

**School of Accounting, Economics and Finance**

**The Dynamics of Inter-Regional Natural Gas Markets**

**Jose Roberto Loureiro**

**0000-0003-4643-1782**

**This thesis is presented for the Degree of  
Doctor of Philosophy  
of  
Curtin University**

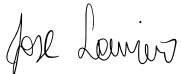
**May 2024**

## DECLARATION

To the best of my knowledge and belief, this thesis contains no material previously published by any other person, except for a paper published in the Journal of Commodity Markets (2023), with the title “*World regional natural gas prices: Convergence, divergence or what? New evidence*” that I co-authored with Julian Inchauspe and Roberto F. Aguilera. This thesis contains no material which has been accepted for the award of any other degree or diploma in any university.

I have obtained permission from the copyright owners to use any third-party copyright material reproduced in the thesis or to use any of my own published work (e.g. journal articles) in which the copyright is held by another party.

As an Elsevier journal author, I have the right to include the journal article “*World regional natural gas prices: Convergence, divergence or what? New evidence*” in a thesis or dissertation where in full or in part; subject to proper acknowledgement. No written permission from Elsevier is necessary.

Signature: 

Date: 18/07/2024

## ABSTRACT

The global trend in policymaking indicates a move towards cleaner energy in the coming years, with natural gas poised to serve as a crucial transition fuel. Understanding the economic dynamics behind regional gas markets is essential. This thesis addresses this need by examining the interactions among key gas prices in Asia Pacific, Europe, and North America and their relationship to the crude oil price.

This thesis employs robust econometric methods to test several hypotheses. It begins by assessing the degree of integration between gas markets through convergence and cointegration testing, aiming to determine if the extent of inter-regional trade eliminates arbitrage opportunities, leading to price co-movement as per the Law of One Price theory. The study also identifies convergence clusters among regional markets and evaluates the significance of the international oil price. It investigates the propagation of shocks in international gas markets using analyses such as causality assessments, impulse-response analysis and forecast variance error decomposition. Robustness tests are implemented to ensure reliability, examining asymmetric shock responses, and identifying structural breaks. Furthermore, the thesis estimates the price and income elasticities of gas demand for two major gas importers, Japan and Korea, that have received limited attention in the literature.

The above investigation leads to a rich collection of results relevant to policymakers relying on gas markets to achieve environmental goals and energy security at stable prices. This study validates earlier research by confirming that natural gas prices in major global trading hubs converged before the 2008-2009 financial crisis, driven by the dominant price-indexation of gas prices to oil. However, after the shift to gas-on-gas pricing started in 2009, this convergence ceased, forming separate clusters in gas prices. European hub prices now form a single cluster, while Asian LNG spot import prices have become increasingly integrated with European prices. Japanese gas import prices remain closely tied to crude oil prices, mostly under long-term contracts. The North American market has largely decoupled from these trends, attributed to the shale revolution's supply growth. These findings challenge the idea that current infrastructure and trade volumes, particularly in LNG, can effectively arbitrage

transoceanic gas prices. Nonetheless, there is evidence of increased integration among spot prices in Europe and Asia, accompanied by reduced influence from oil prices in recent years.

In analysing shock propagation, interesting dynamic patterns emerged. Particularly, it was found that within LNG and gas spot prices in Europe and Asia, the Russian gas export price acts as a leading force, with implications for European policies amid the Russia-Ukraine conflict. Additionally, while European and Asian prices still respond to oil price shocks, the North American price is primarily influenced by its gas market fundamentals. The study discusses further policy implications by aligning the findings with commonly predicted policy scenarios, including the anticipated expansion of LNG trade in the Asia Pacific region, ongoing gas price liberalisation, and the increasing roles of the United States and Qatar as LNG exporters.

Overall, this thesis offers important contributions to understanding economic forces operating in major gas trading hubs with a focus on price dynamics. A chapter of this thesis has already been published in *The Journal of Commodity Markets*, and further dissemination of results is expected soon.

## ACKNOWLEDGEMENTS

Foremost, I express my profound gratitude to the Divine God, whose immense guidance and blessings sustained me through the journey of my doctoral pursuit. Through Jesus Christ's grace, I navigated the challenges of this rigorous academic endeavour, emerging equipped with the necessary knowledge and resilience to complete this demanding chapter in my academic and professional journey.

I am deeply thankful to my main supervisor, Dr. Julian Inchauspe, for his indispensable guidance and support throughout my PhD studies. From the beginning, my goal was to master the time series econometric techniques used in economics, and having Julian as my supervisor was incredible. His extensive expertise in econometrics was crucial in making my doctoral work possible, as he provided constant support and shared his knowledge generously. Julian's dedication, patience, and kindness have profoundly influenced me. Working with him has been incredibly rewarding, and he has inspired me in countless ways.

I also want to thank my co-supervisors, Dr. Roberto Aguilera and Dr. Helen Cabalu. Roberto offered invaluable guidance in understanding global petroleum and natural gas markets and played a significant role in selecting me for this project. I'm also thankful to Helen for stepping into Roberto's position when he moved to Saudi Arabia. Since then, Helen has provided extensive support and essential knowledge, helping me connect my findings with qualitative assessments and policy implications. Her insightful comments greatly contributed to the success of this study.

I extend my gratitude to Dr. Greg White, who consistently entrusted me with the opportunity to collaborate with him in the unit Analytics for Decision Making, where he serves as the unit coordinator. He has gone above and beyond to mentor me in teaching, managing the unit, and handling administrative duties as an academic at the university. I am immensely grateful for his trust in appointing me as the Unit Coordinator during his extended leave in Semester 2, 2023. This opportunity provided

invaluable experience and insights into academic lecturing that I never anticipated gaining as a PhD candidate. Working alongside such an exceptional academic has been an enriching experience, and he has become a great friend and a source of personal and professional inspiration.

Moreover, I want to express my appreciation to numerous academics and staff members within the Faculty of Business and Law who have supported me throughout my journey as a sessional academic. A special thanks to Dr. Khuong Truong, the unit coordinator for Quantitative Techniques for Business, Dr. Aline Lee, Sian Flynne, Donna Miley, Sonnete Dykman, Kathy Gibson, and the Head of School, Felix Chan. Their patience and encouragement have been important in facilitating my various academic responsibilities. I am also thankful to staff members from the Faculty of Science and Engineering, including Sharon Bell and Dr Hai Huyen Heidi Dam, for offering me the opportunity to teach the Foundation of Calculus and Calculus for Science. This opportunity allowed me to leverage my background and collaborate effectively with them.

I extend my gratitude to Curtin's Industry Exchange team, with a special acknowledgment to Nick Schools and all the staff at Assuro, particularly their CEO, Leon Weston. Their collective effort gave me the invaluable opportunity to engage in two industry internships during my PhD candidacy. I am also deeply thankful for the guidance and support of Dr. Felix Chan, who generously shared his expertise despite his heavy academic workload. The summer internships facilitated through Industry Exchange and Assuro were transformative for my professional growth. Working on substantive corporate projects was both challenging and immensely fulfilling.

Many other academics deserve significant gratitude: Dr. Muammar Wali and Dr. Zhiyue Sun provided invaluable software and technical support during my PhD studies. Dr. Hiroaki Suenaga and Dr. Helen Cabalu were reviewers for my Milestone 1 and 2 presentations, offering crucial feedback and suggestions for my thesis work. Dr. Tianpei Luo chaired my Milestone 3 presentation, contributing to its success.

I am deeply thankful for the financial assistance provided by the Graduate Research School (GRS), particularly the Stipend Scholarship, which helped me cover my living expenses throughout my doctoral studies. I extend special thanks to Doreen Phan from

the GRS team, whose great support was essential in successfully securing additional research funding for the international dissemination and engagement of my research.

Lastly, I express my warmest thanks to my father, Antonio Loureiro, whose guidance, wisdom, and support have been constant pillars in my life. I am equally thankful to my mother, Viviane Loureiro, and my children, Gabriel Loureiro and Maya Cabral. My family has been my source of strength, commitment, and focus. I attribute everything I am and all that I have accomplished to my family's emotional support and encouragement. To each of you, I dedicate this achievement. Thank you.

	Conception and Design	Acquisition of Data and Method	Data Conditioning and Manipulation	Analysis and Statistical Method	Interpretation and Discussion
<b>Jose Loureiro</b>	✓	✓	✓	✓	✓
<b>Acknowledgement:</b> I acknowledge that these represent my contribution to the above research output and I have approved the final version. Signature:					
<b>Julian Inchauspe</b>	✓				✓
<b>Acknowledgement:</b> I acknowledge that these represent my contribution to the above research output and I have approved the final version. Signature:					
<b>Roberto Aguilera (Aug-2020 to Jun-2022)</b>				✓	✓
<b>Acknowledgement:</b> I acknowledge that these represent my contribution to the above research output and I have approved the final version. Signature:					
<b>Helen Cabalu (Jun-2022 to May-2024)</b>				✓	✓
<b>Acknowledgement:</b> I acknowledge that these represent my contribution to the above research output and I have approved the final version. Signature:					

## **ACKNOWLEDGEMENT OF COUNTRY**

We acknowledge that Curtin University works across hundreds of traditional lands and custodial groups in Australia, and with First Nations people around the globe. We wish to pay our deepest respects to their ancestors and members of their communities, past, present, and to their emerging leaders. Our passion and commitment to work with all Australians and peoples from across the world, including our First Nations peoples are at the core of the work we do, reflective of our institutions' values and commitment to our role as leaders in the Reconciliation space in Australia.



# TABLE OF CONTENTS

<b>Declaration</b> .....	<b>i</b>
<b>Abstract</b> .....	<b>ii</b>
<b>Acknowledgements</b> .....	<b>iv</b>
<b>Acknowledgement of country</b> .....	<b>vii</b>
<b>Table of Contents</b> .....	<b>viii</b>
<b>List of Tables</b> .....	<b>xiii</b>
<b>List of Figures</b> .....	<b>xvi</b>
<b>CHAPTER 1 Introduction</b> .....	<b>1</b>
1.1 Chapter Outline .....	1
1.2 Context and Motivation for Research .....	1
1.3 Research Questions and Objectives .....	5
1.4 Research Significance .....	6
<b>CHAPTER 2 Integration of the Global Natural Gas Market: A Literature Review</b> .....	<b>9</b>
2.1 Introduction.....	9
2.2 Early Investigations into the Integration of the Global Natural Gas Market (2005 - 2014).....	9
2.3 Recent Investigations into the Integration of the Global Natural Gas Market (2015 - 2020).....	11
2.4 Conclusions and Summary of Literature Review .....	15
<b>CHAPTER 3 Preliminary Overview of the Natural Gas Markets</b> .....	<b>21</b>
3.1 Introduction.....	21
3.2 Overview of the Crude Oil and Gas Price Relationship .....	22
3.3 The Global Gas Supply Chain and its Unique Characteristics .....	23
3.4 Overview of the Natural Gas Market in North America.....	26
3.5 Overview of the Natural Gas Market in Europe .....	27

3.6	Overview of the Natural Gas Market in Asia.....	29
3.7	Concluding Remarks.....	31
<b>CHAPTER 4 Data Description and Time Series Properties of the Thesis Datasets</b>		<b>34</b>
.....		
4.1	Introduction.....	34
4.2	Data Description.....	36
4.3	Unit Root Tests and Stationarity of Time Series .....	41
4.3.1	Stationarity and Unit Root Tests Without Structural Breaks .....	42
4.3.2	Stationarity and Unit Root Tests with Structural Break.....	47
4.3.2.1	Perron’s Model for Unit Root Tests with Structural Break .....	47
4.3.2.2	Zivot and Andrews’ Model for Unit Root Tests with Structural Break.....	50
4.4	Implementation of the Unit Root Tests to the Thesis Dataset .....	52
4.4.1	Graphical Representation of Thesis Data.....	53
4.4.2	Descriptive Statistics of Thesis Data.....	54
4.4.3	Empirical Results of the Unit Root Tests Without Structural Break .....	55
4.4.4	Empirical Results of the Zivot and Andrews (1992) Unit Root Test with Structural Break .....	57
4.5	Concluding Remarks.....	61
<b>CHAPTER 5 Analysing Natural Gas Price Convergence: The Phillips-Sul Relative Convergence Test .....</b>		<b>63</b>
5.1	Introduction.....	63
5.2	Fundamentals of Gas Arbitrage Activities and Key Hypotheses.....	65
5.3	Relative Convergence Testing and Clustering Methodology .....	73
5.4	Phillips and Sul (2007, 2009) Convergence Test Results.....	76
5.4.1	Whole sample convergence clustering.....	76
5.4.2	The Pre-GFC Sub-period .....	78
5.4.3	The 2010-2020 Sub-period with New Gas Price Benchmarks .....	80

5.5	Robustness Check: Kalman Filter (Law of One Price Estimates) .....	82
5.6	Discussion of the Results .....	86
5.7	Concluding Remarks .....	88
<b>CHAPTER 6 Review of Econometric Methods.....</b>		<b>89</b>
6.1	Introduction .....	89
6.2	Introduction to Multivariate Time Series Models .....	90
6.2.1	Simultaneous Equation Method .....	90
6.2.2	ARIMA Method .....	92
6.2.3	Error Correction Model.....	92
6.2.4	Vector Autoregression Method .....	93
6.3	Vector Autoregression Model .....	93
6.3.1	Optimal VAR Lag Length Selection.....	95
6.3.2	Granger Causality Tests .....	96
6.3.3	Impulse Responses in a VAR Model .....	97
6.3.4	Decomposition of Forecast Error Variance.....	99
6.4	Cointegration Techniques .....	101
6.5	ARDL Cointegration Model .....	104
6.6	A Non-linear ARDL (NARDL) Approach to Cointegration .....	108
6.7	Global Economic Conditions Indicator.....	109
6.7.1	The Global Economic Conditions Indicator: A Multi-Dimensional Approach .....	110
6.7.2	Incorporating the GECON Indicator into Econometric Models of this Thesis .....	111
6.8	Concluding Remarks .....	111
<b>CHAPTER 7 Bivariate Linear ARDL Models .....</b>		<b>113</b>
7.1	Introduction .....	113
7.2	Empirical Results of the Bivariate Linear ARDL Models .....	113
7.2.1	Unit Root and Structural Break Analysis.....	115

