***}	(7, 7)	01-Feb-1997 – 28-Dec-2000
	(0.596)	(0.015)		
	0.196	0.145^{***}	(0, 7)	28-Dec-2000 – 16-Nov-2004
	(1.152)	(0.023)		
	7.694^{***}	0.034^{*}	(0, 0)	16-Nov-2004 – 206-Dec-2008
	(1.475)	(0.014)		
	2.737^{***}	0.018^{***}	(0, 9)	06-Dec-2008 – 08-Aug-2017
	(0.421)	(0.004)		
	2.47^{**}	0.012	(0, 0)	08-Aug-2017 – 09-Mar-2020
	(0.789)	(0.011)		
	-0.675	0.07^{***}	(0, 1)	09-Mar-2020 – 27-Nov-2022
	(0.506)	(0.007)		
Brent	0.52	0.125^{***}	(7, 7)	01-Feb-1997 – 28-Dec-2000
	(0.549)	(0.015)		
	2.919^{**}	0.097^{***}	(0, 8)	28-Dec-2000 – 06-Feb-2007
	(0.912)	(0.014)		
	0.645	0.078^{***}	(0, 1)	06-Feb-2007 – 09-Jan-2010,
	(1.092)	(0.01)		
	9.188^{***}	-0.033^{***}	(1, 3)	09-Jan-2010, – 07-Feb-2014
	(0.953)	(0.007)		
	1.234^{**}	0.03^{***}	(0, 8)	07-Feb-2014 – 14-Mar-2020
	(0.239)	(0.003)		
	-0.569	0.066^{***}	(0, 1)	14-Mar-2020 – 27-Nov-2022
	(0.502)	(0.006)		

* Significant at the 10% level; ** Significant at the 5% level; ; *** Significant at the 1% level. Breakpoints selected according to results of Maki test; number of leads (n_1) and number of lags (n_2) chosen according to BIC; Newey-West standard errors; Bartlett kernel used.

4.3 Sub-sample Results

4.3.1 Cointegration Results

Further, cointegration is also tested between natural gas prices in sub samples segmented as pre-2016 and post-2016, as well as from 2016-2020 and post-2020. Engle-Granger (EG) tests are reported in Table 11 for all four sub-periods, where Model 1 allows a constant but no trend, and Model 2 allows for a constant and trend. Considering all variables and models, the EG finds cointegration to be the strongest in the post-2016 period: estimates in Panel B most consistently reject the null of no cointegration at the 1% level during this time period, and every cointegrating relationship is at least significant at 10% level. This is a striking result when compared to the pre-2016 period reported in Panel A, which shows much weaker levels of cointegration. In this pre-2016 period, only the 2nd model for NBP rejects the null at the 1% significance level, while TTF shows no cointegration for the second model, Japanese LNG shows no cointegration, and Brent and WTI reject the null at just the 10% level. Decomposing this post-2016 period into 2016 to 2020 and 2020 onwards in Panel C and Panel D, interesting results are obtained. First, oil prices appear to only reject the null of cointegration post 2020 – only the second model of Brent rejects, and at the 10% level, the null of cointegration in the 2016-2020 period. This is an striking result considering oil prices are highly cointegrated over the entirety of the post-2016 period. Compared to oil prices, natural gas prices generally show higher significance levels in the 2016–2020 period compared to the the post-2020 period, including Japanese LNG prices which are only cointegrated with Henry Hub in the 2016-2020 period.

Table 11: Engle-Granger Cointegration Tests: Pre-2016, Post-2016, 2016-2020 & Post-2020

	Panel A: Pre-2016		Panel B: Post-2016		Panel C: 2016-2020		Panel D: Post-2020	
	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2
NBP	-3.683^{**}	-5.087^{***}	-6.414^{***}	-6.424^{***}	-4.423^{***}	-4.516^{***}	-3.590^{**}	-3.220
TTF	-3.447^{**}	-3.219	-6.830^{***}	-6.903^{***}	-4.282^{***}	-4.319^{**}	-4.298^{***}	-3.985^{**}
JKM	-3.241^{*}	-3.968^{**}	-6.526^{***}	-6.289^{***}	-3.773^{**}	-3.871^{**}	-3.807^{**}	-3.660^{*}
JLNG	-2.669	-2.707	-3.149^{*}	-3.111	-2.902	-4.2^{**}	-2.337	-3.091
Brent	-3.255^{*}	-3.597^{*}	-4.787^{***}	-4.889^{***}	-2.860	-3.698^{*}	-3.659^{**}	-3.750^{*}
WTI	-3.337^{*}	-3.897^{**}	-4.744^{***}	-4.815^{***}	-2.880	-3.048	-3.672^{**}	-3.798^{*}

* Significant at the 10% level; ** Significant at the 5% level; ; *** Significant at the 1% level; Model 1: with constant, no trend ; Model 2: constant and trend; SIC lag selection (max of 8 lags)

4.3.2 Long-term estimates

Table 12 reports the finding of the long-term estimates calculated for the pre-2016 and post-2016 samples, respectively contained in Panel A and Panel B. Moving from the long term estimates reported in Panel A to those reported in Panel B, it can be seen that NBP and JKM see reduced estimates, while WTI see increased estimates. Both the estimates for β_2 for TTF and Brent in the pre-2016 period are not significant, though in the post-2016 period their respective β_2 estimates are both very similar to their commodity groups: natural gas prices, par Japanese LNG prices, revolve around 0.1, while oil prices revolve around 0.06. Notably, none of the estimates for natural gas prices in the post-2016 period are particularly large - they each predict only a 0.1 unit increase in Henry Hub prices for every 1 unit increase in international natural gas market prices. Notably, Japanese LNG import prices appear to have the greatest effect on Henry Hub in this post-2016 period, with an estimate of 0.297. While this is still a relatively small estimate, it is threefold higher than the JKM, NBP, and TTF marker during this time.

Table 13 shows the results of the decomposed time period of post-2016, showing the time span of 2016-2020 in Panel A and of post-2020 in Panel B. Of notice is that reported estimates for Brent and WTI during the post-2020 period are very similar to estimates reported in Table 12, Panel B, which shows the full post 2016 sample. For the post 2020 periods, the β_2 coefficient is 0.067 and 0.072 for Brent and WTI, while for the post 2016 period it is 0.055 and 0.062, respectively. This pattern is also similar for NBP and TTF: their post-2020 estimates of 0.112 and 0.101 are very similar to their entire post-2016 estimates of 0.111 and 0.096, respectively. Further, it can be seen that estimates in the 2016-2020 time period for natural gas prices are much higher than in the post 2020 period. While NBP has a 1.734 positive unit effect on Henry Hub from 2016 to 2020, this reduces to 0.112 post 2020. Similarly, JKM experiences a 14 fold decrease in its effect on Henry Hub moving between these periods.

Visualisations of the predicted Henry Hub price according to estimates in Panel B of Table 12 are provided in Figure 4. The individual figures are split by region: the first two figures show the effect of Asian benchmarks, JKM and Japanese LNG import prices, while the last two figures shows the effect of European benchmarks, TTF and NBP, on Henry Hub.