7.2.2	Bivariate Linear ARDL Models Applied to Sample 1 .....	116
7.2.3	Bivariate Linear ARDL Models Applied to Sample 2 .....	123
7.2.4	Price Discovery Assessment - Leading and Lagging Prices .....	133
7.3	Concluding Remarks .....	137
<b>CHAPTER 8 Bivariate Asymmetric ARDL Models.....</b>		<b>140</b>
8.1	Introduction .....	140
8.2	Empirical Results of the Bivariate NARDL Models.....	141
8.2.1	Bivariate NARDL Models Applied to Sample 1 .....	141
8.2.2	Bivariate NARDL Models Applied to Sample 2 .....	152
8.3	Concluding Remarks .....	165
<b>CHAPTER 9 VAR Assessment of Causality Relationships Between Natural Gas and Oil Prices.....</b>		<b>168</b>
9.1	Introduction .....	168
9.2	Empirical Results of the VAR Model Applied to Sample 2 .....	169
9.2.1	Optimal VAR Lag Length.....	171
9.2.2	Toda and Yamamoto (1995) Causality Test .....	173
9.2.3	VAR Analysis of Impulse Responses .....	177
9.2.4	VAR Analysis of Forecast Error Variance Decomposition .....	186
9.3	Concluding Remarks .....	189
<b>CHAPTER 10 Price and Income Elasticity of Natural Gas Demand: A Multi-Sector Analysis in Japan and Korea.....</b>		<b>192</b>
10.1	Introduction .....	192
10.2	Overview of the Natural Gas Markets in Japan and South Korea .....	195
10.2.1	Japan.....	195
10.2.2	South Korea.....	196
10.3	Literature Review on the Energy Demand Equation .....	197
10.4	Methodology and Data .....	199
10.4.1	Model Specification .....	199

10.4.2	Variables and Data Sources .....	200
10.5	Empirical Results and Discussion.....	202
10.5.1	Unit Root Test.....	202
10.5.2	Analysis of the Natural Gas Demand in the Industrial Sector: Japan and Korea .....	203
10.5.3	Analysis of the Natural gas Demand in Residential Sector: Japan and Korea .....	207
10.6	Concluding Remarks.....	210
<b>CHAPTER 11</b>	<b>Conclusions and Policy Implications .....</b>	<b>213</b>
11.1	Introduction.....	213
11.2	Discussion of the Results and Policy Implications .....	213
<b>REFERENCES</b>	.....	<b>221</b>

## LIST OF TABLES

Table 1.1	International Natural Gas Trade Statistics by Region in 2021 .....	2
Table 2.1	Summary of Literature Review Findings (2005 - 2014) .....	16
Table 2.2	Summary of Literature Review Findings (2015 - 2020) .....	17
Table 4.1	Correlation Matrix of Price Time Series from 2001M01 to 2020M02. .....	40
Table 4.2	Correlation Matrix of Price Time Series from 2010M07 to 2020M02. .....	41
Table 4.3	Descriptive Statistics of the Dataset with Sample Period between 2001M01 and 2020M02. ....	54
Table 4.4	Descriptive Statistics of the Dataset with Sample Period between 2010M07 and 2020M02. ....	54
Table 4.5	Standard Unit Root Testing Results for the 2001-2020 Data Sample. .....	56
Table 4.6	Standard Unit Root Testing Results for the 2010-2020 Data Sample. .....	56
Table 4.7	Zivot and Andrews (1992) Unit Root Testing Results for the 2001- 2020 Data Sample in Levels.....	58
Table 4.8	Zivot and Andrews (1992) Unit Root Testing Results for the 2001- 2020 Data Sample in First Differences. ....	59
Table 4.9	Zivot and Andrews (1992) Unit Root Testing Results for the 2010- 2020 Data Sample in Levels.....	59
Table 4.10	Zivot and Andrews (1992) Unit Root Testing Results for the 2010- 2020 Data Sample in First Differences. ....	60
Table 4.11	Summary of the Zivot and Andrews (1992) Unit Root Testing Results for the 2001-2020 Data Sample.....	60
Table 4.12	Summary of the Zivot and Andrews (1992) Unit Root Testing Results for the 2010-2020 Data Sample.....	61
Table 5.1	Variables Description and Sourcing .....	71
Table 5.2	Descriptive Statistics for Monthly Regional Price Variables.....	72
Table 5.3	Convergence Test Results for 01:2001-02:2020.....	77

Table 5.4	Convergence Test Results for 01:2001-08:2008. ....	79
Table 5.5	Convergence Test Results for 01:2010-02:2020. ....	80
Table 7.1	Summary of Sample 1 Dataset (Jan 2001 – Feb 2020). ....	114
Table 7.2	Summary of Sample 2 Dataset (Jul 2010 – Feb 2020). ....	114
Table 7.3	Linear ARDL Bivariate Models Applied to Sample 1 – Cointegration and ECM Long-Run Estimates. ....	117
Table 7.4	Linear ARDL Models with GECON Indicator as Exogenous Variable Applied to Sample 1 – Cointegration and ECM Long-Run Estimates. .....	119
Table 7.5	Linear ARDL Bivariate Models Applied to Sample 1 – ECM Short- Run Estimates. ....	121
Table 7.6	AIC-Augmented Linear ARDL Bivariate Models Applied to Sample 1 – Short-Run Estimates. ....	122
Table 7.7	Linear ARDL Bivariate Models Applied to Sample 2 – Cointegration and ECM Long-Run Estimates. ....	123
Table 7.8	Linear ARDL Models with GECON Indicator as Exogenous Variable Applied to Sample 2 – Cointegration and ECM Long-Run Estimates. .....	127
Table 7.9	Linear ARDL Bivariate Models Applied to Sample 2 – ECM Short- Run Estimates. ....	130
Table 7.10	AIC-Augmented Linear ARDL Bivariate Models Applied to Sample 2 – Short-Run Estimates. ....	131
Table 7.11	Summary of the Speed of Adjustment Coefficients of Cointegrated Bivariate Models in Sample 1. ....	134
Table 7.12	Summary of the Speed of Adjustment Coefficients of Cointegrated Bivariate Models in Sample 2. ....	135
Table 7.13	Price Discovery Assessment of Cointegrating Pairs in Sample 2. ...	136
Table 7.14	Summary of the Linear ARDL Bivariate Models Results. ....	137
Table 8.1	Asymmetric ARDL (NARDL) Bivariate Models Applied to Sample 1 - Cointegration and ECM Long-Run Estimates. ....	142
Table 8.2	Asymmetric ARDL (NARDL) Models with GECON Indicator as Exogenous Variable Applied to Sample 1 - Cointegration and ECM Long-Run Estimates. ....	144

Table 8.3	Welch's t-test and Welch-Satterthwaite results for null hypothesis, $H_0: \varphi_2 + = \varphi_3 -$ (Sample 1). .....	149
Table 8.4	Asymmetric ARDL (NARDL) Bivariate Models Applied to Sample 1 - ECM Short-Run Estimates. ....	149
Table 8.5	AIC-Augmented Asymmetric ARDL (NARDL) Bivariate Models Applied to Sample 1 – Short-Run Estimates.....	151
Table 8.6	Asymmetric ARDL (NARDL) Bivariate Models Applied to Sample 2 - Cointegration and ECM Long-Run Estimates. ....	152
Table 8.7	Asymmetric ARDL (NARDL) Models with GECON Indicator as Exogenous Variable Applied to Sample 2 - Cointegration and ECM Long-Run Estimates. ....	156
Table 8.8	Welch's t-test and Welch-Satterthwaite results for null hypothesis, $H_0: \varphi_2 + = \varphi_3 -$ (Sample 2). ....	160
Table 8.9	Asymmetric ARDL (NARDL) Bivariate Models Applied to Sample 2 - Cointegration and ECM Short-Run Estimates. ....	161
Table 8.10	AIC-Augmented Asymmetric ARDL (NARDL) Bivariate Models Applied to Sample 2 - Short-Run Estimates.....	162
Table 8.11	Summary of the Asymmetric ARDL (NARDL) Bivariate Models Results. ....	166
Table 9.1	Optimal Lag Length Selection Using Information Criterion. ....	171
Table 9.2	LM Test Results for Serial Correlation in the Residuals.....	172
Table 9.3	Toda and Yamamoto Granger Causality Test in VAR model.....	173
Table 9.4	Toda and Yamamoto Granger Causality Test in VAR model with GECON Indicator as Exogenous Variable. ....	175
Table 9.5	Forecast Error Variance Decomposition from the VAR Model.....	187
Table 10.1	Variables' Description and Abbreviation.....	202
Table 10.2	Results of Unit Root Tests.....	203
Table 10.3	Coefficient Estimation of Natural Gas Demand in Japan and Korea: Industrial Sector.....	205
Table 10.4	Coefficient Estimation of Natural Gas Demand in Japan and Korea: Residential Sector.....	208



## LIST OF FIGURES

Figure 1.1	Share of Spot and Short-term vs. Total LNG Trade (MTPA/%). Source: GIIGNL (2023). .....	4
Figure 3.1	Natural Gas Major Trade Movements in BCM (Pipeline and LNG) in 2009. Source: BP (2010) .....	32
Figure 3.2	Natural Gas Major Trade Movements in BCM (Pipeline and LNG) in 2021. Source: BP (2022) .....	33
Figure 4.1	Global Gas and Brent Prices: Jan 2018 - Dec 2022. Source: Heather (2023) .....	35
Figure 4.2	Total Traded Volumes in European Gas Hubs in 2022 (TWh). Source: Heather (2023).....	35
Figure 4.3	Price Time Series in US\$/MMBtu from 2001M01 to 2020M02. ....	40
Figure 4.4	Price Time Series in US\$/MMBtu from 2010M07 to 2020M02. ....	40
Figure 4.5	Graphical Representation of Time Series in Natural Logs.....	53
Figure 5.1	Arbitrage in International Markets and Long-run Dynamics Implications. ....	66
Figure 5.2	Schematic Representation of the Price-Levelling Arbitrage Hypothesis. Source: Data from BP (2020) for 2019. ....	71
Figure 5.3	Relative Convergence Coefficients for the 2001-2019 Panel Data of 6 Stacked Prices.....	78
Figure 5.4	Relative Convergence Coefficients for the 2010-2019 Panel Data of 8 Stacked Prices.....	81
Figure 5.5	Traditional Gas Price Benchmarks and Oil Price: Kalman-Filter Estimates with 2-Standard Deviation Prediction Bands.....	85
Figure 5.6	The New Gas Price Benchmarks (AS and TTF): Kalman-Filter Estimates with 2-Standard Deviation Prediction Bands.....	86
Figure 7.1	Flowchart of the ARDL Approach. ....	115
Figure 7.2	Diagram of the Linear Long-Run Causality Directions of Sample 1. .....	120
Figure 7.3	Diagram of the Linear Short-Run Causality Directions of Sample 1. .....	123

Figure 7.4	Diagram of the Linear Long-Run Causality Directions of Sample 2. .....	129
Figure 7.5	Diagram of the Linear Short-Run Causality Directions of Sample 2. .....	133
Figure 8.1	Diagram of the Asymmetric Long-Run Causality Directions of Sample 1. ....	146
Figure 8.2	Diagram of the Positive Asymmetric Long-Run Causality Directions of Sample 1. ....	147
Figure 8.3	Diagram of the Negative Asymmetric Long-Run Causality Directions of Sample 1. ....	147
Figure 8.4	Diagram of the NARDL Short-Run Causality Directions of Sample 1. .....	152
Figure 8.5	Diagram of the Asymmetric Long-Run Causality Directions of Sample 2. ....	158
Figure 8.6	Diagram of the Positively Asymmetric Long-Run Causality Directions of Sample 2. ....	159
Figure 8.7	Diagram of the Negatively Asymmetric Long-Run Causality Directions of Sample 2. ....	159
Figure 8.8	Diagram of the NARDL Short-Run Causality Directions of Sample 2. .....	165
Figure 9.1	Causality Detection Trap in Bivariate Models. ....	169
Figure 9.2	Diagram of the Granger Causality Directions Based on the T-Y Causality Test Applies to Sample 2. ....	177
Figure 9.3	Responses to a Henry Hub (HH) gas price shock. ....	179
Figure 9.4	Responses to a National Balancing Point (NBP) gas price shock. ....	180
Figure 9.5	Responses to a Title Transfer Facility (TTF) gas price shock. ....	181
Figure 9.6	Responses to a Russian (RUS) gas price shock. ....	182
Figure 9.7	Responses to a Japan's Monthly Average LNG Import (JPN) gas price shock. ....	183
Figure 9.8	Responses to a shock in LNG prices imported into Northeast Asia (ALNG). ....	184
Figure 9.9	Responses to a Brent crude oil (OIL) price shock. ....	185

Figure 11.1 European Gas Hub and Asian LNG Prices 2020-22 (\$/MMBtu).  
Source: Stern (2023).....217

# CHAPTER 1

## INTRODUCTION

### 1.1 Chapter Outline

This chapter is organised into five sections, each serving a specific purpose in introducing the thesis framework. Section 1.2 presents the background and motivation that underlie the broader context of the thesis. Section 1.3 presents the main research questions and objectives, which define the specific boundaries of the study's investigation. Section 1.4 explores the contributions to the existing literature, explaining its unique value in academic discourse, while Section 1.5 explains the organisation and structure of the thesis.

### 1.2 Context and Motivation for Research

Natural gas is an important alternative for the growing global energy demand. Its appeal lies in two main factors. First, it offers a cleaner environmental profile than conventional energy sources like oil and coal, which are major contributors to greenhouse gas emissions. In the transition to renewable energy and carbon emission reduction, natural gas plays a crucial role by providing energy resilience and affordability while facilitating the decommissioning of highly polluting assets like coal and oil power plants. Investments in lower-emission fuel production, addressing methane emissions, and electrifying oil and gas operations are essential for an orderly transition. Given recent challenges like the war in Ukraine and the COVID-19 pandemic, the urgency for energy resilience has heightened, underscoring the importance of natural gas as a transitional fuel.

McKinsey & Company (2022) highlights uncertainties despite progress at COP27 towards global cooperation on emissions reduction targets. To bridge the gap between current trajectories and desired commitments, substantial scaling up of renewable energy capacity is imperative, with McKinsey's analysis indicating a need for nearly tripling annual solar and wind installations over the next decade. In this scenario, natural gas emerges as a crucial element in the energy transition, facilitating the move

towards renewable energy sources while guaranteeing a dependable and cost-effective energy supply.

Although natural gas presents itself as a competitive option and complements emerging renewable energies, a notable challenge arises from the regionalisation of its markets. This regionalisation is primarily a result of substantial investments in natural gas infrastructure and transportation systems. The intricate network of pipelines, terminals, and transportation routes poses barriers to entry for a globally integrated natural gas market. The strategic positioning of liquefied natural gas (LNG) facilities, crucial for facilitating natural gas transportation over long distances via ships, introduces an additional layer of complexity to the investment landscape, reinforcing the regional nature of natural gas markets.

At present, the natural gas trade is primarily divided into three main trading regions: North America, Europe, and Asia. Each region functions within its unique market dynamics, influenced by geographical, infrastructural, and geopolitical factors, domestic natural gas demand and production capacities. The availability and scale of domestic natural gas production significantly shape the market dynamics in these regions. Geographical factors play a role in determining the accessibility and cost-effectiveness of natural gas supply. At the same time, the development of infrastructure, including LNG facilities and pipeline networks, further defines the regional landscape. Table 1.1 highlights the significance of each region in the global natural gas trade.

Table 1.1<sup>1</sup> International Natural Gas Trade Statistics by Region in 2021

<b>Natural Gas Inter-Regional Trade</b>		
<b>United States</b>	<b>2021 (BCM)</b>	<b>World Share 2021 (%)</b>
Total imports	76.5	7.5
Total exports	179.3	17.5
<b>Europe</b>	<b>2021 (BCM)</b>	<b>World Share 2021 (%)</b>
Total imports	341	33.4
Total exports	3.8	0.4
<b>Asia (China + OECD)</b>	<b>2021 (BCM)</b>	<b>World Share 2021 (%)</b>

<sup>1</sup> Table 1.1 continues on the next page.

<b>Asia (China + OECD)</b>	<b>2021 (BCM)</b>	<b>World Share 2021 (%)</b>
Total imports	332.9	32.6
Total exports	108.3	10.6
<b>LNG Trade</b>		
<b>United States</b>	<b>2021 (BCM)</b>	<b>World Share 2021 (%)</b>
Total imports	0.6	0.1
Total exports	95	18.4
<b>Europe</b>	<b>2021 (BCM)</b>	<b>World Share 2021 (%)</b>
Total imports	108.2	21
Total exports	3.8	0.7
<b>Asia Pacific</b>	<b>2021 (BCM)</b>	<b>World Share 2021 (%)</b>
Total imports	371.8	72
Total exports	176.3	34.2

Source: BP Statistical Review of World Energy 2022.

Aguilera, Inchauspe and Ripple (2014) explored regional market structures within the natural gas industry. North America is described as competitive, with abundant shale gas supply maintaining prices at relatively low levels, averaging around US\$2 to \$4 per million British thermal units (MMBtu) for much of the 2010–20 period. In contrast, Europe, traditionally characterised as an oligopolistic gas market with few sellers and many buyers, has experienced a shift toward competitiveness in recent years. During the same period, prices in Europe have hovered around \$10. Asia, on the other hand, has conventionally been depicted as a bilateral monopoly with few sellers and few buyers. Prices in the Asian natural gas market have averaged approximately \$20 over the past decade.

However, over the past decade, the natural gas industry has undergone a rapid and transformative expansion, reshaping the dynamics of international gas trading. A significant driver of this expansion lies in the substantial growth of LNG liquefaction capacity and trade. This surge in LNG trade has contributed to faster transactions and increased the number of players in the natural gas trade, which played a role in an attempt to diminish gas price differentials.

Fulwood (2023) highlights a significant transformation in the industry's dynamics, transitioning from traditional oil indexation to adopting gas-on-gas (spot and short-

term) pricing mechanisms. Oil indexation is a gas pricing mechanism that correlates with other fuel prices, primarily oil or refined products. On the other hand, gas-to-gas pricing involves indexation to spot prices that mirror the fluctuations in supply and demand for natural gas within a market. This increased contractual flexibility reflects a departure from the previously entrenched link between gas and oil prices and long-term contracts. This change in pricing structures has been particularly pronounced in Europe and, to a lesser extent, Asia.

Furthermore, the emergence of gas hubs and spot price benchmarks in Europe and Asia have added dynamism to global gas trading. The acceleration of LNG trade, the departure from oil indexation, and the establishment of these benchmarks have diversified trading options and marked an evolution in the industry's overall trading and pricing mechanisms. Figure 1.1 illustrates the evolution of the total LNG trade in MTPA (left axis) versus the proportion of spot and short-term transactions in the total LNG purchase agreements (right axis) over the last decade.

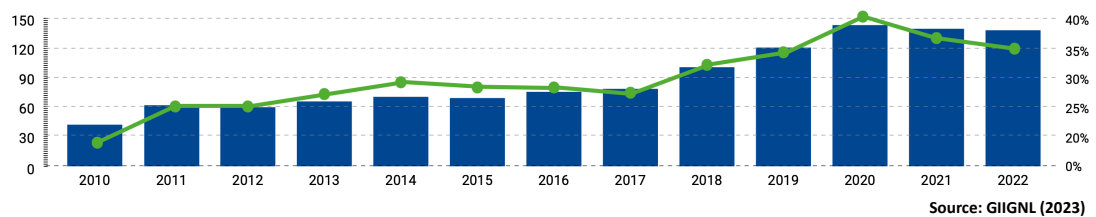


Figure 1.1 Share of Spot and Short-term vs. Total LNG Trade (MTPA/%). Source: GIIGNL (2023).

The continuous changes could have potential impacts on the gas market. If the LNG market strengthens, with lower transaction costs and increased flexibility, it could lead to a significant transformation, creating a more globally integrated gas market. The issue of market integration is fundamentally structural and requires a long-term perspective. Consequently, it prompts questions about the current stability relationships of the natural gas markets, specifically between oil and gas prices and within gas prices themselves. The past decade underscores this point, with oil prices significantly diverging from gas prices after the 2008 financial crisis and the US shale gas revolution. Similar trends were observed in the US during the eighties and nineties. Since the 1990s, the increasing prominence of combined cycle gas turbines and more

recent innovations like horizontal drilling and rock fracturing techniques may have significantly changed the relationship between natural gas and crude oil over time.

Given the critical importance of the global integration of natural gas markets, there is an increasing need to investigate the current dynamics among international natural gas prices. This proposed research aims to comprehensively assess the spot and long-term contract gas prices within the three main trading regions. The research uses rigorous econometric analysis involving convergence, cointegration, and causality assessments to identify gas prices that exhibit greater reactivity to market fundamentals and other gas prices, contrasting those that still maintain a significant influence from oil price indexation. Another crucial aspect of this research is identifying market leaders shaping the natural gas price formation.

This thesis also examines factors that influence the demand for natural gas. It looks closely at Japan and Korea, significant importers of natural gas in Asia. It will analyse economic factors that impact their natural gas markets more closely. Through this detailed analysis, the research aims to provide valuable insights into how natural gas markets are changing globally and in specific regions. This helps us understand both the major trends and region-specific characteristics.

### **1.3 Research Questions and Objectives**

This thesis aims to explore the subsequent overarching research questions:

1. To what degree have geographically distant natural gas markets achieved integration due to developments in the LNG industry over the past decade?
2. Which natural gas markets exhibit higher reactivity to market fundamentals, and conversely, which markets are more significantly influenced by oil price indexation, and what factors contribute to these distinctions in pricing mechanisms?
3. How does price discovery occur in the global natural gas markets, specifically focusing on the Asian and European markets, and what can be discerned about leading and lagging markets in this context?
4. How can econometric models be utilised to assess the factors influencing natural gas demand in Japan and Korea, two significant Asian gas importers?



Based on the outlined research questions, this research expects to achieve the following objectives:

1. Using unique econometric approaches and considering new natural gas price benchmarks, assess the extent of integration in geographically distant natural gas markets resulting from advancements in the natural gas industry over the last decade.
2. Identify natural gas markets that demonstrate greater sensitivity to market fundamentals, leading to price formation through gas-on-gas mechanisms and aligned with gas hubs' spot prices. Concurrently, assess the recent influence of oil price indexation on natural gas prices, providing an up-to-date analysis and validating or challenging existing findings from the literature.
3. Examine the price discovery mechanisms in global natural gas markets, identifying leading and lagging markets within this framework. Consequently, propose policy implications related to the co-movements of natural gas prices.
4. Build a comprehensive model to examine the elasticities of natural gas demand in key natural gas importing nations, comparing both industrial and residential sectors. Specifically, the research will evaluate the price and income elasticities of natural gas demand in Japan and Korea while also analysing the substitutive or complementary dynamics between natural gas and other energy sources.

#### **1.4 Research Significance**

Given the significant transformations in the global natural gas market in recent years, this thesis proposes an enhanced and up-to-date sampling framework to evaluate the current dynamics between interregional gas markets. This includes examining recently established gas price benchmarks in Asia and Europe. By doing so, the research not only updates existing literature but also addresses a gap in the current body of knowledge concerning these recent developments. Methodologically, the study contributes to the existing field of research by applying a rigorous set of state-of-the-art econometrics tools.

Assessing the degree of integration in the global natural gas market and the current role of the traditional oil indexation price mechanism holds great importance for industry stakeholders and policymakers. A more integrated global market potentially reduces price risk for importing countries by fostering diversification in gas imports and decreasing dependence on major gas suppliers. Moreover, a globally integrated, market-oriented gas market will likely draw in a broader range of market participants and enhance liquidity in spot gas trading compared to long-term contracts, often tied to oil prices. As a result, it is expected that the price of natural gas will reflect the combined expectations of all participants in the market, considering both current and future supply and demand conditions. This shift carries significant implications for businesses and consumers. Additionally, it provides policymakers valuable insights into the commoditisation of gas and the potential expansion of LNG trade, facilitating the reduction of greenhouse gas emissions and the transition from coal to natural gas in the global energy mix. Also, it provides the groundwork for developing effective energy supply-demand policies and addressing risks related to market price fluctuations.

As a complementary study, this research extends beyond existing literature focused on investigating the elasticity of natural gas demand, particularly in Europe and China. The current study shifts its focus to other key importer markets, Japan and Korea, the second and third-largest consumers of LNG globally, offering a more comprehensive perspective on assessing natural gas demand.

The thesis is structured across chapters to address the research questions systematically. Chapter 2 reviews the relevant literature and theoretical perspectives to contextualise the research, while Chapter 3 provides an overview of the main natural gas trading markets – North America, Europe, and Asia. Chapter 4 outlines the main gas price datasets used throughout the thesis and assesses the main properties of the thesis' data, such as the stationarity and structural breaks, to ensure accurate modelling in later analysis. Chapter 5 uses Phillips and Sul's growth convergence testing and clustering algorithms to examine natural gas price relationships. Chapter 6 provides a theoretical foundation for the methods used for the subsequent chapters. Chapter 7 presents the empirical results from implementing bivariate Autoregressive Distributed Lag (ARDL) models along with the bounds cointegration test, while Chapter 8 applies

the Non-linear ARDL (NARDL) model, also assessing cointegration in our main thesis' data. Chapter 9 assesses the dynamic and causality relationships among gas prices by proposing a unified VAR composed of all gas prices in our dataset. Chapter 10 assesses natural gas demand in Japan and Korea, revealing distinct price and income elasticities across industrial and residential sectors. Lastly, Chapter 11 concludes the thesis with a summary of key findings, policy implications, and directions for future research.

## **CHAPTER 2**

# **INTEGRATION OF THE GLOBAL NATURAL GAS MARKET: A LITERATURE REVIEW**

### **2.1 Introduction**

This chapter thoroughly examines existing literature that investigates the dynamics between natural gas prices across three primary gas trading regions. It is crucial to note that before the ground-breaking study by Siliverstovs et al. (2005), research on the relationship between natural gas prices predominantly concentrated on regional assessments. The focus was on assessing the degree of market liberalisation in each region and the impact of oil prices on gas price formation. Initially centred on the North American market (late 1990s), these studies shifted their attention to the European market in the 2000s and early 2010s, coinciding with its gas market liberalisation process. The literature is notably deficient in a regional evaluation of the Asian market due to significant constraints on market liberalisation persisting into recent periods.

This chapter, however, refrains from an in-depth review of these regional assessments. Instead, a brief discussion will be included when presenting an overview of each regional market in the subsequent chapter. This decision aligns with the thesis's primary focus on examining the relationship of natural gas prices on an interregional level.

### **2.2 Early Investigations into the Integration of the Global Natural Gas Market (2005 - 2014)**

Siliverstovs et al. (2005) assess the degree of integration in natural gas markets across Europe, North America, and Japan from 1993 to 2004. It utilises principal components analysis and the Johansen likelihood-based cointegration procedure to evaluate the connection between international gas market prices and their correlation with oil prices. The study's primary findings reveal the successful integration of regional gas prices in North America, driven by factors like gas-to-gas competition and regulatory

changes. European and Japanese gas markets show integration, attributed to similar contract structures and oil-price indexation. The study highlights the need for further integration between natural gas markets across the Atlantic, underscoring distinct market dynamics and limited arbitrage opportunities during the observed period.

Brown and Yücel (2009) and Neumann (2009) addressed analogous questions concerning the integration of interregional natural gas markets and the impact of oil indexation. The focus was on assessing the impact of increased LNG trade, particularly its import into the North American market.

Brown and Yücel (2009) analyse the transmission of natural gas price shocks across the Atlantic, examining the influence of oil prices on this transmission. Using data from 1997 to 2008, their study employs causality tests on North American and European natural gas prices. Bivariate tests reveal bidirectional causality between the US Henry Hub and the UK's National Balancing Point (NBP) in natural gas prices, indicating coordinated movements. Multivariate models emphasise the role of crude oil prices in coordinating natural gas prices across the Atlantic. LNG pricing reinforces the link between European crude oil and natural gas prices.

Neumann (2009) extends the analysis of the North American and European gas markets from 1999 to 2008. Utilising the Kalman Filter, the study indicates a growing convergence in spot prices across transoceanic markets. This convergence is attributed primarily to LNG market arbitrage, with a lesser impact coming from oil prices, contrary to the findings of Brown and Yücel (2009).

Erdös (2012) examines the integration between gas prices in the UK and the US and utilises spot gas prices from the Henry Hub and the National Balancing Point, along with WTI crude oil prices. It models data from 1997 to 2011 using vector error correction models to examine the relationship between oil and natural gas prices. The models' estimates indicate the decoupling of US natural gas prices from European gas and crude oil prices after 2009. However, US gas and oil prices demonstrate long-term integration in the preceding period (1997-2008). The author explains that the absence of liquefying and export capacities in the US has constrained interregional price arbitrage since 2009, preventing the flow of oversupply caused by the commencement of shale production. This results in lower gas prices from the US to Europe and

maintained the decoupling of North American gas prices from oil-integrated prices in Europe and Asia.

Nick and Tischler (2014) conduct a pairwise price convergence test on international gas prices, focusing on the convergence of gas prices in the US and the UK from 2000 to 2012. They employ a threshold error correction model with time-varying parameters obtained through the Kalman filter. Although their findings suggest integration in the US and UK markets from 2000 to 2008, the period spanning 2009 to 2012 reveals substantial threshold estimates that underscore challenges beyond justifiable transportation costs that have not been previously considered.