Table 12: DOLS: Long Term Estimates for Pre-2016 & Post-2016

	Panel A: Pre-2016			Panel B: Post-2016		
	δ_0	β_2	(n_1, n_2)	δ_0	β_2	(n_1, n_2)
NBP	4.889***	0.493***	(0,1)	2.462***	0.111***	(8,3)
	(0.816)	(0.109)		(0.147)	(0.009)	
TTF	3.681***	0.087	(0,8)	2.561***	0.096***	(6,8)
	(0.685)	(0.063)		(0.118)	(0.007)	
JKM	1.236***	0.224***	(0,0)	2.28***	0.111***	(8,0)
	(0.262)	(0.023)		(0.136)	(0.01)	
JLNG				−0.024	0.297***	(3,0)
				(0.575)	(0.05)	
Brent	5.655***	0.013	(0,0)	−0.238	0.055***	(0,1)
	(0.67)	(0.007)		(0.463)	(0.006)	
WTI	4.799***	0.024**	(0,0)	−0.508	0.062***	(0,1)
	(0.709)	0.008		(0.444)	(0.006)	

* Significant at the 10% level; ** Significant at the 5% level; ; *** Significant at the 1% level; Number of leads (n_1) and number of lags (n_2) chosen according to BIC; Newey-West standard errors; Bartlett kernel

Table 13: DOLS: Long Term Estimates for 2016-2020 & Post-2020

	Panel A: 2016 – 2020			Panel B: Post-2020		
	δ_0	β_2	(n_1, n_2)	δ_0	β_2	(n_1, n_2)
NBP	1.13	1.734***	(4, 0)	2.3***	0.112***	(3, 0)
	(1.19)	(0.352)		(0.277)	(0.012)	
TTF	2.193***	0.17***	(0, 0)	2.292***	0.101***	(6, 7)
	(0.241)	(0.035)		(0.216)	(0.008)	
JKM	2.132***	0.143***	(0, 0)	2.014***	0.01***	(5, 0)
	(0.212)	(0.024)		(0.251)	(0.116)	
JLNG	2.507**	0.074	(1, 0)			
	(0.743)	(0.067)				
Brent				−0.755	0.067***	(0, 1)
				(0.514)	(0.007)	
WTI				−0.858	0.072***	(0, 1)
				(0.51)	(0.007)	

* Significant at the 10% level; ** Significant at the 5% level; ; *** Significant at the 1% level; Number of leads (n_1) and number of lags (n_2) chosen according to BIC; Newey-West standard errors; Bartlett kernel

4.4 Short-term estimates

4.4.1 Threshold Cointegration Tests

Threshold cointegration test are reported in Table 14, which reports the results of the SupLM test for data subsetted in the pre-2016, post-2016, 2016 – 2020, and 2020 onwards sub-samples. The SupLM test finds that in the post-2016 period, Brent Crude and WTI are better modelled by a threshold ECM rather than a linear ECM, while Japanese LNG import prices are found to have threshold effects at the 10% level for the 2016 to 2020 period. This indicates that many of the bivariate relationships between Henry Hub and international natural gas prices are better modelled by linear ECMs as opposed to threshold ECMs. It also infers that asymmetric adjustment towards the equilibrium is not necessarily a key feature of the equilibrating relationships between Henry Hub and foreign natural gas prices.

Table 14: SupLM Tests for Threshold Cointegration

	Pre-2016		Post-2016		2016-2020		Post-2020	
	Stat	Threshold	Stat	Threshold	Stat	Threshold	Stat	Threshold
Brent	47.073	4.095	67.949**	-1.795			14.794	0.289
WTI	63.954	2.329	64.156*	-1.139			16.72	0.26
NBP	15.079	3.642	13.31	3.611	18.957	-8.42	15.564	3.569
TTF	11.433	3.882	10.644	3.944	12.572	1.53	11.983	3.514
JKM	20.11	1.189	53.436	3.109	17.774	1.531	46.758	5.828
JLNG			10.787	0.956	14.789*	2.397		

* Significant at the 10% level; ** Significant at the 5% level; ***0 Significant at the 1% level; 1000 bootstraps utilised, grid search of 300; p values determined by fixed regressor bootstrap while test statistics are maximised for their threshold values; number of lags chosen according to AIC

4.4.2 Linear Error Correction Model Estimates

The short-term error correction term (ECT) estimates are reported in Table 15 for variables that are linearly cointegrated with Henry Hub prices. The table is interpreted such that the Δhh column reports the ECT for the speed of adjustment for Henry Hub prices to its long run equilibrium with the referenced price in the first column, while the Δp column reports the speed of adjustment for the referenced price to Henry Hub prices. As such, the table reports that U.S.

and British prices adjust faster to each other moving into the post-2016 period than in the pre-2016 period, with estimates increasing from 0.0031 to 0.0135 for Henry Hub and from 0.0025 to 0.0925 for NBP prices. This adjustment of NBP prices to Henry Hub prices increases further to 0.1351 moving into the post-2020 period. This pattern also holds for TTF prices, which display a higher speed of adjustment estimate to Henry Hub prices of 0.1574 in the post 2020 period compared to the 0.116 estimate in the post 2016 period. Henry Hub prices also adjust faster to TTF prices moving from the pre-2016 period to the 2016-2020 period and the post-2020 period. For Asian prices, the short-run adjustments of Henry Hub to the Japan/Korea Marker (JKM) prices between the pre-2016 and 2016-2020 estimates are relatively similar. The estimates of Japanese LNG import pries are largely insignificant, except for the 2016-2020 period which shows adjustment of Henry Hub prices to Japanese LNG import prices to the long run equilibrium at a rate of 33.91% per month. This coefficient is expectedly larger than the other variables shown since it is quoted in a monthly frequency.

For oil prices, we see that there are no significant adjustment coefficients of oil prices to Henry Hub prices in any period. However, adjustment coefficients for Henry Hub to oil prices are significant and increasing from pre-2016 to post-2016 through to post-2020. Across the post-2020 period, Henry Hub prices adjust quickest to TTF prices out of all natural gas prices and oil prices, while European prices adjust the quickest to Henry Hub prices.

Table 15: Error Correction Model (ECM): Error Correction Term Estimates

	Pre-2016		Post-2016		2016-2020		Post-2020	
	Δhh	Δp	Δhh	Δp	Δhh	Δp	Δhh	Δp
NBP	-0.0031^{**}	0.0025^{*}	-0.0135^{***}	0.0925^{**}	-0.0009	0.0021	-0.0127^{*}	0.1351^{**}
	(0.0010)	(0.0011)	(0.0039)	(0.0281)	(0.0011)	(0.0014)	(0.0054)	(0.0462)
TTF	-0.0095^{***}	-0.0011	-0.0005	0.116^{***}	-0.0227^{***}	0.0059	-0.0217^{***}	0.1574^{**}
	(0.0027)	(0.0052)	(0.0026)	(0.0203)	(0.0066)	(0.0108)	(0.0064)	(0.0597)
JKM	-0.0270^{**}	0.0378	-0.0006	0.0305	-0.0256^{***}	0.0069	-0.0069	0.0120
	(0.0088)	(0.0265)	(0.0040)	(0.0202)	(0.0070)	(0.0143)	(0.0036)	(0.0210)
JLNG			-0.0831	0.199	-0.3391^{*}	0.1562		
			(0.1247)	(0.1032)	(0.1521)	(0.1361)		
Brent	-0.0031^{**}	0.0024	-0.0128^{***}	0.0285			-0.021^{***}	0.0295
	(0.0009)	(0.0048)	(0.0035)	(0.0280)			(0.006)	(0.0450)
WTI	-0.0034^{***}	0.0028	-0.0147^{***}	0.0232			-0.0213^{***}	0.0359
	(0.0010)	(0.0054)	(0.0037)	(0.0279)			(0.0061)	(0.0452)

* Significant at the 10% level; ** Significant at the 5% level; *** Significant at the 1% level; Lags are chosen according to the AIC. Δhh column shows the estimated ECT for the short-run adjustment of Henry Hub prices to the referenced price in the left-hand column; Δp shows the adjustment of the price series in the left-hand column to Henry Hub prices. The long-run coefficients are taken from the results of the DOLS estimates from Table 12 and Table 13.