Li, Joyeux, and Ripple (2014) serve as a precursor to the methodology discussed in Chapter 5 of this thesis. Inspired by Neumann's (2009) research, Li, Joyeux, and Ripple (2014) tackle the same question by employing the methodology introduced by Phillips and Sul (2007, 2009). They employ a method that includes examining a set of compiled time series of prices to assess long-term trends that characterise growth convergence groups or clusters. This dataset integrates gas prices from the US, the UK, and three regional Asian markets (Japan, Taiwan, and South Korea). Three notable findings are observed: (1) the three Asian gas prices converge toward a shared trend; (2) there is limited evidence of convergence between Asian and European prices, which is linked to substantial oil indexation in both markets; and (3) the Henry Hub price in North America demonstrates a unique and dynamic pattern.

### **2.3 Recent Investigations into the Integration of the Global Natural Gas Market (2015 - 2020)**

Caporin and Fontini (2017) investigate the impact of the rise of shale gas production in the US on the long-term relationship between oil and the US natural gas price by testing monthly data from 1997 to 2013, accounting for potential breaks. The authors propose an evaluation using a Vector Error Correction Model (VECM), demonstrating that, in the short run, gas production significantly influences the formation of the Henry Hub gas price following the beginning of shale gas production, with the impact of oil price doubling. However, based on the available data, the long-run relationship between WTI oil and Henry Hub is still to be determined.

Aruga (2016) expands on this research by investigating the impact of changes in the US natural gas market structure after the shale gas revolution on the gas markets in Japan and Europe. The study employs the Bai–Perron test to pinpoint the break date associated with the shale gas revolution and assessed through the Johansen cointegration method changes in market linkages among US, Japanese, and European gas markets before and after the determined break date, using data from 1992 to 2012. After the shale gas revolution, the US gas market's ties with Japanese and European markets weakened, suggesting the global impact of increased US gas supply is limited, likely due to constrained exports. Furthermore, this indicates that the influence of the US shale gas revolution remains largely regional.

Employing a Markov regime-switching model, Geng, Ji, and Fan (2016) evaluate how the North American shale gas revolution affects gas prices in Europe and North America from 1998 to 2015. They demonstrate that, although the revolution altered North American natural gas price movements, this shift was minimal. Before the revolution, North American gas prices showed a seasonal pattern, which disappeared afterwards, indicating a decoupling from the crude oil price. However, the relationship between crude oil and gas prices varied around an equilibrium level in Europe. The sudden increase in natural gas output due to the shale gas revolution did not immediately influence the European price, which still followed the crude oil price dynamics.

Utilising a novel systemic time series methodology, Zhang et al. (2018) explores how oil prices and market fundamentals impact gas prices in Japan, the United States, and Germany from 2000 to 2016. Results show that in Japan and Germany, gas prices are less influenced by supply and demand factors than by oil prices, whereas, in the US market, where gas-on-gas competition prevails, these factors play a more prominent role than oil prices. The study suggests that the Asian premium is primarily due to the influence of oil prices through the oil-indexed pricing mechanism rather than inherent market factors. This emphasises the potential to create reference gas prices in Asia, backed by gas-on-gas trading hubs, aiming for better alignment with local market fundamentals and efficient utilisation of gas resources.

Zhang et al. (2018) use a multiple bubble test to compare oil indexation and hub-based pricing, using gas price data from Japan, Europe, and the US from 1982 to 2017. The

research reveals that Japan's oil-indexed natural gas pricing results in higher prices, increased volatility, and more frequent pricing bubbles during crude oil price surges, influenced by increased financialisation and post-2008 uncertainty in natural gas fundamentals. In contrast, the US hub pricing system, more connected to natural gas fundamentals, experienced fewer explosive movements post-2008. Europe reflected a mixed system, exhibiting fewer bubbles than Japan but more than the US, indicating evidence of oil indexation, particularly in fluctuations mirroring Brent crude oil bubble periods.

Mu and Ye (2018) explore the degree of integration over time in spot LNG markets across various regions, including East Asia, Iberia, Northwest Europe, and South America, through time-varying models and price convergence methods, considering a sample period from 2010 to 2015. Employing Phillips and Sul (2007, 2009) approach, the study assesses if market indices post-Fukushima are moving towards perfect competition. Results demonstrate evident price convergence in spot LNG markets and between spot LNG and NBP prices, especially towards the end of the sample period. This suggests a more integrated market may emerge as LNG availability for spot and short-term markets increases, coupled with enhanced infrastructure interconnection.

Chai et al. (2019) focus on the Chinese gas market and examine the recent links between the world's natural gas markets, specifically the Far East, the Middle East, the United States, Europe, China, and Japan, characterising the time-varying characteristics of these relationships. The authors employ a methodology based on the DCC-GARCH- NARDL-ARDL-ECM models and data ranging from 2014 to 2018, intending to capture the impact of China's natural gas market price reforms in 2014. The study reveals that global natural gas markets have limited integration with distinct regional characteristics. Japan, the United States, and Europe are identified as highly representative of the international natural gas market. Despite China's natural gas market being subject to strict residential gas price controls, effective integration with the global market is essential. The examination of risk spillover suggests that China's natural gas market is not aligned with the global market, as differences between domestic and international markets impede the prompt transmission of gas prices.

In a similar investigation, Wang et al. (2020) explore the US gas price and regional natural gas market fundamentals influencing China's natural gas import prices. Using



a VECM framework and data from 2006 to 2019, the results are summarised as follows: the crude oil price primarily influences China's natural gas import prices. Regional market dynamics within China significantly impact import prices, underscoring the success of gas market reforms. Following the 2014 pricing mechanism reforms, crude oil's importance in import prices diminished for both LNG and pipeline natural gas. Post-2014, the influence of Henry Hub natural gas price changes on LNG import prices strengthened, indicating a greater connection between China's LNG market and the US natural gas market. These results are significantly different from those presented by Chai et al. (2019).

A significant investigation of the global natural gas market's integration in recent years was undertaken by Chiappini, Jégourel, and Raymond (2019). The study evaluates the integration levels among European, North American, and Asian markets, considering the growth in global LNG export capacity and the increase of spot and short-term gas trades. Analysing gas price data from 2004 to 2018, the authors employ cointegration tests with one or more structural breaks and threshold error correction models in their methodology. The results show that an increased integration in gas markets has occurred, with stronger connections between gas prices and a weakened link to oil prices. European gas markets show significant integration, with a notable rise in integration between European and American markets post-2014. In Asia, European markets are closely linked with the Japan-Korea Marker (JKM) gas spot price, while the Henry Hub (HH) showed no significant connection. Asymmetric adjustments to long-term equilibrium are observed between HH and the European National Balancing Point (NBP) spot price and between HH and JKM. Adjusting to positive deviations in NBP or JKM prices requires more time than negative ones, a pattern possibly influenced by market arbitrages from exporting nations experiencing a rise in spot transactions.

Finally, Kim and Kim (2019) and Kim et al. (2020) assess the integration of North American, Asian, and European gas markets, focusing on the impact of swing suppliers like Russia and Qatar in the 2010s and Japan's increased gas imports after the Fukushima disaster in 2011. The first study, with data spanning from 2000 to 2017, employs a VECM price discovery framework to analyse leading and lagging markets. Initially, the Asian market took the lead in the natural gas market dynamics, but a shift

occurred after 2011, with the European market assuming the role of price discovery. Following a market shock, the Asian market emerged as the primary adjuster, indicating that changes in the European market have a consequential impact on the Asian market. The second study explores the integration of North American, European, and Asian markets through Engle-Granger cointegration and error correction model on data samples from 2000 to 2019. European and Asian markets display integration, with econometric findings suggesting that Qatar's emergence might have indirectly enhanced integration. Furthermore, the integration between the North American and European markets, evident in the early 2000s, has diminished.

## **2.4 Conclusions and Summary of Literature Review**

This chapter offers a comprehensive overview of the key findings that address the central questions of this thesis. The studies are divided into two periods: 2005-2014, which covers pioneering research on the subject, and 2015-2020, which includes updated investigations reflecting the ongoing development of the natural gas industry and trade. Additionally, the impact of the US shale gas revolution, which began boosting domestic gas production in late 2008 and significantly influenced the global gas industry in the early 2010s, is considered when organizing the literature review based on their data and modelling sample periods.

A synthesis of the key findings from two decades of literature on global gas market integration and the influence of oil prices on gas prices is presented. Notably, conclusions have evolved over time, indicating an increasing integration of natural gas markets, particularly between Asia and Europe. Furthermore, the once robust correlation between oil prices and natural gas markets appears to be decreasing. This comprehensive review underscores the need to continue exploring this topic using diverse econometric models and, critically, with a more recent dataset. Incorporating spot prices from Europe, notably the European Transfer Title Facility (TTF), and Asian spot LNG prices, which have gained prominence, is essential for a thorough understanding. This thesis aims to address this gap in the literature, offering updated results and valuable insights for future studies.

Tables 2.1 and 2.2 provide a concise overview of the literature review outlined in this chapter. The next chapter introduces this thesis's three main natural gas markets,

detailing their historical evolution, market structures, and the oil-gas price relationship. It also presents the principal gas and oil prices used as crucial data for the thesis's econometric models.

Table 2.1<sup>2</sup> Summary of Literature Review Findings (2005 - 2014)

Author(s)	Sample Period	Methodology	Variables (Natural gas and Oil prices)	Main Findings
Silverstovs et al. (2005)	1993 – 2004	Johansen cointegration test	<ul style="list-style-type: none"> <li>• HH</li> <li>• PIPE US</li> <li>• LNG US</li> <li>• LNG EUR</li> <li>• LNG EUR</li> <li>• LNG JPN</li> <li>• BRENT</li> </ul>	<ul style="list-style-type: none"> <li>• European and Japanese gas markets exhibit cointegration, emphasizing oil-indexation price formulas.</li> <li>• Natural gas markets lack integration across the Atlantic, specifically with the North American Henry Hub.</li> <li>• Henry Hub is driven by regional market fundamentals rather than the oil price.</li> </ul>
Brown and Yucel (2009)	1997 – 2008	Johansen Full cointegration test and Error Correction Model	<ul style="list-style-type: none"> <li>• HH</li> <li>• NBP</li> <li>• BRENT</li> <li>• WTI</li> </ul>	<ul style="list-style-type: none"> <li>• Henry Hub and NBP demonstrate bidirectional causality, with both prices adjusting in response to each other.</li> <li>• Crude oil is more significant in price arbitrage across the Atlantic than LNG shipments.</li> </ul>
Neumann (2009)	1999 – 2008	Kalman Filter	<ul style="list-style-type: none"> <li>• HH</li> <li>• NBP</li> <li>• ZEE</li> <li>• WTI</li> <li>• BRENT</li> </ul>	<ul style="list-style-type: none"> <li>• Results indicate integration between Henry Hub and NBP.</li> <li>• Rather than oil, LNG is identified as the key factor for transmitting regional impacts on gas prices.</li> </ul>
Erdös (2012)	1997 – 2011	Vector error correction models (VECM)	<ul style="list-style-type: none"> <li>• HH</li> <li>• NBP</li> <li>• WTI</li> </ul>	<ul style="list-style-type: none"> <li>• US natural gas prices decoupled from European gas and crude oil prices after 2009.</li> <li>• Long-term integration between US gas and oil prices occurred in 1997-2008.</li> <li>• The absence of US LNG export capacities since 2009 maintained decoupling from oil-integrated prices in Europe and Asia.</li> </ul>

<sup>2</sup> Table 2.1 continues on the next page.

Table 2.1 Summary of Literature Review Findings (2005 – 2014), Continued.

Author(s)	Sample Period	Methodology	Variables (Natural gas and Oil prices)	Main Findings
Nick and Tischler (2014)	2000 – 2011	Threshold error correction model and Kalman Filter	<ul style="list-style-type: none"> <li>• HH</li> <li>• NBP</li> </ul>	<ul style="list-style-type: none"> <li>• Integration was observed in US and UK markets from 2000 to 2008.</li> <li>• 2009 to 2012 revealed significant threshold estimates, highlighting challenges beyond justifiable transportation costs.</li> </ul>
Li, Joyeux, and Ripple (2014)	1997 – 2011	Phillips and Sul (2007, 2009) convergence test and Kalman Filter	<ul style="list-style-type: none"> <li>• HH</li> <li>• NBP</li> <li>• LNG JPN</li> <li>• LNG KOR</li> <li>• LNG TWN</li> <li>• BRENT</li> </ul>	<ul style="list-style-type: none"> <li>• Three Asian gas prices show a shared trend.</li> <li>• Limited evidence of convergence between Asian and European prices is linked to substantial oil indexation in both markets.</li> <li>• Henry Hub price in North America exhibits a unique and dynamic pattern.</li> </ul>

Table 2.2<sup>3</sup> Summary of Literature Review Findings (2015 - 2020)

Author(s)	Sample Period	Methodology	Variables (Natural gas and Oil prices)	Main Findings
Caporin and Fontini (2017)	1997 – 2013	Vector error correction model (VECM)	<ul style="list-style-type: none"> <li>• HH</li> <li>• WTI</li> <li>• US Gas Output</li> </ul>	<ul style="list-style-type: none"> <li>• Short-run causality: Shale gas production notably impacted Henry Hub gas prices; oil price had a doubling effect on this relationship.</li> <li>• Long-run causality: The cointegration between WTI oil and Henry Hub in the long term has yet to be confirmed, requiring further analysis.</li> </ul>
Aruga (2016)	1992 – 2012	Johansen Full cointegration test and Bai-Perron test	<ul style="list-style-type: none"> <li>• HH</li> <li>• LNG JPN</li> <li>• RUS GAS</li> <li>• US Gas Output</li> </ul>	<ul style="list-style-type: none"> <li>• The US gas market links to Japan and Europe weakened, signalling a limited global impact due to export constraints.</li> <li>• The US shale gas revolution primarily influences local dynamics, with global markets not significantly affected despite increased supply.</li> </ul>

<sup>3</sup> Table 2.2 continues on the next pages.

Table 2.2 Summary of Literature Review Findings (2015 – 2020), Continued.

Author(s)	Sample Period	Methodology	Variables (Natural gas and Oil prices)	Main Findings
Geng, Ji, and Fan (2016)	1998 – 2015	Markov regime-switching model	<ul style="list-style-type: none"> <li>• HH</li> <li>• NBP</li> <li>• WTI</li> <li>• BRENT</li> </ul>	<ul style="list-style-type: none"> <li>• The shale gas revolution changed the HH gas price, eliminating seasonal patterns and decoupling from crude oil.</li> <li>• The European gas price remained tied to crude oil dynamics despite the shale gas revolution's surge in natural gas output.</li> </ul>
Zhang et al. (2018)	1997 – 2011	VAR approach by Diebold and Yilmaz (2009)	<ul style="list-style-type: none"> <li>• HH</li> <li>• JPN LNG</li> <li>• RUS GAS</li> <li>• WTI</li> <li>• BRENT</li> <li>• DUBAI</li> </ul>	<ul style="list-style-type: none"> <li>• Japan and Germany: Gas prices in these countries are more tied to oil prices than supply and demand, contrasting with the US market.</li> <li>• US: Gas-on-gas competition in the US makes supply and demand factors more influential than oil prices.</li> <li>• The oil prices largely drive the Asian premium through the oil-indexed pricing mechanism rather than inherent market factors.</li> </ul>
Zhang et al. (2018)	1982 – 2017	Multiple bubble test	<ul style="list-style-type: none"> <li>• HH</li> <li>• NBP</li> <li>• LNG JPN</li> <li>• WTI</li> <li>• BRENT</li> </ul>	<ul style="list-style-type: none"> <li>• Oil-indexed natural gas pricing in Japan leads to higher prices, heightened volatility, and more frequent pricing bubbles during crude oil price surges.</li> <li>• The US hub pricing system, tied closely to natural gas fundamentals, experienced fewer explosive movements post-2008.</li> <li>• Europe shows a mixed system with fewer bubbles than Japan but more than the US, suggesting evidence of oil indexation, especially during Brent crude oil bubble periods.</li> </ul>

Table 2.2 Summary of Literature Review Findings (2015 – 2020), Continued.

Author(s)	Sample Period	Methodology	Variables (Natural gas and Oil prices)	Main Findings
Mu and Ye (2018)	2010 – 2015	Phillips and Sul (2007, 2009) convergence test	<ul style="list-style-type: none"> <li>• LNG Asia</li> <li>• LNG NW Europe</li> <li>• LNG Iberia</li> <li>• LNG South America</li> <li>• NBP</li> </ul>	<ul style="list-style-type: none"> <li>• Results show clear price convergence between spot LNG prices and between spot LNG and NBP prices.</li> <li>• The findings suggest the potential emergence of a more integrated market as LNG availability for spot and short-term markets increases, coupled with improved infrastructure interconnection.</li> </ul>
Chai et al. (2019)	2014 – 2018	DCC-GARCH-NARDL-ARDL-ECM	<ul style="list-style-type: none"> <li>• US FOB</li> <li>• Middle East FOB</li> <li>• JPN CIF</li> <li>• Europe CIF</li> <li>• China LNG</li> </ul>	<ul style="list-style-type: none"> <li>• Global natural gas markets show limited integration with distinct regional characteristics.</li> <li>• Japan, the United States, and Europe highly represent the international natural gas market.</li> <li>• The examination of risk spillover indicates that China's natural gas market is not aligned with the global market, hindered by differences between domestic and international markets, impeding prompt price transmission.</li> </ul>
Wang et al. (2020)	2006 – 2019	Vector error correction model (VECM)	<ul style="list-style-type: none"> <li>• HH</li> <li>• China LNG</li> <li>• China Pipe</li> <li>• BRENT</li> <li>• WTI</li> <li>• China Gas Output</li> <li>• China Gas Imports</li> <li>• Climate factors</li> </ul>	<ul style="list-style-type: none"> <li>• Regional market dynamics significantly influence import prices in China, showcasing the success of gas market reforms.</li> <li>• After the 2014 pricing reforms in China, the significance of crude oil in import prices decreased for both LNG and pipeline natural gas prices.</li> <li>• Post-2014, the influence of Henry Hub natural gas price changes on LNG import prices strengthened, highlighting a stronger connection between China's LNG market and the US natural gas market.</li> </ul>

Table 2.2 Summary of Literature Review Findings (2015 – 2020), Continued.

Author(s)	Sample Period	Methodology	Variables (Natural gas and Oil prices)	Main Findings
Chiappini, Jégourel, and Raymond (2019)	2004 – 2018	Cointegration test with structural breaks and Threshold error correction models	<ul style="list-style-type: none"> <li>• HH</li> <li>• NBP</li> <li>• NCG</li> <li>• TTF</li> <li>• JKM</li> <li>• WTI</li> </ul>	<ul style="list-style-type: none"> <li>• Gas prices exhibit increased correlations, whereas the link between gas and oil prices has weakened.</li> <li>• Significant integration in European gas markets; rise in European-American integration post-2014.</li> <li>• Asymmetric adjustments in long-term equilibrium seen between HH and NBP/JKM spot prices</li> </ul>
Kim and Kim (2019)	2000 – 2017	VECM price discovery framework	<ul style="list-style-type: none"> <li>• NBP</li> <li>• JKM</li> </ul>	<ul style="list-style-type: none"> <li>• The Asian market initially led the market, but the European gas price took over as price discovery after 2011.</li> <li>• After the market break in 2011, the Asian market became the primary adjuster, impacted by European market changes.</li> </ul>
Kim et al. (2020)	2000 – 2019	Granger cointegration and Error correction model	<ul style="list-style-type: none"> <li>• HH</li> <li>• NBP</li> <li>• JKM</li> </ul>	<ul style="list-style-type: none"> <li>• European and Asian markets are integrated, with econometric findings suggesting Qatar's emergence indirectly contributed to integration.</li> <li>• Integration between North American and European markets, present in the early 2000s, has diminished over time.</li> </ul>

# **CHAPTER 3**

## **PRELIMINARY OVERVIEW OF THE NATURAL GAS MARKETS**

### **3.1 Introduction**

The previous chapter has discussed how studies in the literature have tested relationships between regional gas markets relying on reduced-form models that primarily use price data. This chapter will explore the major regional gas trading hubs and present a historical economic account of the development of these markets.

Li, Joyeux, and Ripple (2014) and Zhang et al. (2018) highlight the distinct nature of natural gas markets, contrasting it with the more universally structured crude oil market. They emphasise that natural gas trading is predominantly regional, occurring in three major regions: North America, Europe, and Asia. In the United States, natural gas pricing mechanisms are primarily influenced by gas-on-gas competition, exemplified by the pricing formation at the Henry Hub. European markets also rely heavily on gas hubs such as the United Kingdom's National Balancing Point (NBP) and the Dutch Title Transfer Facility (TTF), where prices are determined through competitive interactions between regional market players. Nonetheless, certain European gas markets, particularly those involving Russian natural gas imports, continue to employ oil indexation in their pricing strategies. The Asian natural gas market has historically adhered to an oil indexation pricing model. Oil-indexation has the disadvantage of not fully reflecting the natural gas sector's specific supply and demand conditions. Consequently, this approach results in the Asian market being distinct from the supply and demand-driven pricing mechanism that is becoming predominant in other regions.

This chapter offers a historical overview of the above markets and proceeds as follows. Section 3.2 initially explores the historical relationship between gas and oil prices. Section 3.3 presents the distinctiveness of the global natural gas supply chain compared to conventional fossil fuels. A comprehensive examination of the major



natural gas markets under investigation, North America, Europe, and Asia, are detailed in Sections 3.3, 3.4, and 3.5, respectively.

## **3.2 Overview of the Crude Oil and Gas Price Relationship**

One basic notion is that oil prices significantly influence natural gas supply and demand, largely due to their substitution relationship. Villar and Joutz (2006) highlight that this belief stems from the interchangeable use of oil and gas, particularly before the 2000s when gas was frequently used as an alternative to oil in electricity generation and industrial applications. For instance, when oil prices rise, gas often becomes a more cost-effective option for energy production, increasing demand. This increased demand for gas, in turn, applies upward pressure on its prices, demonstrating the interconnected nature of oil and gas markets.

Since the mid-2000s, studies such as Stern (2007, 2009) have increasingly suggested that the traditional oil-gas price substitution model is no longer applicable. The main reasons are that crude oil became almost extinct from the stationary energy sectors, tight environmental standards favour the demand for gas over oil, and the emergence of new gas-burning technologies has made the utilisation of oil inefficient in the industrial sector. In addition, several other major differences make the short-run substitutability inefficient. The costs associated with the two commodities are very distinct, such as the costs related to production, processing, transportation, and storage.

Since the rapid development of gas in the late 1990s, the substitution of gas-oil demand has not provided the main link between oil and gas prices. The main link is oil-price-indexation in long-term natural gas contracts. The latter is a practical solution that traders have applied to set gas prices between points with very few transactions and rely on long-term contracts. As fossil fuel producers often produce oil and gas, sometimes from the same fields, the oil price becomes a good proxy for the opportunity cost of supplying gas. Under oil-indexation, the price of natural gas in long-term contracts is set with a formula that ties it to an average price of a basket of oil prices and includes bounds. These contractual conditions are typically known as S-shape pricing clauses.

Despite all the above, the divergence in oil and gas prices across different markets is a significant phenomenon. In the case of crude oil, despite the distinct characteristics of different benchmarks like Brent and WTI, their prices tend to be closely related. In contrast, regional supply and demand dynamics are crucial in shaping the price formation of spot natural gas prices at major hubs. As a result, the factors driving natural gas prices vary significantly by region, which will be further elaborated in subsequent sections. However, Stern and Rogers (2014) point out that the interdependency of regional natural gas markets has increased, promoting a favourable ground for global gas market integration.

The impact of oil prices on gas markets remains apparent due to the correlation between them, as outlined in some purchase agreements, particularly in the Asian gas markets. This approach, however, goes against oil and gas's contrasting supply and demand patterns and their declining interchangeability as commodities. Therefore, determining the extent of oil price influence on gas price formation in different markets is relevant when investigating the integration of the global natural gas markets. For this purpose, the Brent crude oil spot price is included in the panel data set of this thesis, as it is the most appropriate benchmark when considering the three different natural gas markets: Europe, Asia, and North America.

### **3.3 The Global Gas Supply Chain and its Unique Characteristics**

Another important discussion when assessing the relationship between gas and other energy sources is the distinctiveness of the global gas supply chain compared to other conventional fossil fuel sources, such as crude oil. The global natural gas industry requires considerable investment concentration due to the substantial capital needed for infrastructure projects. Projects such as extraction facilities, pipelines, and LNG terminals require significant long-term financial commitments. Unlike oil, which benefits from well-established and widespread infrastructure, natural gas projects typically require extensive initial investments in production and transportation systems. This concentrated and intensive investment requirement shapes distinctive market dynamics, impacting aspects ranging from regional supply security to global trade patterns.

The transportation and logistics of the natural gas supply chain present unique challenges and complexities that are distinct from other energy sources. It primarily relies on extensive pipeline networks that demand substantial upfront investment and long-term planning. Furthermore, liquefying natural gas for shipping involves intricate and costly infrastructure, requiring specialized tankers and regasification terminals at the receiving end. These facilities add layers of complexity and cost to the logistics of natural gas transportation (Energy Working Group 2018).

Putišek and Karasz (2017) outline the key aspects of the natural gas supply chain, which includes:

- **Pipeline networks:** Extensive networks are crucial for transporting natural gas from production sites to consumers, demanding substantial capital investment and careful route planning to ensure efficient distribution. The transportation pipelines, typically large steel pipes over 1 meter in diameter, operate at high pressures of up to 125 bar and pass through compressor stations, metering stations, and storage facilities. They are crucial in directing the flow of natural gas. They undergo stringent monitoring and inspection to mitigate risks such as seam failures, corrosion, and materials failure. From these transportation pipelines, natural gas moves through distribution pipelines (also known as mains), which are medium-sized pipes ranging from 0.1 to 0.7 meters in diameter. Distribution pipelines carry gas at varying pressure levels and are constructed from steel, cast iron, plastic, and occasionally copper. Finally, service pipelines smaller than 0.05 meters in diameter connect distribution pipelines to meters and typically carry odorised natural gas at low pressures.
- **Liquefaction and LNG transportation:** Natural gas often undergoes liquefaction to facilitate transportation over long distances. Natural gas is liquefied in specialized liquefaction plants where it is cooled to approximately -161.5 °C, reducing its volume significantly for easier transportation by ship or truck. These liquefaction plants typically operate multiple processing units called "trains" that work in parallel to maximize efficiency and allow for continuous operation during maintenance. The liquefaction process involves purifying natural gas to near-pure methane, removing impurities like sulphur

and CO<sub>2</sub>, and separating natural gas liquids (NGLs) such as propane and butane. The LNG is then stored onsite before being loaded onto LNG vessels — these vessels, equipped with either spherical tanks or membrane tanks, transport LNG across seas. Upon arrival at regasification plants, LNG is returned to its gaseous state through vaporization using heat exchangers fuelled by seawater. The regasified natural gas is then injected into the pipeline grid, which mixes with existing natural gas supplies for distribution.

- **Storage and distribution challenges:** Storing natural gas presents unique challenges due to its lower density than oil, necessitating advanced technologies and secure facilities for safe and efficient storage and distribution. Natural gas storage balances supply and demand fluctuations, especially between seasons. When demand is lower in the summer, natural gas is injected into storage facilities and withdrawn during the winter when demand peaks. Storage options include underground facilities such as depleted reservoirs, aquifers, caverns (salt and rock), and aboveground storage like gasholders and LNG tanks. Additionally, line pack storage within pipelines helps manage short-term demand variations. These solutions ensure a steady supply of natural gas despite fluctuations in consumer demand, providing essential flexibility and reliability to the distribution network.

In the long term, the natural gas supply chain offers a valuable model for developing hydrogen as an emerging energy carrier. In particular, green hydrogen, produced from renewable sources, requires a similarly extensive infrastructure and significant investment. The current natural gas infrastructure, including pipelines, LNG and storage facilities, could be adapted to accommodate hydrogen, easing the transition to a hydrogen-based economy. The expertise developed in managing the complexities of the natural gas supply chain will be crucial for overcoming the challenges related to hydrogen transportation and distribution. This knowledge base can significantly contribute to the growth of hydrogen as a clean energy source, potentially revolutionizing global energy markets by leveraging insights gained from the natural gas industry (International Energy Agency 2019).

The unique attributes of the natural gas supply chain (its complexity, high investment needs, and specialized infrastructure) not only define the current global gas market but

also provide a strategic framework for future energy transitions, notably the integration of hydrogen into the energy landscape.

### **3.4 Overview of the Natural Gas Market in North America**

Before 1978, pipeline owners purchased most of the gas they transported for resale from the United States. The trade transactions were made in long-term (20 or more years) purchase contracts with field producers, bundling supplied marketing and transportation as a single product. In addition, federal regulation at that time resulted in higher transaction costs for natural gas production and transmission. However, the US was the first region to reorganise its natural gas market by promoting the deregulation of wellhead prices and opening access to gas pipeline infrastructure by implementing the Natural Gas Policy Act (NGPA) in 1978 (Neumann 2009). Doane and Spulber (1994) state that after the NGPA, considerable changes happened in the pricing and transportation of natural gas in North America. These sudden changes resulted in many contract disputes as pipelines had to prevent losses from take-or-pay obligations. According to Doane and Spulber (1994), a major milestone in the unbundling of gas marketing and transportation in the US was taken by the Federal Energy Regulatory Commission (FERC) in Order No. 436 in 1985, which considered pipelines as open-access transporters for gas purchased by all classes of buyers directly from the gas producers. This created an incentive for the complete separation of the marketing and transportation services of the interstate pipeline. Moreover, it allowed the entrance of local distribution and electric utility companies and industrial customers as natural gas buyers.