4.4.3 Threshold Error Correction Model Estimates

For those time periods and variables that are reported to have threshold effects, the threshold ECM results are reported in Table 16. Regime 1 refers to when $ECT_{t-1} < \theta_L$; Regime 2 refers to when $\theta_L < ECT_{t-1} < \theta_U$; Regime 3 refers to when $\theta_U < ECT_{t-1}$, where the $ECT = hh_t - \beta p_t$, and β is the long run estimate from Table 12 and Table 13. The 'Split (%)' mentioned in the row below the provided standard errors refers to the percentage of observations in each regime. The table reports that in Regime 1, when the price spread between Henry Hub and Brent and Henry Hub and WTI is negative, both WTI and Brent prices strongly over-adjust back to the equilibrium (as the ECT term in this regime is negative). However, this coefficient is out of the bounds of what should be expected from ab=n error correction adjustment processes (around 0 to -2) and thus is evidence of a misspecified model or extremee explosive behavior in the underlying time series. Further, this regime only has 5% of the total observations within it, indicating it is an 'extreme' regime. In Regime 3, when the price differential between Henry Hub and Brent is positive, Henry Hub prices will adjust down to Brent prices at a rate of 4%

day. This rate is slightly faster than the single-regime linear estimate for post-2016 period from Table 15 of 1.28% per day.

Further, when the price spread between Henry Hub and Japanese LNG import prices is positive, Henry Hub prices will adjust back to the long-run equilibrium after a single period. The results for Regime 2 are across all variables are largely significant, and results for the Japanese LNG import price appear spurious given the extremely large coefficient value of -32.19.

Table 16: Threshold Error Correction Model: Error Correction Term Estimates

	Regime 1		Regime 2		Regime 3		θ_L	θ_U
	Δhh	Δp	Δhh	Δp	Δhh	Δp		
Brent	0.218	4.23^{***}	-0.004	0.0200	-0.04^{**}	-0.207^{*}	-1.576	0.939
Post-2016	(0.06)	(0.00)	(0.51)	(0.633)	(0.001)	(0.029)		
$\beta = 0.055$, Split = 5% / 84.4% / 10.6%								
WTI	0.139	4.042^{***}	-0.005	0.07	-0.035^{***}	-0.011	-1.721	-0.096
Post-2016	(0.261)	(0.00)	(0.643)	(0.336)	(0.00)	(0.847)		
$\beta = 0.062$, Split = 5% / 62% / 32.9%								
JLNG	-0.225	0.839^{**}	-32.19^{***}	-10.309	-1.087^{***}	-0.038	2.397	2.519
2016-2020	(0.789)	(0.00)	(0.558)	(0.772)	(0.315)	(0.536)		
$\beta = 0.297$, Split = 39.1% / 10.9% / 50%								

* Significant at the 10% level; ** Significant at the 5% level; *** Significant at the 1% level; Lags are chosen according to the AIC. Split refers to the percentage of observations in each regime.

5 Discussion

The differing results for the presence of a unit root amongst different frequencies of commodity prices reflects ongoing discussion of the topic. This paper concludes that natural gas are non-stationary at levels, which is in alignment with other studies that similarly concluded that spot and future energy prices are non-stationary (Pindyck, 1999; Maslyuk & Smyth, 2007). Further, this paper's conclusion is in alignment with many studies specific to the analysis of natural gas, which similarly fail to reject the presence of a unit root at levels across differing frequencies of observation (Gebre-Mariam, 2011; Walls, 1994; Serletis, 1997; Presno et al., 2014).

However, in an analysis specific to natural gas spot and future prices observed at the daily frequency, Mishra and Smyth (2016) find that unit roots tests with two structural breaks that account for heteroscedasticity will conclude natural gas prices are mean reverting. Further, in a recent meta analysis, Winkelried (2021) finds evidence against unit roots in real commodity prices. Interpreting this from an economic perspective, a stationary process implies that the real prices of natural gas prices will revert to their long-run marginal cost and any shocks will not induce persistent nor permanent change. This implies that events such as hydraulic fracturing, the emergence of a global LNG market, and the recent European crisis have not substantially impacted gas prices. Intuitively, this appears to be a spurious conclusion, and as the CHLT test that accounts for more parameter adjustments reports, there is significant evidence supporting the non-stationary properties of the data in this study.

Turning to the cointegration results amongst the full sample and their endogenously determined breakpoints, there appears to be a high level of cointegration between all international natural gas prices and oil prices with U.S. prices. The exception to this is Japanese LNG import prices, though this could be attributable to the lower power tests of cointegration have when data is temporally aggregated (as discussed in section 2.2). For both the Brent Crude and WTI data, both the Gregory-Hansen and Maki test detect breaks occurring around the 2008 to 2009. These dates occur around the Global Financial Crisis (GFC), which caused a significant pertubation in oil markets as oil prices, which are denominated in US dollars, adjusted upwards as the U.S. dollar fell relative to other currencies. As oil prices recovered, natural gas prices remained depressed due to to high storage levels from the expansion of shale gas production and a weak economic recovery (Brown, 2017). The effect of this is supported by the long term estimates of Table 8, which show a marked decrease in the long run effects of WTI and Brent Crude on Henry Hub prices after 2009. However, this result is slightly complicated by the finding that long term estimates for the time period after 2016 are higher (0.04 units) for WTI than in the pre-2016 period. Notably, this trend does not significantly increase moving into the post-2020 period, with estimates for both WTI and Brent in the post-2020 period just 0.01 higher than in the post-2016 period. This indicates that even if the influence of oil prices on Henry Hub natural gas prices has strengthened in the period after 2016 relative to the entire period of 2016, this influence has not continued to substantially increase through to the post 2020 period.

Further, this papers finds strong evidence that European prices exerted a stronger influence on U.S. natural gas prices between 2016 and 2020. In particular, the long term estimate on British prices increases from pre-2016 levels of 0.493 to 1.734, while the estimate on Dutch prices moves from insignificance to a 0.17 effect. Further, while we see no evidence for cointegration between Japanese LNG import prices prior to 2016, we see evidence of cointegration moving into the 2016-2020 period at the 5% level. These findings are even stronger considering the lack of cointegration with oil prices during this period, which diminishes the extent to which oil-indexation could attribute to the price co-movement during this time. However, the decrease in the long-term estimate on the Japan/Korea Marker is in contrast to the aforementioned patterns. This could be attributable to the noted decoupling the S&P, who constructs the Asian marker, find between JKM and other LNG price indexes during 2018[17]. In particular, they note during this time that the pro-active buying of LNG by Chinese and South-Korean markets for the winter drove up JKM prices in the summer of 2018, while increasing Atlantic and Pacific LNG charter rates reduced the competitiveness of U.S. LNG. These patterns are reflected in the bottom panel of Figure 4, where perturbations to the JKM price surrounding 2018 are not well reflected in the price of Henry Hub.

Across the European prices, NBP and TTF, the Asian prices, JKM, and well as the two oil prices, it can be seen that cointegration tests endogenously confirm a relatively common break-

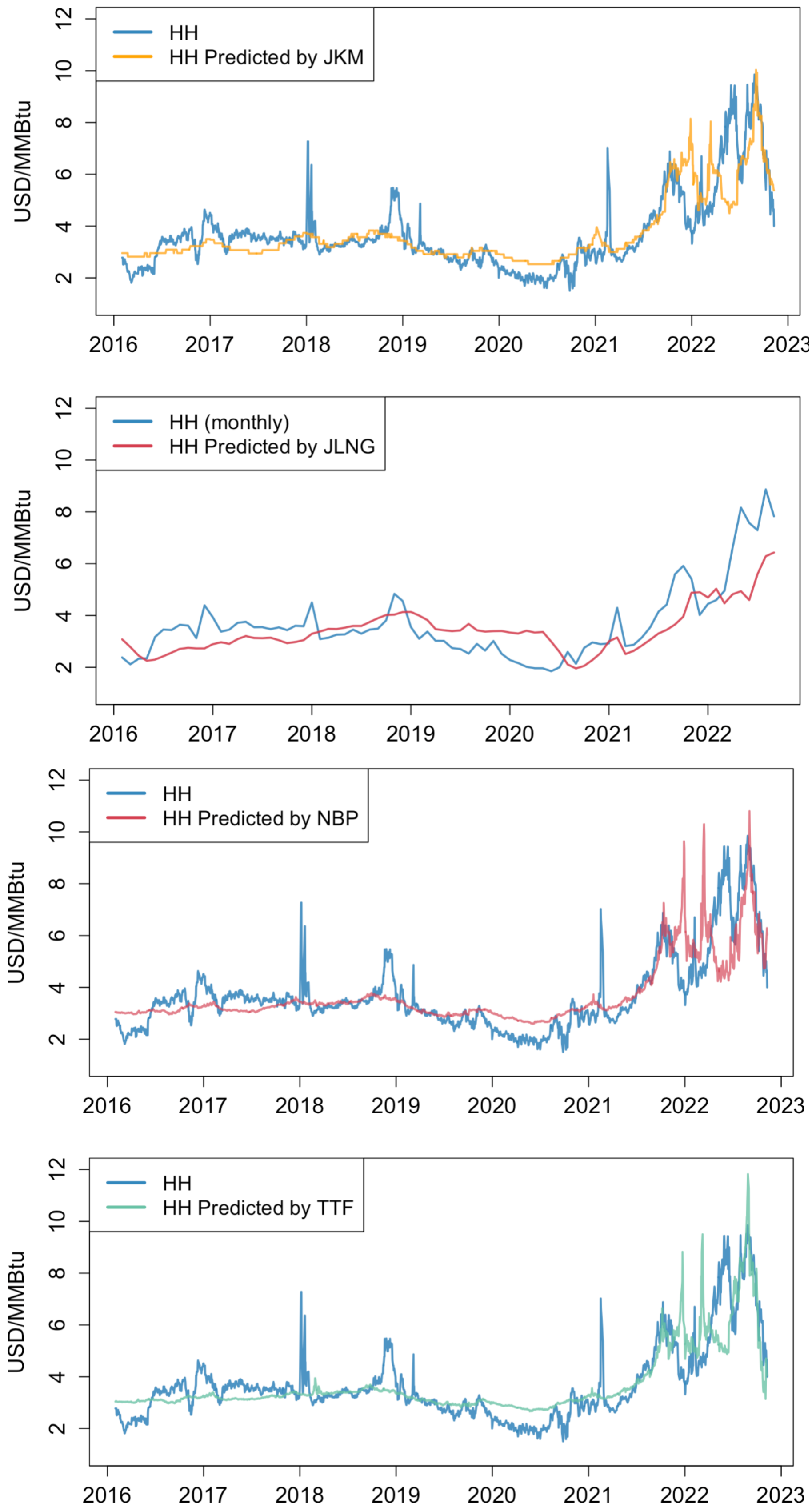
[17]Source: S&P Global

point just prior to the Russian invasion of Ukraine. Figure 9 and Figure 10 visually show the commonality of this breakpoint. Notably, long term estimates for the effect of international natural gas prices after these common breaks are both nominally and relatively small compared to prior estimated periods. This result is confirmed in the post-2020 sub-sample in Panel B of Table 13, which shows lower long term estimates between natural gas prices and Henry Hub after 2020 compared to before 2020. While long-term estimates for natural gas prices in this post-2020 period are relatively weaker, the short-term adjustment effects are significantly stronger than in previous periods. Compared to the entirety of the post-2016 period, in the post-2020 period TTF prices adjust 1.35 times as fast and NBP prices adjust 1.46 times as fast. Further, the speed of adjustment of Henry Hub to NBP and TTF price moving from the pre-2016 to the post-2020 period increases by a factor of 4 and 2.28, respectively. This could indicate that while prices have not necessarily moved towards a law of one price since 2020, the speed at which price transmissions are disseminated around global markets has increased since 2020. Indeed, plotting the predicted price of Henry Hub based off of the estimated long-term relationship between European & Asian prices (Figure 4), the similarity of price patterns can clearly be seen. In particular, the two spikes occurring just prior to 2022 and just before the end of the sample period are well predicted by the JKM, NBP, and TTF markers.

There is limited evidence to support asymmetric adjustment towards the equilibrium relationship among natural gas prices; however, there is evidence of threshold adjustment between U.S. natural gas prices and Brent Crude and WTI during the post-2016 period. From these models, it appears that Henry Hub prices adjust down at a rate of 4% to Brent prices when the price differential between Henry Hub and Brent is positive, while it adjusts at less than half that rate in the linear error correction model. While comparison is difficult due to the absence of threshold effects in other time periods and significance in other threshold regimes, these results indicate that the rate of convergence to equilibrium between oil and gas prices is particularly strongest when gas prices are adjusting downwards.

Further, this analysis reveals that cointegration may hold over an entire sample but breakdown in sub-periods of the sample. This could potentially be due to the effect of decreasing the sample size, as cointegration tests are asymptotic and have less power in being able to identify non-stationarity in shorter periods of time. As such, further research should be conducted into tests that can account for the simultaneous presence of unit roots in parts of the sample and no unit roots in other parts of the sample.

Figure 4: Predictions of Henry Hub Post-2016 Based on Asian & European Benchmarks



All shown prices are realised to 2022 prices. Predicted Henry Hub prices are calculated from estimates given in Panel B of Table 12

Figure 5: Henry Hub vs Brent Crude and WTI: Long Term Estimates Based on Maki Sub-samples