After the 1980s regulatory changes, there was a notable shift in the natural gas industry from long-term to short-term contracts, typically of a duration of around one week. This shift facilitated the development of natural gas futures contracts. According to Doane and Spulber (1994), open-access transportation policies led to trading these commodity contracts, making spot prices a benchmark for long-term contract pricing and reducing transaction costs.

A significant outcome of these changes was the establishment of the Henry Hub in Louisiana as a key natural gas trading location. Brown and Yücel (2008) highlight that Henry Hub became a crucial market centre, connecting at least 16 pipelines, LNG

infrastructures, and underground storage facilities. Since its inception, it has been the reference point for the New York Mercantile Exchange (NYMEX) gas futures contract, fostering a liquid market and becoming the primary price reference for most natural gas traded in North America (Neumman 2009).

The Henry Hub can be considered the most important gas exchange centre in the world. Heather and Petrovich (2017) assesses natural gas hubs based on liquidity and transparency, focusing on five aspects: traded volume, product variety, tradability index, market participant diversity, and churn rates. By these criteria, the Henry Hub is deemed the most mature gas hub, excelling in all five aspects, and is widely regarded as a model for other gas markets globally.

Despite all the liberalisation measures since the 1980s and the emergence of a liquid and mature natural gas market in the US, the initial penetration of natural gas has historically been dependent on crude oil-linked contracts, with gas serving as a direct substitute for crude oil for heating and industrial purposes. However, Perifanis and Dagoumas (2018) stated that the US was a pioneer in weakening the linkages between crude oil and natural gas with the development of the shale gas revolution in the 2000s. The increase in unconventional shale oil and gas production oversupplied the US domestic market.

Furthermore, the Organization of the Petroleum Exporting Countries (OPEC) adopted a strategy to maintain production, plunging the crude oil price to \$26 a barrel in 2016. This situation resulted in US oil and gas companies improving efficiency and reducing costs. In addition, shale gas producers continued their increase in production without interruptions, as it became clear that demand for natural gas had different drivers than crude oil, especially considering the increased substitution of coal for natural gas in the industrial power generation units. The evolution of the natural gas market in the US has made this region the most mature and integrated regional gas market compared to the European and Asian gas markets.

### **3.5 Overview of the Natural Gas Market in Europe**

In the past, the production and export of natural gas in Europe were controlled by large publicly owned companies, which exerted monopolistic or oligopolistic market

behaviour and had extensive market power, resulting in gas prices higher than the competitive level (Hulshof, Maat and Mulder 2016). However, since the late 1990s, a series of regulatory policies has significantly changed the natural gas markets in Europe. Renou-Maissant (2012) argues that these changes resulted from European energy regulations targeting a single European gas market. These policies aimed to offer a more competitive environment to all gas consumers in the European Union (EU), as well as to promote the efficiency of the natural gas sector by providing new opportunities and enhancing cross-border trade.

The historical convention of oil indexation in the gas price, which linked the price of gas to the price of oil, has lost momentum. Oil indexation was the leading pricing mechanism in Europe until the early 21st century. However, the energy sector has developed, and oil products have yet to remain an optimal substitute for gas. Stern (2007) demonstrated that the fuel switching between gas and oil products in the European energy markets had already plunged to low levels in the mid-2000s. The share of oil-linked gas trades in Europe had recently fallen to 25% in 2018 from 80% in 2005 (Stern, 2020).

During the latter half of the 2000s, the European Union made further progress toward promoting market liberalisation in the natural gas sector by enhancing the efficacy of third-party access and ownership unbundling. Furthermore, the extinction of destination clauses in purchase contracts drastically changed the regulatory and market context in which they operate. According to Stern and Rogers (2014), the European gas market underwent a contextual shift due to two main factors: the rise of third-party access and the creation of gas hubs that offered market-based pricing of gas, accessible to any market player both within and outside Europe.

Ache et al. (2013) point out that the UK was the first European country to deregulate its gas market in the 1990s following the privatisation of the state-owned company British Gas in 1986. Following deregulation policies in the UK, gas trading has witnessed a surge in the Over-the-Counter (OTC) market, facilitated by the creation of the National Balancing Point (NBP) virtual gas hub. The NBP has established a price reference for several forward transactions and future contracts in the UK. Further to this, the fundamentals of gas-to-gas competition and spot gas prices spread to other European countries with the development of the Interconnector (pipeline connecting

the UK gas market to continental Europe), which started its operations in 1998. Consequently, spot and derivative gas trading also emerged in continental Europe, such as the Zeebrugge (ZEE) hub, the arrival point of the Interconnector pipeline system in Belgium, and the Netherlands' Title Transfer Facility (TTF) hub. In the last decade, the TTF emerged as the main gas price reference in Europe, surpassing the NBP as the most liquid and mature European gas hub and further promoting a gas-on-gas competition in the northwest European region (Bennet 2019, Liao and Skykes 2019).

The European gas market largely depended on gas imports, accounting for 73% of the EU gas demand in 2015, because of its declining domestic gas production and the low probability of developing unconventional gas reserves. Russia is a significant participant in the European gas market. Eurostat data indicates the EU's substantial reliance on gas imports from Russia, reaching nearly 40% of all European gas supply in 2020. Stern and Rogers (2014) argues about the importance of the Russian gas supply for EU energy security as it still has several long-term contracts that will expire in the next decade. In the 2000s, the Russian state-owned company Gazprom was the primary opponent of shifting from oil indexation to hub-based gas prices. In 2012, the company made several adjustments to its gas contracts, reducing the base price by 7-10% and reducing take or pay obligations to avoid losing market share. However, oil indexation still accounts for a considerable proportion of the gas pricing imported from Russia to the EU.

### **3.6 Overview of the Natural Gas Market in Asia**

Natural gas is more critical in Asia's energy portfolio than anywhere else. According to the BP (2018) report, from 2010 to 2018, Asia's contribution to global natural gas demand almost doubled from 12% to 21%. This region faces a significant challenge due to its limited domestic gas production, leading to a substantial gap between demand and supply and a consequent reliance on imports. A particular obstacle is the insufficient regional transport infrastructure, like pipelines, making Asian markets heavily dependent on imported LNG. In fact, in 2021, Asian gas markets accounted for 72% of global LNG trade, as per BP's 2022 findings.



The LNG sector is marked by high production, storage, and transport costs, influencing the nature of contractual arrangements in Asian markets, which differ notably from those in the US and Europe. As Vivoda (2019) notes, to protect the interests of both importers and exporters, LNG contracts in Asia often involve long-term agreements spanning 15 to 25 years, including take-or-pay clauses. However, Rodríguez (2008) suggests that these contractual restrictions lead to market inefficiencies. He argues that adopting a more flexible 'free destination' approach in LNG trading could potentially enhance the value of long-term LNG supplies by up to 40%.

Moreover, the traditional Asian pricing mechanism in LNG contracts linked the gas price to the Japan Crude Cocktail (JCC) price (Aguilera, Inchauspe and Ripple 2014) - the Japan Crude Cocktail is the average price of crude oil products imported into Japan (Stern and Rogers 2014). Vivoda (2014) explains that the traditional pricing formula for the cost of Asian LNG imports is  $(JCC \times \text{price slope}) + \text{an established premium}$ . The agreed premium is fixed and unaffected by any changes in oil prices, as it is typically based on shipping costs. The slope defines the relationship between LNG and oil prices. Most of the traditional Asian oil-indexed gas long-term contracts present a slope so-called "S-curve", which has a flatter slope at low and high oil prices, protecting both buyers and sellers from adverse oil price fluctuations (Ernst & Young, 2013).

While the oil-indexed pricing model for natural gas was effective in Asian markets for many years, Stern and Rogers (2014) contends that its efficiency diminished following the surge in oil prices above US\$100 per barrel post-2008. This decline in effectiveness was further aggravated by the 2011 Fukushima nuclear disaster, which led to a spike in Japan's natural gas demand and a consequent surge in Japanese LNG import prices. Consequently, natural gas prices in Asia began to deviate significantly from those in Europe and even more so from US prices, especially after America's shale gas revolution. This disparity in pricing is often referred to in the literature as the "Asian premium" (Shi and Shen 2021).

Another factor fuelling criticism of oil indexation in Asian gas markets is the increasingly apparent competition between gas and coal rather than the oil-gas substitution in the industrial energy sector. Shi and Variam (2016) argue that oil

indexation fails to accurately reflect the supply and demand dynamics in the gas markets for determining wholesale prices. Consequently, they believe this pricing strategy is no longer suitable for the current market conditions.

Most oil-indexed LNG procurement contracts in Asia are set to expire within the next 20 years, by 2030-40, thereby enhancing the importers' bargaining power to lower the oil indexation slope or shift towards spot and short-term LNG procurement contracts (Shahati, Khadadeh and Al-Aradah 2019). Another major development in the LNG industry in the last decade was the increase in the shares of LNG spot trading because of the recent surplus of supply from the shale gas revolution in the US and from new LNG projects coming online in Australia and Russia (Alim, Hartley and Lan 2018). In 2000, spot LNG trade constituted merely 5% of the total trade, whereas, in 2020, spot and short-term imports accounted for a significant 40% share of the total LNG imports (GIIGNL 2021). That is an 800% per cent increase in 20 years. These recent developments have already decreased the number of new traditional long-term agreements and stimulated the emergence of Asian LNG pricing indexes associated with the spot transactions.

To evaluate gas-on-gas spot prices, the Japan-Korea Marker (JKM) is the most widely used price index among LNG traders in the Asian market. It is published by Platts and is based on prices reported in spot market trades from Japan, South Korea, Taiwan, and China. Another alternative is the Northeast Asia Spot LNG DES (ALNG) price index, which is evaluated every week by Refinitiv (Reuters). Although assessed by different agencies, the two price indexes reflect identical transactions, resulting in a similar price trend. Furthermore, their prices are expected to have less contribution from the oil-indexed LNG prices and more influence from the local supply/demand factors.

### **3.7 Concluding Remarks**

This chapter provides a comprehensive overview of the natural gas markets, highlighting key distinctions regarding the natural gas supply chain characteristics and its price formation across the primary trading regions of Asia, Europe, and North America. It also elaborates on the complexities of natural gas pricing mechanisms,

including oil indexation, its influence in various markets, and the emergence and role of different natural gas benchmarks.

This thesis gains relevance from the notable rise in global natural gas trades over the past decade. Figures 3.1 and 3.2 illustrate the major natural gas movements in 2009 and 2021, respectively, using data extracted from the BP Statistical Review of World Energy 2010 and 2022 editions - the latest available as of May 2024. These figures reveal a roughly 50% increase in trade volumes, primarily driven by a significant surge in LNG imports to Asia.

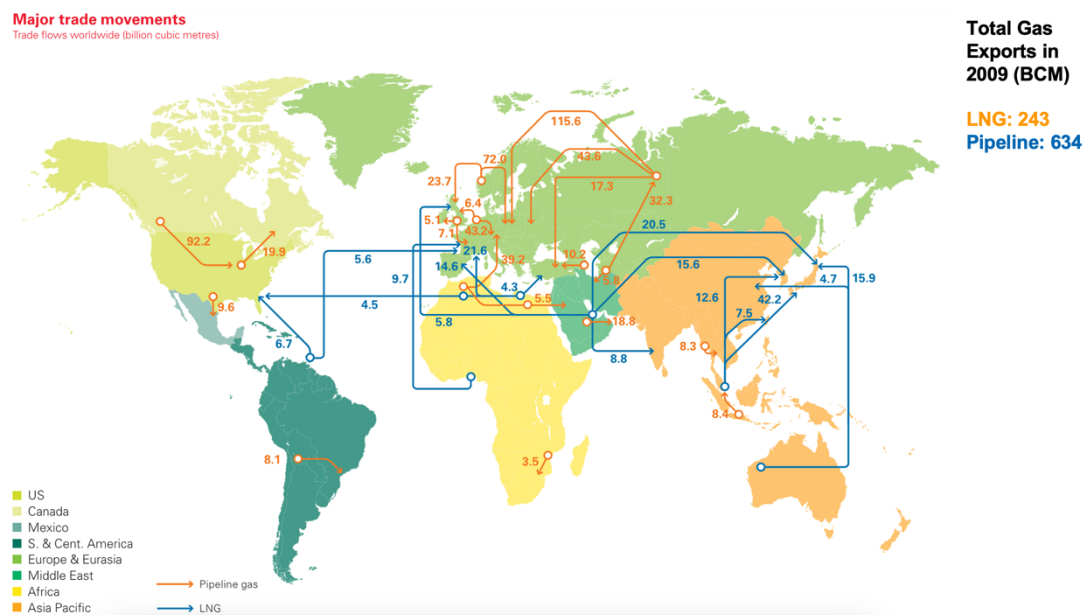


Figure 3.1 Natural Gas Major Trade Movements in BCM (Pipeline and LNG) in 2009. Source: BP (2010)

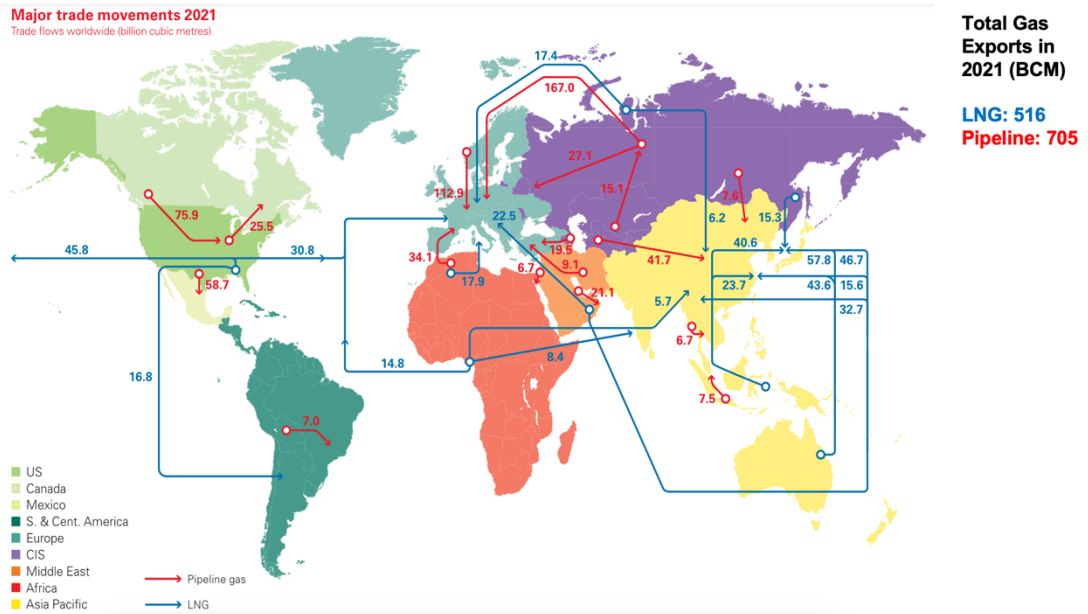


Figure 3.2 Natural Gas Major Trade Movements in BCM (Pipeline and LNG) in 2021. Source: BP (2022)

Overall, this chapter underscores the significant regional variations in natural gas price drivers and the evolution of global natural gas markets, illustrating the evolving landscape of natural gas economics in the last 20 years.