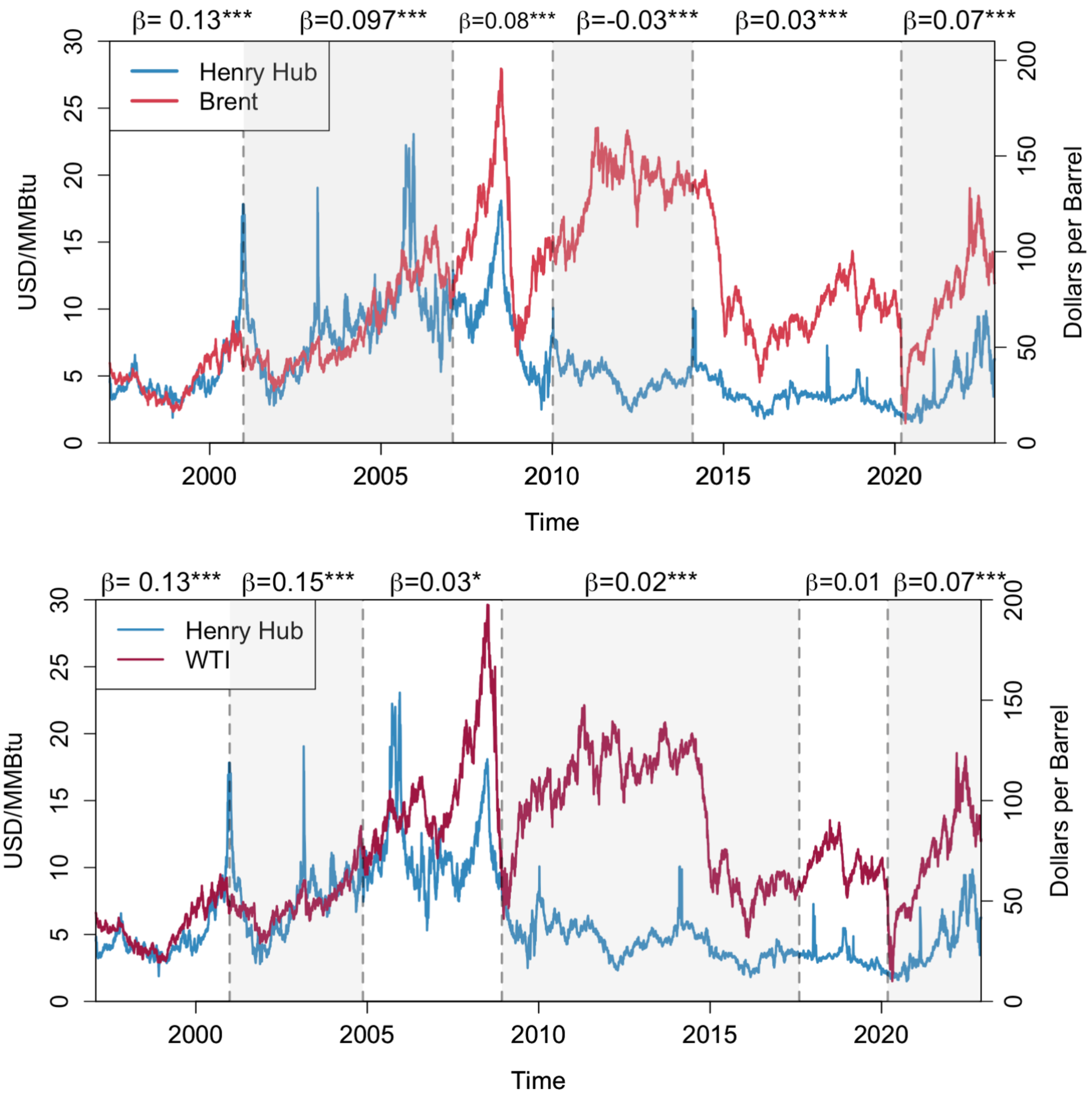
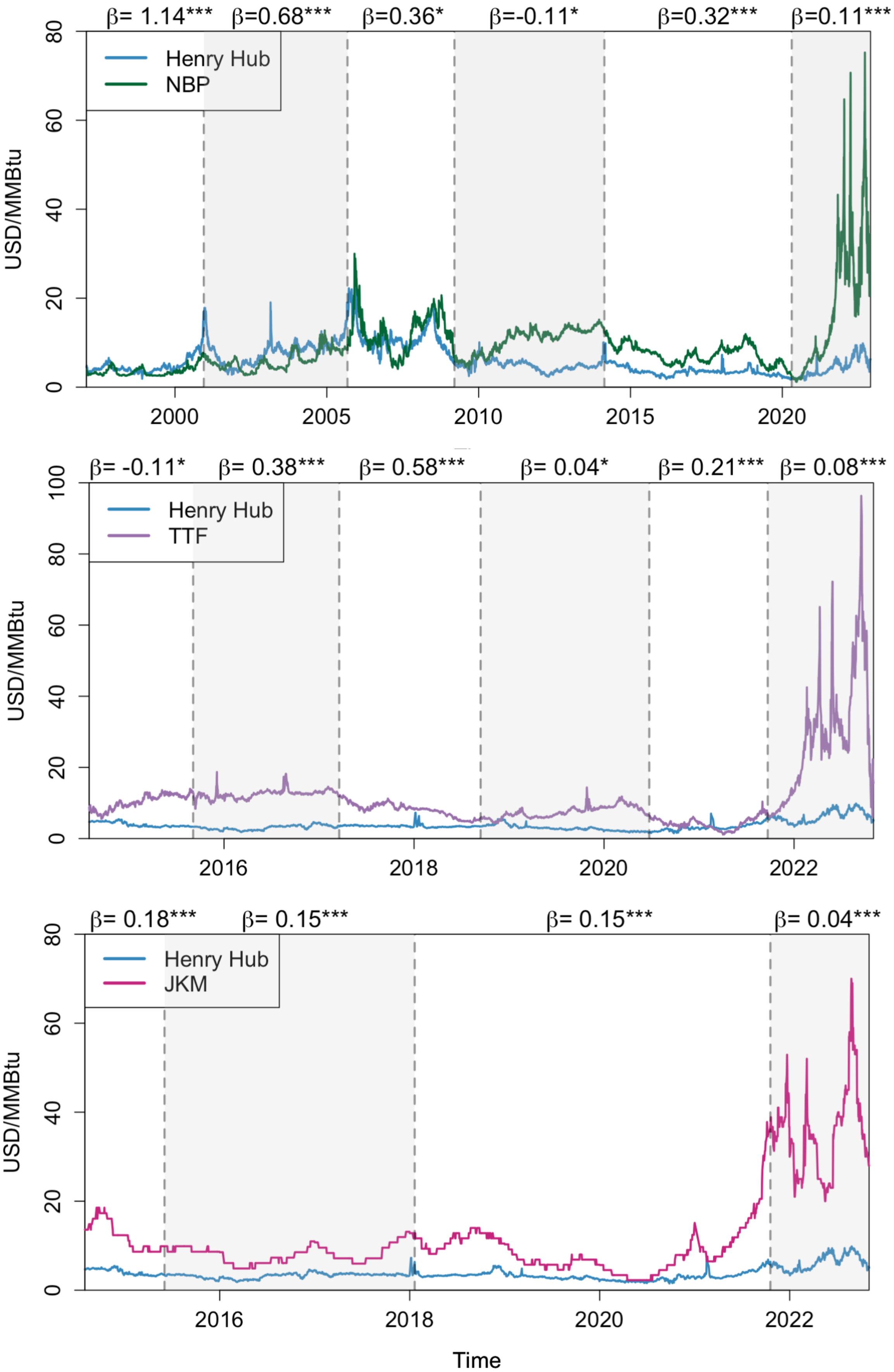


Figure 6: Henry Hub vs European & Asian Prices: Long Term Estimates based on Maki Sub-samples



6 Conclusion & Policy Implications

The energy crisis in Europe has caused a reverberating shock throughout global energy markets, with the potential to permanently change the dynamic relationships between energy commodities. In particular, the crisis has presented a serendipitous opportunity to the burgeoning U.S. LNG market, whose institutional development has been accelerated by a harmonious interlinking of public and private desires to ensure energy security and profit off of substantial price differentials.

As such, this paper has sought to assess the impacts of U.S. LNG exports on cointegration between European & Asian natural gas benchmarks and oil prices. Further, it has also attempted to assess how this integration has changed in the wake of the constriction of energy exports to Europe from its largest energy partner.

This paper confirms recent research that European prices and U.S. prices became more integrated in the 2016-2020 period whilst U.S. natural gas prices decoupled from oil prices. However, extending this analysis to current price data presents mixed results. In particular, the long-run equilibrium relationship between Asian and European prices and Henry Hub prices is weaker in the post-2020 period than in prior periods. Further, oil prices appear to have a stronger long-term relationship with U.S. gas prices in the post 2020 period than prior to 2016. Notably, the short run adjustment processes between Henry Hub prices and international natural gas prices have strengthened in the post-2020 period compared to all other periods. Results thus indicate that while prices do not resemble a law of one price, the speed at which prices respond to changes in foreign markets has increased substantially after the emergence of U.S. LNG exports and after the European energy crisis.

While this paper has placed a focus on bivariate frameworks, further research may utilise a multivariate framework to account for extenuating factors or interacting effects. This is a limitation of our paper, since both the short-term and long-term effects of international prices on Henry Hub do not directly account for the presence of oil-indexation. Further areas of research may also include the effect of natural gas prices on coal prices, which have recently experienced a surge in prices due to recent spikes in electricity demand,. In tandem with this, other areas of investigation could pertain to the direct effect of increased natural gas prices on energy consumption & the predicted effect on emissions levels.

The broader implications of this study first include the observation that oil prices have not necessarily decoupled from U.S. natural gas prices. This may be due to how the rise in LNG trade within the U.S. has equilibrated Henry Hub prices to both natural gas and LNG export prices, the latter of which, especially in Asia, are still predominantly oil-indexed. Further, U.S. natural gas prices have responded more quickly to changes in international natural gas prices since 2020, indicating greater market integration regarding price transmission. While this doesn't necessarily indicate a persistent increase in the mean level of Henry Hub prices, it does indicate that domestic gas prices could become less robust to international macroeconomic shocks in the future. Accordingly, investment in energy sources that are less prone to market turmoil, such as renewables, may reduce the United States exposure to international financial and geopolitical shocks.

Appendix A

A.1 Imputation choices

Details concerning the dates of imputation are included below, listing the variable and the date:

1. Henry Hub, February 2003[1]: a price spike in physical next-day delivery prices was caused by a combination of high demand induced by a cold front across the eastern half of the United States, and limited supply induced by well freeze-offs in the Mid-continent. In addition, the unusually cold winter that had preceded the cold front had lowered gas storage inventories below the late-February 5-year average by 42%.

2. Henry Hub , February 2021[2]: a price spike occurs here due to the arctic storm that preceded the Texas power crisis. Specifically, there was a large increase in demand from high energy consumption due to the intense winter storm whilst a severe decrease in supply from natural gas refineries along the Gulf coast whilst they suffered well freeze-offs.

3. JKM, January 2021[18]: JKM spiked due to demand driven by a preceding cold spell in Bejing; a lack of seasonal storage which concurrently limited supply; and a lack of spare LNG tanker capacity which created transportation constraints.

4. TTF, February 2018[19]: a price spike was caused by extremely low temperatures in the UK and surrounding regions, as well as limited gas in Dutch storages at the time.

5. WTI, April 2020[20]: oil prices temporarily turned negative with the onset of the coronavirus pandemic, which sharply decreased demand in a time when there was high oil inventories.

[1]Source: FERC
[2]Source: S&P Global and Oxford Energy
[18]Source: S&P Global
[19]Source: REKK, Regional Centre for Energy Policy Research
[20]Source: NYT

Appendix B

B.1 Supplementary Unit Root Tests

Table B.1: Monthly ADF tests: Levels & First Differences

	Levels			Differences		
	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
HH	$-1.14(5)$	$-2.73(5)^{*}$	$-3.11(5)$	$-7.58(5)^{***}$	$-7.57(5)^{***}$	$-7.56(5)^{***}$
TTF	$2.594(4)$	$2.159(4)$	$2.144(4)$	$-4.32(4)^{***}$	$-4.44(4)^{***}$	$-4.76(4)^{***}$
JKM	$0.384(3)$	$-0.399(3)$	$-1.23(3)$	$-6.18(3)^{***}$	$-6.24(3)^{***}$	$-6.69(3)^{***}$
NBP	$-0.2887(5)$	$-1.93(5)$	$-2.60(5)$	$-8.54(5)^{***}$	$-8.61(5)^{***}$	$-8.71(5)^{***}$
WTI	$-0.655(5)$	$-2.49(5)$	$-2.52(5)$	$-8.49(5)^{***}$	$-8.48(5)^{***}$	$-8.48(5)^{***}$
Brent	$-0.629(5)$	$-2.31(5)$	$-2.36(5)$	$-8.15(5)^{***}$	$-8.15(5)^{***}$	$-8.15(5)^{***}$

* Significant at the 10% level; ** Significant at the 5% level; ; *** Significant at the 1% level; () denote lag length. Model 1: No drift nor trend ; Model 2: With drift, no trend ; Model 3: With drift and trend

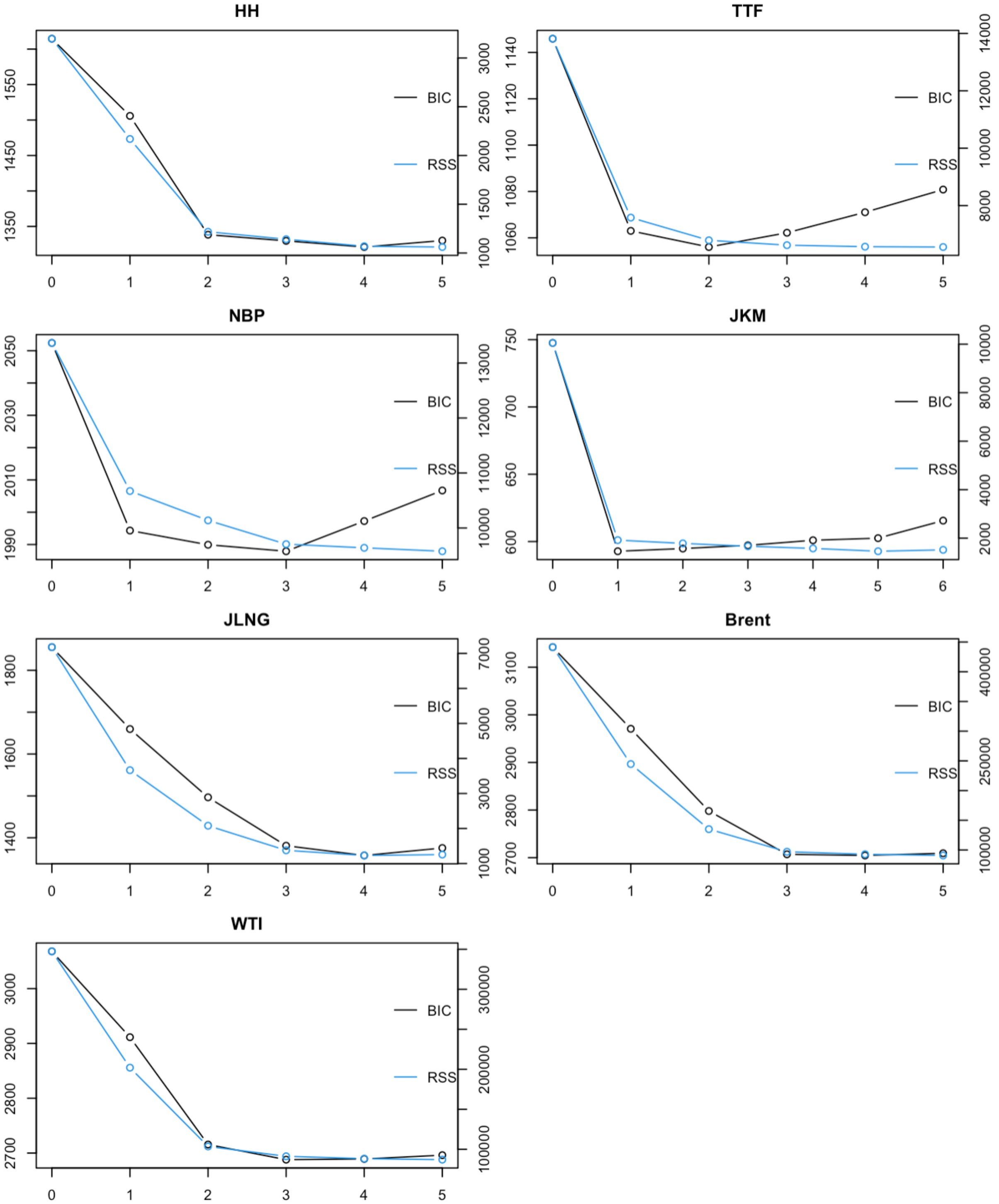
Table B.2: Daily and Weekly ADF tests: differences

	Daily			Weekly		
	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
Differences						
HH	$-31.1(10)^{***}$	$-31.1(10)^{***}$	$-31.1(10)^{***}$	$-13.6(7)^{***}$	$-13.6(7)^{***}$	$-13.6(7)^{***}$
TTF	$-22.8(9)^{***}$	$-22.8(9)^{***}$	$-22.8(9)^{***}$	$-7.02(6)^{***}$	$-6.99(6)^{***}$	$-6.93(6)^{***}$
JKM	$-17.8(8)^{***}$	$-17.8(8)^{***}$	$-17.8(8)^{***}$	$-9.44(5)^{***}$	$-9.44(5)^{***}$	$-9.53(5)^{***}$
NBP	$-31.6(10)^{***}$	$-31.6(10)^{***}$	$-31.7(10)^{***}$	$-17.0(7)^{***}$	$-17.0(7)^{***}$	$-17.0(7)^{***}$
WTI	$-30.1(10)^{***}$	$-30.1(10)^{***}$	$-30.1(10)^{***}$	$-10.3(7)^{***}$	$-10.3(7)^{***}$	$-10.3(7)^{***}$
Brent	$-29.9(1)^{***}$	$-28.9(10)^{***}$	$-28.9(10)^{***}$	$-11.1(7)^{***}$	$-11.1(7)^{***}$	$-11.1(7)^{***}$