# **CHAPTER 4**

## **DATA DESCRIPTION AND TIME SERIES PROPERTIES OF THE THESIS DATASETS**

### **4.1 Introduction**

This thesis uses established econometric techniques to analyse the current relationships within the global natural gas markets and their connections to crude oil. The primary data for this study are solely time series of natural gas and oil prices, deliberately excluding variables such as regional production, consumption, and volumes of gas exports and imports. This focused approach aligns with a well-established price-based assessment methodology outlined in Chapter 2. Despite the limited data selection, we anticipate that this thesis will provide a comprehensive assessment of global natural gas market dynamics supported by the latest data.

This chapter details the six distinct natural gas prices used as key inputs in the econometric tests intrinsic to this thesis's methodology. Additionally, the Brent crude oil benchmark is incorporated into the data set to explore the influence of oil indexation across different gas markets.

One might question the criteria utilised for selecting the six natural gas prices included in this thesis' analysis. These prices represent the most significant distribution hubs and globally traded natural gas benchmarks. Heather (2023) report confirms this selection. It highlights the main European gas hubs and other key global gas prices. This data is depicted in Figures 4.1 and 4.2, which were extracted from their report. Figure 4.1 illustrates the price trends of global natural gas from 2018 to 2022 as identified by the Oxford Institute for Energy Studies (OIES), aligning with the price data selected for this thesis. Additionally, Figure 4.2 ranks European gas hubs based on their trading volumes in 2022, which corresponds with the chosen European gas prices for this study.

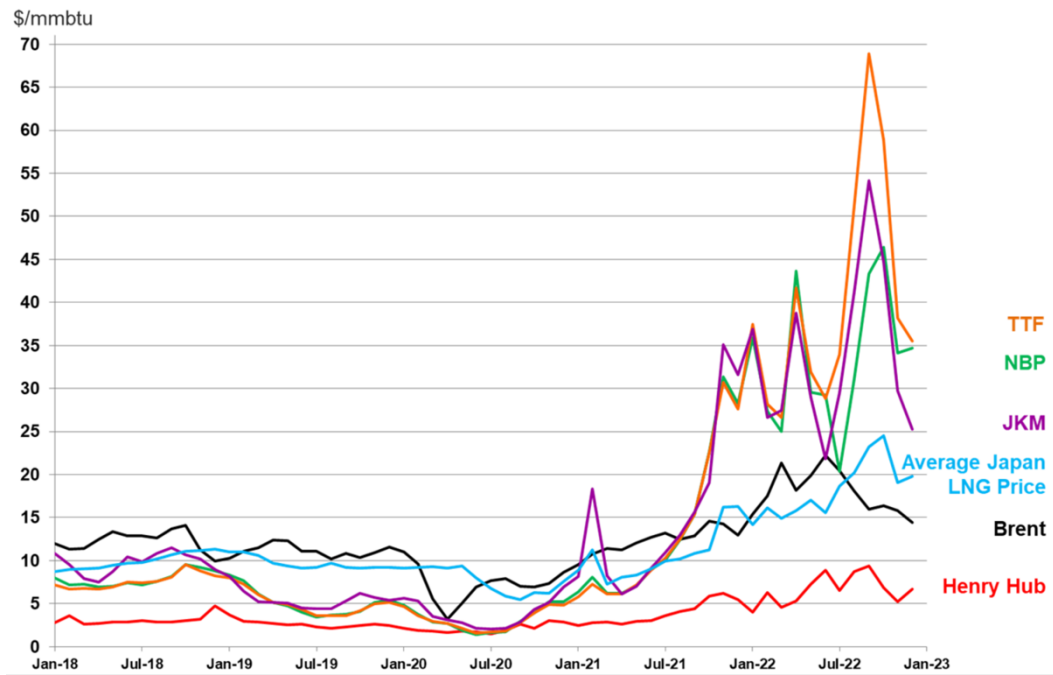


Figure 4.1 Global Gas and Brent Prices: Jan 2018 - Dec 2022. Source: Heather (2023)

2022		TOTAL TRADED VOLUMES* (TWh)						
HUB		2008	2011	2020	Δ% =>	2021	Δ% =>	2022
TTF		560	6295	46690	+14	53430	-19	43135
NBP		10620	18000	10060	-34	6640	-5	6335
NCG	THE	EST. 215	880	1965	-5	3155	+5	3305
GPL		EST. 145	310	1350				
TRF		PEG N 185	PEG N 430	890	-4	855	+66	1415
PSV		160	185	1455	-21	1155	-19	940
VTP		CEGH 165	CEGH 170	1010	-9	920	-25	685
ZTP		n/a	n/a	235	⇔	235	+138	560
PVB		n/a	n/a	145	+17	170	+61	260
VOB		n/a	n/a	95	⇔	95	-5	90
ZEE		500	870	235	-66	80	-56	35

\*rounded to nearest 5TWh; not the same data sources in all years.

Figure 4.2 Total Traded Volumes in European Gas Hubs in 2022 (TWh). Source: Heather (2023)

This chapter also examines the statistical characteristics of the time series used in Chapters 6, 7, and 8. The choice of econometric techniques to be applied to the data later will consider these statistical features. In particular, the presence or absence of a unit root in a time series can significantly impact its behaviour and properties over time. Stationary time series should be treated differently than non-stationary series in

time series econometric models (Brooks 2008). The second important aspect of this chapter is the investigation of structural breaks, represented by external shocks that can permanently change the time series behaviour over time. Considering the presence of structural breaks is essential in preventing biased results in multivariate modelling.

The definition of stationarity is crucial when exploring econometric methods such as Autoregressive Distributed Lag models (ARDL), Vector Auto Regression models (VAR), and Vector Error Correction Models (VECM), which are used to test the cointegration and causality relationships amongst the time series. Additionally, when examining the stationarity characteristics of a time series, it is important not to overlook the potential presence of structural breaks. According to Perron (1989), failure to account for structural breaks could result in unit root tests being unable to reject the null hypothesis of a spurious unit root and incorrectly identifying a time series as non-stationary. Consequently, Section 4.2 will include a unit root test that considers the presence of a structural break. This test will determine an endogenous break date based on recursive estimation of t-statistics.

The organisation of this chapter is as follows: Section 4.2 provides the data description of the gas prices used throughout this thesis' methodologies, highlighting its relevance in elaborating the main thesis' objectives. Section 4.3 offers a comprehensive overview of the main unit root tests, detailing their application in scenarios without and with a single structural break in the time series variable. Section 4.4 presents the outcomes of these unit root tests, conducted on the time series data pertinent to this thesis, both in the absence and presence of structural breaks. Sections 4.4.1 and 4.4.2 display the time series plots and their descriptive statistics, respectively. Subsequently, Sections 4.4.3 and 4.4.4 present the key findings, and Section 4.4 provides concluding remarks.

## **4.2 Data Description**

We study natural gas prices in the six major trading hubs in the world. All these prices were obtained on the standard unit of US\$ per million British thermal units (US\$/MMBtu), except the National Balancing Point (NBP) and Title Transfer Facility (TTF) prices. Unit conversion to US\$/MMBtu was implemented for these European benchmarks to align with the common unit used in gas price integration research,

enabling comparative analyses between the findings of this thesis and existing literature. A key methodological step involves transforming these time series into their natural logarithmic values, a common initial step in analysing trended time series. This transformation is implemented to facilitate the interpretation of results, as data in this format readily convey elasticity changes in percentage.

The Henry Hub gas price in North America has been the primary benchmark in the region in the last two decades. The European gas market's selection of datasets effectively represents the region's price formation dynamics. The NBP and TTF hubs are highlighted as key indicators of the EU's spot gas market prices. The Russian gas export price (RUS) is also noted for its hybrid pricing model, which combines the average import price at the Russian border with a component based on spot prices. This study will represent the Asian gas market by examining spot and contract prices. The evaluation of spot prices will be carried out through the weekly Northeast Asia Spot LNG DES assessment. In contrast, Japan's LNG customs average import prices will be utilised to evaluate long-term contracts. By analysing the Asian market through both spot and long-term contract prices, it will be feasible to examine the recent disconnection of spot and long-term oil-indexed gas prices in Asia, as suggested by Fulwood (2019).

This thesis selects the Brent crude oil price as the crude oil benchmark due to its relevance to the European and Asian markets, accurately reflecting the European energy market's crude oil pricing. Although the West Texas Intermediate (WTI) crude oil price, the US benchmark, has sometimes deviated from international crude oil prices due to domestic supply management issues. Outside these periods, both Brent and WTI benchmarks are generally closely aligned. Consequently, the choice of either benchmark is not expected to affect the econometric outcomes of the study substantially. A concise summary of each time series is provided below.

1. The variable LnOIL represents the natural logarithm of the Europe Brent Crude Oil Spot price time series. This benchmark is initially traded in US\$ per barrel. A conversion factor is employed to align with the methodology's unit convention, equating each oil barrel to 5.698 MMBtu (U.S. Energy Information Administration 2020). This time series is obtained in a monthly frequency, spanning from January 2001 (2001M01) to February 2020 (2020M02), and



includes a total of 230 data points. The source of this data is the US Energy Information Administration (EIA) database.

2. The variable LnHH is defined as the time series of the natural logarithm of the Henry Hub (HH) gas spot price, which is traded in the United States. This time series reflects the North American natural gas market. The data is in monthly frequency and spans again from 2001M01 to 2020M02. The data for LnHH was sourced from the US Energy Information Administration (EIA) database.
3. The LnNBP variable represents the time series of the natural logarithm of the National Balancing Point (NBP) gas one-day-ahead price, a virtual trading platform in the United Kingdom. The trading unit for these contracts is Pence per Therm. The conversion involves two steps to standardise this unit to US\$/MMBtu: first, converting Therms to MMBtu, and then converting Pence to US\$ using the monthly average exchange rate time series from the OzForex Limited database (OzForex Limited 2021). This time series, representing a key European natural gas price, includes monthly observations from 2001M01 to 2020M02. The data for LnNBP was sourced from Datastream and is identifiable by the Reuters Instrument Code (RIC) TRGBNBPD1.
4. The LnTTF variable represents the time series of the natural logarithm of the Title Transfer Facility (TTF) gas one-day-ahead price, which is virtually traded in the Netherlands. In the day-ahead market, participants buy or sell financial assets a day before the delivery day. The TTF price is sourced in US\$ per megawatt-hour (US\$/MWh). To align this with the study's standard unit of US\$/MMBtu, a conversion factor from the U.S. Energy Information Administration (2021) was used to transform MWh into MMBtu. This time series represents another European benchmark. Its dataset comprises 116 monthly observations covering July 2010 (2010M07) to February 2020 (2020M02). The timeframe is notably shorter than most series in this thesis, limited by data availability before 2010. The data for LnTTF is obtained from Datastream and is identified by the Reuters Instrument Code (RIC) TRNLTTFD1.
5. The LnRUS variable represents the time series of the natural logarithm of the Russian natural gas (RUS) price. This series includes the average export border

price and spot price components. This thesis considers the RUS time series as one of the European natural gas prices. This time series is characterised by a monthly frequency comprising 230 observations from 2001M01 to 2020M02. The data for LnRUS were sourced from the World Bank open database, ensuring reliable and accessible information for this analysis (The World Bank 2021).

6. The LnJPN is the time series of Japan's (JPN) monthly average LNG import prices, historically influenced by long-term contracts tied to oil prices. This thesis recognises the JPN time series as one of Asia's key natural gas prices. This time series is compiled with a monthly frequency from 2001M01 to 2020M02. The data for LnJPN was sourced from Datastream and is identifiable under the Reuters Instrument Code (RIC) LNG-TOT-JP.
7. The LnALNG variable is a time series representing the natural logarithm of spot LNG prices imported into Northeast Asia (ALNG), as assessed by Refinitiv (Reuters). This natural gas price is anticipated to exhibit a lower influence of oil indexation in its price formation, given that it is derived from spot price transactions rather than long-term contractual agreements. In the context of this thesis, the ALNG time series is identified as one of Asia's natural gas prices. This time series maintains a monthly frequency and includes 116 observations, covering the timeframe from 2010M07 to 2020M02. This period was constrained by the unavailability of data before 2010. The data for LnALNG was obtained from Datastream and is designated by the Reuters Instrument Code (RIC) LNG-AS.

Figures 4.3 and 4.4 present a graphical visualization of the price time series in two datasets in US\$/MMBtu from 2001M01 to 2020M02 and from 2010M07 to 2020M02, respectively. In addition to the time series graphs, Tables 4.1 and 4.2 present correlation matrixes from both datasets. These initial analyses will offer insights into the relationships between the prices under study. Furthermore, it serves as a foundation for the subsequent analysis in this thesis, aimed at challenging conclusions drawn solely from visual observations by examining short-run and long-run dynamic links and causality directions, identifying leading time series in price formation, and assessing long-term convergence.

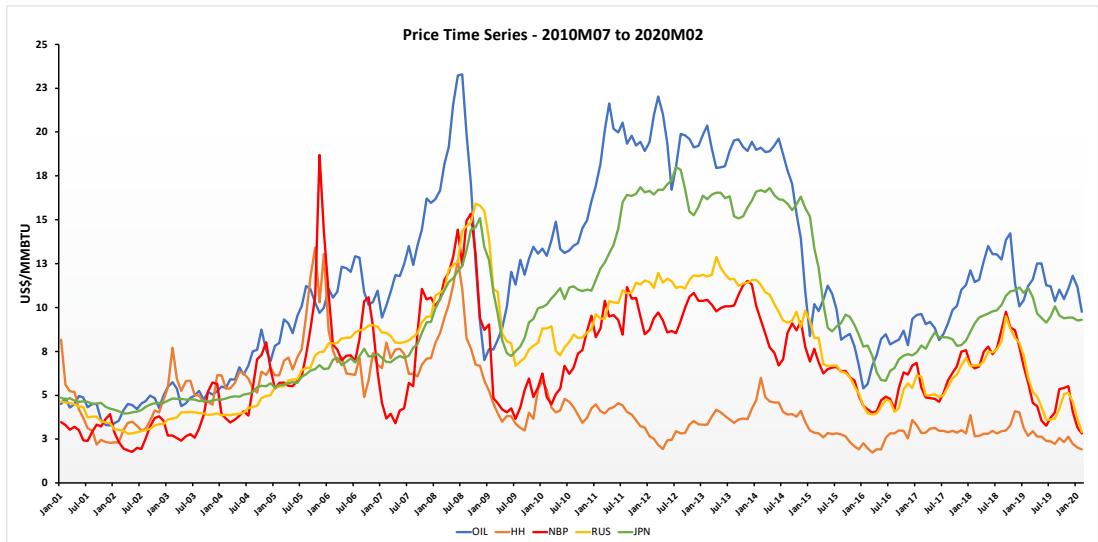


Figure 4.3 Price Time Series in US\$/MMBTu from 2001M01 to 2020M02.