* Significant at the 10% level; ** Significant at the 5% level; ; *** Significant at the 1% level; () denote lag length. Model 1: No drift nor trend ; Model 2: With drift, no trend ; Model 3: With drift and trend

B.2 Structural Breaks: BIC & RSS Scores

Figure B.1: BIC & RSS Results



* Left axis is BIC scores; right axis is RSS scores; bottom axis is breakpoints

B.3 Supplemntary Cointegration Tests

Table B.3: Results of Maki Cointegration Test: $m = 3$

vs. HH	Model	Statistic	Breakpoints
JKM	2	-5.895^{**}	04-Jun-2015, 20-Jan-2018, 19-Oct-2021
	3	-6.199	
JLNG [1]	2	-5.398	
	3	-5.924	
NBP	2	-8.295^{***}	30-May-2000, 30-Dec-2002, 14-Feb-2014
	3	-8.924^{***}	11-Dec-2000, 30-Aug-2005, 07-Mar-2009
TTF[2]	2	-6.812^{***}	06-Sep-2011, 18-Oct-2015, 24-Apr-2021
	3	-6.663^{**}	17-Sep-2011, 07-Feb-2014, 15-Feb-2021
Brent	2	-7.483^{***}	5-Jun-2000, 26-Feb-2003, 20-Jan-2018
	3	-7.252^{***}	28-Dec-2000, 6-Feb-2007, 14-Mar-2020
WTI	2	-7.538^{***}	27-Mar-2000, 5-Nov-2004, 14-Feb-2014
	3	-7.348^{***}	28-Dec-2000, 8-Aug-2017, 9-Mar-2020

*** Significant at the 1% level; ** Significant at the 5% level, * Significant at the 10% level. Model 2 = Regime Shift (11), Model 3 = Regime Shift and trend shift (12). Trimming parameter $eta = 0.10$. Up to 3 structural breaks allowed in the cointegrating vector. Tests conducted in Gauss (Nazlioglu, 2021)

1. Use $\eta = 0.11$ for model 3 as unable to compute with $\eta = 0.10$
2. Use $\eta = 0.12$ for model 2 as unable to compute with $\eta = 0.10$

Table B.4: Results of Maki Cointegration Test: $m = 4$

Vs. HH	Model	Statistic	Breakpoints
JKM	2	-5.895^{*}	04-Jun-2015, 09-Dec-2016, 20-Jan-2018, 19-Oct-2021
	3	-6.315	
Jap LNG[1]	2	-5.398^{*}	04-Jun-2015, 09-Dec-2016, 20-Jan-2018, 19-Oct-2021
	3	-6.008	
NBP	2	-8.547^{***}	30-Jun-2000, 30-Dec-2002, 28-Mar-2009, 14-Feb-2014
	3	-8.924^{***}	11-Dec-2000, 30-Aug-2005, 07-Mar-2009, 06-Apr-2020
TTF[2]	2[1]	-7.706^{***}	06-Sep-2011, 27-Mar-2013, 18-Oct-2015, 24-Apr-2021
	3	-7.428^{**}	17-Sep-2011, 07-Feb-2014, 03-Jun-2016, 15-Feb-2021
Brent	2	-7.733^{***}	05-Jun-2000, 26-Feb-2003, 18-Nov-2005, 20-Jan-2018
	3	-7.355^{**}	18-Dec-2000, 06-Feb-2007, 07-Feb-2014, 14-Mar-2020
WTI	2	-7.906^{***}	27-Mar-2000, 05-Nov-2004, 23-Dec-2008, 14-Feb-2014
	3	-7.365^{**}	28-Dec-2000, 6-Dec-2008, 08-Aug-2017, 09-Mar-2020

*** Significant at the 1% level; ** Significant at the 5% level, * Significant at the 10% level. Model 2 = Regime Shift, Model 3= Regime Shift and trend shift. Trimming parameter $\eta = 0.10$. Up to 4 structural breaks allowed in the cointegrating vector. Tests conducted in Gauss (Nazlioglu, 2021)

2. Use an $\eta = 0.11$ for model 3 as test would not run with $\eta = 0.10$

2. Use an $\eta = 0.12$ for model 2 as test would not run with $\eta = 0.10$

References

Bai, J., & Perron, P. (1998). Estimating and testing linear models with structural changes. *Econometrica*, *66*(1), 47–78.

Bai, J., & Perron, P. (2003). Computation and analysis of multiple structural change models. *Journal of Applied Econometrics*, *18*(1), 1–22.

Balke, N. S., & Fomby, T. B. (1997). Threshold cointegration. *International Economics Review*, *38*(3), 627–645.

Barbe, A., & Riker, D. (2015). Obstacles to international trade in natural gas.

Barnes, R., & Bosworth, R. (2015). Lng is linking regional natural gas markets: Evidence from the gravity model. *Energy Economics*, *47*, 11–17.

Batten, J. A., Ciner, C., & Lucey, B. M. (2017). The dynamic linkages between crude oil and natural gas markets. *Energy Economics*, *62*, 155–170.

Brown, S. P. A., & Yücel, M. K. (2009). Market arbitrage: European and north american natural gas prices. *The Energy Journal*, *30*, 167–185.

Brown, S. P. (2005). Natural gas pricing: Do oil prices still matter? *Federal Reserve Bank of Dallas*.

Brown, S. P. (2017). Natural gas vs. oil in u.s. transportation: Will prices confer an advantage to natural gas? *Energy Policy*, *110*, 210–221.

Buberkoku, O. (2017). Examining energy futures market efficiency under multiple regime shifts. *International Journal of Energy Economics and Policy*, *7*(6), 61–71.

Busetti, F., & Taylor, A. M. R. (2003). Variance Shifts, Structural Breaks, and Stationarity Tests. *Journal of Business & Economic Statistics*, *21*(4), 510–531.

Campbell, J. Y., & Perron, P. (1991). *Pitfalls and opportunities: What macroeconomists should know about unit roots* (Working Paper No. 100). National Bureau of Economic Research.

Campos, J., Ericsson, N. R., & Hendry, D. F. (1996). Cointegration tests in the presence of structural breaks. *Journal of Econometrics*, *70*(1), 187–220.

Cavaliere, G., Harvey, D. I., Leybourne, S. J., & Taylor, A. R. (2011). Testing for unit roots in the presence of a break in trend and nonstationary volatility. *Econometric Theory*, *27*(5), 957–991.

Cavaliere, G., & Taylor, A. M. R. (2006). Testing for a Change in Persistence in the Presence of a Volatility Shift. *Oxford Bulletin of Economics and Statistics*, *68*(s1), 761–781.

Chiappini, R., Jégourel, Y., & Raymond, P. (2019). Towards a worldwide integrated market? New evidence on the dynamics of U.S., European and Asian natural gas prices. *Energy Economics*, *81*, 545–565.

Choi, I. (1992). Effects of data aggregation on the power of tests for a unit root: A simulation study. *Economics Letters*, *40*(4), 397–401.

Choi, I., & Kurozumi, E. (2012). Model selection criteria for the leads-and-lags cointegrating regression [Recent Advances in Nonstationary Time Series: A Festschrift in honor of Peter C.B. Phillips]. *Journal of Econometrics*, *169*(2), 224–238.

De Bock, R., & José, G. (2011). Will natural gas prices decouple from oil prices across the pond? *IMF Working Paper*.

Dickey, D., & Fuller, W. (1979). Distribution of the estimators for autoregressive time series with a unit root. *JASA. Journal of the American Statistical Association*, *74*.

Engle, R. F., & Granger, C. W. J. (1987). Co-integration and error correction: Representation, estimation, and testing. *Econometrica*, *55*(2), 251–276.

Franses, P. H., & Haldrup, N. (1994). The effects of additive outliers on tests for unit roots and cointegration. *Journal of Business and Economic Statistics*, *12*(4), 471–478.

Gebre-Mariam, Y. K. (2011). Testing for unit roots, causality, cointegration, and efficiency: The case of the northwest us natural gas market. *Energy*, *36*(5), 3489–3500.

Ghoshray, A. (2018). How persistent are shocks to energy prices? *The Energy Journal*, *39*, 175–192.

Gregory, A., & Hansen, B. (1996). Residual-based tests for cointegration in models with regime shifts. *Journal of Econometrics*, *70*, 99–126.

Gregory, A. W., & Hansen, B. E. (1996). Tests for cointegration in models with regime and trend shifts. *Oxford bulletin of economics and statistics*, *58*, 555–560.

Hakkio, C. S., & Rush, M. (1991). Cointegration: How short is the long run? *Journal of International Money and Finance*, *10*(4), 571–581.

Hansen, B. E., & Seo, B. (2002). Testing for two-regime threshold cointegration in vector error-correction models. *Journal of Econometrics*, *110*, 293–318.

Harris, D., Harvey, D. I., Leybourne, S. J., & Taylor, A. M. R. (2009). Testing for a unit root in the presence of a possible break in trend. *Econometric Theory*, *25*(6), 1545–1588.

Hartley, P. R., Medlock, K. B., & Rosthal, J. E. (2008). The relationship of natural gas to oil prices. *The Energy Journal*, *29*(3), 47–65.

Hatemi-J, A. (2008). Tests for cointegration with two unknown regime shifts with an application to financial market integration. *Empirical Economics*, *35*(3), 497–505.

Hayashi, M., & Hughes, L. (2013). The fukushima nuclear accident and its effect on global energy security. *Energy Policy*, *59*, 102–111.

Heather, P. (2010). The evolution and functioning of the traded gas market in britain. *Oxford Institute for Energy Studies*, *NG 44*.

Heather, P. (2019). European traded gas hubs: A decade of change. *Oxford Institue for Energy Studies*.

In Choi & Bhum Suk Chung. (1995). Sampling frequency and the power of tests for a unit root: A simulation study. *Economics Letters*, *49*(2), 131–136.

Inclan, C., & Tiao, G. C. (1994). Use of cumulative sums of squares for retrospective detection of changes of variance. *Journal of the American Statistical Association*, *89*(427), 913–923.

Jensen, J. T. (2003). The lng revolution. *The Energy Journal*, *24*(2), 1–45.

Ketenci, N. (2016). The bilateral trade flows of the eu in the presence of structural breaks. *Empirical Economics*, *51*, 1369–1398.

Kim, S.-H., Lim, Y.-Y., Kim, D.-W., & Kim, M.-K. (2020). Swing suppliers and international natural gas market integration. *Energies*, *13*(18), 1–12.

Lalenti, R. (2021). Rising lng exports and global natural gas price convergence.

Lee, C.-C., & Lee, J.-D. (2009). Energy prices, multiple structural breaks, and efficient market hypothesis. *Applied Energy*, *86*(4), 466–479.

Lee, J., Huang, C. J., & Shin, Y. (1997). On stationary tests in the presence of structural breaks. *Economics Letters*, *55*(2), 165–172.

Li, R., Joyeux, R., & Ripple, R. D. (2014). International natural gas market integration. *The Energy Journal*, *35*(4), 159–179.

Lo, M. C., & Zivot, E. (2001). Threshold cointegration and nonlinear adjustment to the law of one price. *Macroeconomic Dynamics*, *5*(4), 533–576.

Loureiro, J. R., Inchauspe, J., & Aguilera, R. F. (2022). World regional natural gas prices: Convergence, divergence or what? new evidence.

Maki, D. (2012). Tests for cointegration allowing for an unknown number of breaks. *Economic Modelling*, *29*(5), 2011–2015.

Maslyuk, S., & Smyth, R. (2007). *Unit Root Properties of Crude Oil Spot and Futures Prices* (Monash Economics Working Papers No. 40-07). Monash University, Department of Economics.

Mishra, V., & Smyth, R. (2016). Are natural gas spot and futures prices predictable? *Economic Modelling*, *54*, 178–186.

Mokhatab, S., Mak, J. Y., Valappil, J. V., & Wood, D. A. (Eds.). (2014). Chapter 1 - lng fundamentals. In *Handbook of liquefied natural gas* (pp. 1–106). Gulf Professional Publishing.

Nazlioglu, S. (2021). Tspdlib: Gauss time series and panel data methods (version 2.0).

Neumann, A. (2009). Linking natural gas markets – is lng doing its job? *The Energy Journal*, *30*, 187–199.

Nick, S., & Tischler, B. (2014). The law of one price in global natural gas markets, 26.

Noriega, A. E., & Ventosa-Santaulària, D. (2012). The effect of structural breaks on the engle-granger test for cointegration. *Estudios Económicos*, *27*(1 (53)), 99–132.

Perron, P. (1989a). The Great Crash, the Oil Price Shock, and the Unit Root Hypothesis. *Econometrica*, *57*, 1361–1401.

Perron, P. (1989b). Testing for a random walk: A simulation experiment of power when the sampling interval is varied. In B. Raj (Ed.), *Advances in econometrics and modelling* (pp. 47–68). Springer Netherlands.

Pierse, R., & Snell, A. (1995). Temporal aggregation and the power of tests for a unit root. *Journal of Econometrics*, *65*(2), 333–345.

Pindyck, R. S. (1999). The long-run evolution of energy prices. *The Energy Journal*, *20*(2), 1–27.

Presno, M. J., Landajo, M., & Fernandez, P. (2014). Non-renewable resource prices: A robust evaluation from the stationarity perspective. *Resource and Energy Economics*, *36*(2), 394–416.

Ramberg, D. J., & Parsons, J. E. (2012). The Weak Tie Between Natural Gas and Oil Prices. *The Energy Journal*, *0*(Number 2).

Sanso, A., Arago, V., & Carrion, J. (2003). Testing for changes in the unconditional variance of financial time series. *Universitat de les Illes Balears, Departament d'Economia Aplicada, DEA Working Papers*, *4*.

Serletis, A. (1997). Is there an east-west split in north american natural gas markets? *The Energy Journal*, *18*(1), 47–62.

Siliverstovs, B., L'Hégaret, G., Neumann, A., & Hirschhausen, C. v. (2005). International market integration for natural gas? A cointegration analysis of prices in Europe, North America and Japan. *Energy Economics*, *27*(4), 603–615.

Stauffer, T. R. (1995). The diseconomics of long-haul lng trading. *The Journal of Energy and Development*, *20*(2), 159–174.

Stigler, M. (2020). Nonlinear time series in r: Threshold cointegration with tsdyn. In H. D. Vinod & C. Rao (Eds.), *Financial, macro and micro econometrics using r* (pp. 229–264). Elsevier.

Stock, J. H., & Watson, M. W. (1993). A simple estimator of cointegrating vectors in higher order integrated systems. *Econometrica*, *61*(4), 783–820.

Walls, W. D. (1994). Price convergence across natural gas fields and city markets. *The Energy Journal*, *15*(4), 37–48.

Winkelried, D. (2021). Unit roots in real primary commodity prices? a meta-analysis of the grilli and yang data set. *Journal of Commodity Markets*, *23*, 100168.

Zhang, D., & Ji, Q. (2018). Further evidence on the debate of oil-gas price decoupling: A long memory approach. *Energy Policy*, *113*, 68–75.

Zhou, S. (2001). The power of cointegration tests versus data frequency and time spans. *Southern Economic Journal*, *67*(4), 906–21.