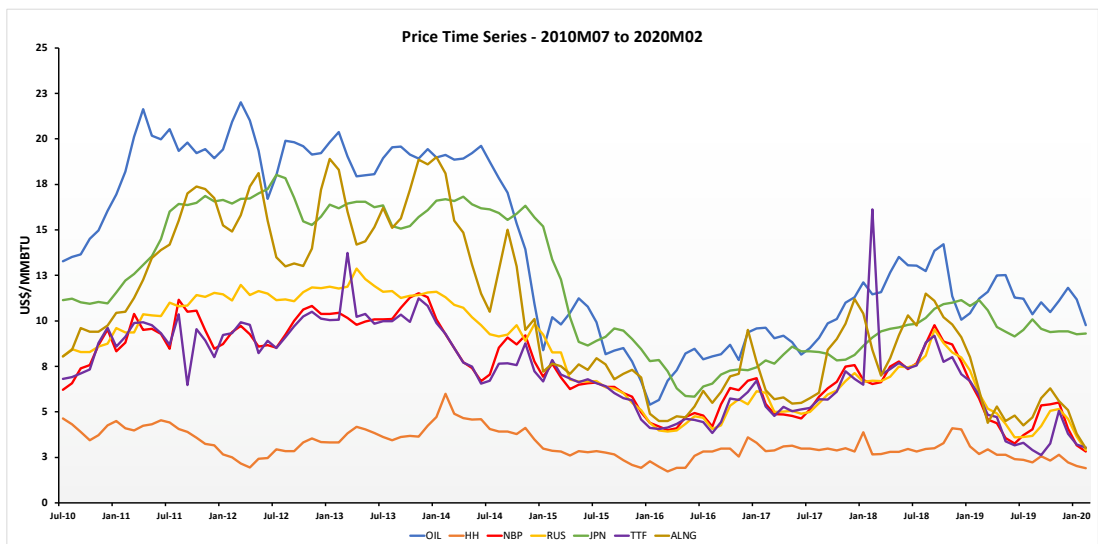


Figure 4.4 Price Time Series in US\$/MMBTu from 2010M07 to 2020M02.

Table 4.1 Correlation Matrix of Price Time Series from 2001M01 to 2020M02.

	OIL	HH	NBP	RUS	JPN
OIL	-	0.088	0.751	0.822	0.86
HH	0.088	-	0.357	0.236	-0.173
NBP	0.751	0.357	-	0.806	0.66
RUS	0.822	0.236	0.806	-	0.825
JPN	0.86	-0.173	0.66	0.825	-

Table 4.2 Correlation Matrix of Price Time Series from 2010M07 to 2020M02.

	<b>OIL</b>	<b>HH</b>	<b>NBP</b>	<b>RUS</b>	<b>JPN</b>	<b>TTF</b>	<b>ALNG</b>
<b>OIL</b>	-	0.569	0.824	0.893	0.879	0.741	0.878
<b>HH</b>	0.569	-	0.577	0.59	0.513	0.509	0.544
<b>NBP</b>	0.824	0.577	-	0.93	0.779	0.879	0.90
<b>RUS</b>	0.893	0.59	0.93	-	0.913	0.852	0.935
<b>JPN</b>	0.879	0.513	0.779	0.913	-	0.686	0.856
<b>TTF</b>	0.741	0.509	0.879	0.852	0.686	-	0.80
<b>ALNG</b>	0.878	0.544	0.90	0.935	0.856	0.80	-

### 4.3 Unit Root Tests and Stationarity of Time Series

A time series is considered stationary if its mean and variance-covariance structure remain constant over time (Lütkepohl and Krätzig 2004). Formally, the following two conditions define stationarity:

1.  $E(y_t) = \mu_y$  for all  $t \in T$
2.  $E[(y_t - \mu_y)(y_{t-h} - \mu_y)] = \gamma_h$  for all  $t \in T$  and all integers  $h$  such that  $t - h \in T$

The first condition is that the observations of a time series representing a stationary stochastic process must have a constant and equal mean. In other words, the time series should have no trend. The second condition is that the variance is also constant over time, as the variance  $\sigma_y^2 = E[(y_t - \mu_y)^2] = \gamma_0$  and covariances  $E[(y_t - \mu_y)(y_{t-h} - \mu_y)] = \gamma_h$  do not depend on  $t$ .

If a time series is non-stationary, the regression equations estimated by the OLS method will generally not result in accurate economic modelling (Glynn, Perera and Verma 2007). If variables are non-stationary, exploring the potential for a long-run cointegrating relationship between them is necessary. Cointegration occurs when a linear combination of (non-stationary and possibly stationary) time series become a stationary process. If it exists, this long-run relationship is known as an error-correction term when incorporated in models as an explanatory variable. For bivariate analysis, the ARDL models are a popular choice that can effectively determine both

long-run and short-run relationships between time series variables when they are either stationary in first difference,  $I(1)$  hereafter, or a mix of stationary in first difference and in levels,  $I(0)$  and  $I(1)$  variables (Pesaran, Shin and Smith 2001).

In assessing the unit root hypothesis for time series, it is essential to consider the findings of Nelson and Plosser (1982). This research suggests that random shocks can have enduring impacts on the long-term behaviour of time series. Such shocks may permanently alter the dynamic characteristics of the series over time. Highlighting this aspect, Zaklan, Abrell, and Neumann (2016) noted that strong autocorrelation is a key feature of long-run commodity prices, like natural gas prices, underscoring the importance of investigating whether price shocks have persistent effects on future observations. The forthcoming section will concisely review the literature on major unit root tests for time series, including those able to detect endogenous structural break and permanent shock effects.

#### **4.3.1 Stationarity and Unit Root Tests Without Structural Breaks**

A stationary time series always fluctuates around its mean with broadly constant amplitude (variance). It will always return to its mean value (mean reversion) (Hall, Taylor and Cuthbertson 1993). Therefore, a stationary time series will present the same mean, variance, and autocovariance value independent of the period measured. The importance of stationarity in time series is to define its deterministic and random components. It is essential to note that non-stationary time series can lead to erroneous regression analysis, commonly called spurious regression. This issue occurs because non-stationary time series may exhibit similar trends, leading to statistically significant coefficients in regression analysis, even when there is no genuine relationship between the variables. Yule (1926) first introduced this concept. The idea of spurious regression was further developed by Granger and Newbold (1974) as they presented an important rule of thumb to identify spurious problems in the regression:  $R^2 > d$  (Durbin-Watson value).

The literature has developed several methods to test for unit roots in time series. Next, we will briefly introduce three of the most popular methods. While a wide range of tests exist, varying in application and complexity, this discussion will be limited to the fundamental tests applied to the data in this thesis.

The first unit-root test presented in this chapter is based on Dickey and Fuller (1979) and allows the detection of a unit root in univariate time series data. This test assesses the existence of a unit root in the following stochastic process:

$$Y_t = \rho Y_{t-1} + u_t \quad (4.1)$$

Where  $Y_t$  is the time series, and  $u_t$  is the white-noise error term. The Dickey and Fuller (DF) test considers a null hypothesis on the coefficient  $\rho$ . When  $|\rho| = 1$ , (4.1) can be simplified as  $\Delta Y_{t-1} = u_t$ , and the time series  $Y_t$  is cointegrated of order 1. In other words,  $Y_t$  has a unit root and is considered non-stationary. Differencing is applied to the data to convert  $Y_t$  to a stationary time series.

There are two other possibilities to test a null hypothesis on the value  $\rho$ . When  $|\rho| < 1$ , time series  $Y_t$  will converge to a stationary series as  $t$  increases. On the other hand, if  $|\rho| > 1$ ,  $Y_t$  becomes non-stationary, and its variance increases exponentially. Manipulating (1.1) by subtracting  $Y_{t-1}$  from both sides, we obtain a more autocorrelation-friendly regression model for the DF tests, which can be expressed as:

$$\Delta Y_t = \delta Y_{t-1} + u_t ; \delta = (\rho - 1) \quad (4.2)$$

Therefore, if  $\rho = 1$ ,  $\delta = 0$ . In this case, the null hypothesis of a unit root is confirmed, and the time series tested is said to be non-stationary. It is important to notice that if  $\delta = 0$ , then (4.2) can be written as:

$$\Delta Y_t = Y_t - Y_{t-1} = u_t \quad (4.3)$$

As previously stated,  $u_t$  is a white-noise error term assumed to be stationary. It implies that the first difference of a random walk time series is also stationary. Model (4.2) may include an intercept and/or a linear time trend. If stationarity is found in these formulations, it will be concluded that the process is stationary with a drift and/or deterministic trend. In its general form, the Dickey-Fuller (DF) test, as shown in Equation (4.2), treats  $u_t$  as uncorrelated over time. However, Dickey and Fuller introduced an enhanced version known as the Augmented Dickey-Fuller (ADF) test for scenarios where the white-noise error is correlated. The ADF test modifies the equation by including lagged values of the changes in  $Y_t$  on the right side to mitigate autocorrelation in the residuals. It allows for including a constant and a trend variable

in the regression. This extension makes it possible to consider trend stationarity as an alternate hypothesis. The equation for the ADF test, based on Equation (4.1), is reformulated as follows:

$$Y_t = \alpha_1 Y_{t-1} + \dots + \alpha_p Y_{t-p} + u_t \quad (4.4)$$

Where  $u_t$  is the white-noise error term, and  $\alpha_i$  are fixed coefficients. By applying the lag operator, Equation (4.4) can be reduced as:

$$(1 - \alpha_1 L - \dots - \alpha_p L^p) Y_t = u_t \text{ or } \alpha(L) Y_t = u_t \quad (4.5)$$

With  $\alpha(L) = 1 - \alpha_1 L - \dots - \alpha_p L^p$ .

Equation (4.5) is integrated when  $\alpha(1) = 1 - \alpha_1 - \dots - \alpha_p = 0$ . For the ADF test, the hypothesis of interest is  $\alpha(1) = 0$ . However, a reparameterization of the model is required to test for the null hypothesis against the alternative of stationarity. By subtracting  $Y_{t-1}$  on both sides, (4.4) can be rearranged as follows:

$$\Delta Y_t = \phi Y_{t-1} + \sum_{j=1}^{p-1} \alpha_j^* \Delta Y_{t-j} + u_t \quad (4.6)$$

Where  $\phi = -\alpha(1)$  and  $\alpha_j^* = -(\alpha_{j+1} + \dots + \alpha_p)$ . In (4.6), the null hypothesis to be tested is  $H_0: \phi = 0$  against the alternative  $H_1: \phi < 0$ . In the ADF test, a test is conducted based on the  $t$ -statistic of the coefficient  $\phi$  from an OLS estimation of (4.6) (Dickey and Fuller 1979). The Tao-distribution of this statistic and critical values are obtained through simulation methods; MacKinnon's (1991) work is a renowned reference for these critical values in the literature.

It is crucial to choose a parsimonious lag length when using this test. Including more lag lengths than necessary in the ADF test can decrease its efficiency since it involves estimating more parameters and considering fewer effective observations. Conversely, using a small number of lag lengths may cause the null hypothesis of a unit root to be excessively rejected at any significant level (Campbell and Perron 1991). Information criterion methods provide an intriguing approach to determining the optimal lag lengths. More details about the optimal lag length selection through the main information criterion methods are in Section 6.2 of Chapter 6.

A disadvantage of the ADF test is that it may become less reliable when applied to a stationary time series with a unit root near the non-stationary boundary (Brooks 2008). We can find other methods in the literature that provide a greater testing power. We will briefly explore two of them, Phillips and Perron (1988) and Kwiatkowski et al. (1992), known as the PP and KPSS unit root tests.

A prerequisite for the Dickey-Fuller (DF) test is that the residuals should be uncorrelated and identically distributed. To address issues of autocorrelation in DF test residuals, the Augmented Dickey-Fuller (ADF) test incorporates lagged terms of the dependent variable. Phillips and Perron (1988) introduced a unit root test that resolves the serial correlation problems in residuals without adding lagged dependent variables. Their approach employs a non-parametric statistical method and applies a correction to the standard deviation, providing a robust variance estimate. The regression formula used in the Phillips-Perron (PP) unit root test is as follows:

$$Y_t = \pi Y_{t-1} + u_t \quad (4.7)$$

Where  $u_t$  may be heteroskedastic and  $I(0)$ . The PP tests account for serial correlation and heteroskedasticity in the errors  $u_t$  by manipulating the test statistics  $t_{\pi=0}$  and  $T_{\hat{\pi}}$ . The modifications in the t-statistics are represented by  $Z_t$  and  $Z_\pi$ , which are defined as follows:

$$Z_t = \left(\frac{\hat{\sigma}^2}{\hat{\lambda}^2}\right)^{1/2} t_{\pi=0} - \frac{1}{2} \left(\frac{\hat{\lambda}^2 - \hat{\sigma}^2}{\hat{\lambda}^2}\right) \left(\frac{T * SE(\hat{\pi})}{\hat{\sigma}^2}\right) \quad (4.8)$$

$$Z_\pi = T_{\hat{\pi}} - \frac{1}{2} \left(\frac{T^2 * SE(\hat{\pi})}{\hat{\sigma}^2}\right) (\hat{\lambda}^2 - \hat{\sigma}^2) \quad (4.9)$$

The variables  $\hat{\sigma}^2$  and  $\hat{\lambda}^2$  are consistent estimates of the variance parameters, represented as:

$$\sigma^2 = \lim_{T \rightarrow \infty} T^{-1} \sum_{t=1}^T E[u_t^2]$$

$$\lambda^2 = \lim_{T \rightarrow \infty} \sum_{t=1}^T E[T^{-1} S_T^2]$$



Where  $S_T = \sum_{t=1}^T u_t$ . The residuals  $u_t$  in (4.7) have a consistent estimate of  $\sigma^2$ , and the Newey-West long-run variance of the residuals is a consistent estimate of  $\lambda^2$ . The null hypothesis tested is that  $\pi = 0$ . When the null is not rejected, the PP test statistics  $Z_t$  and  $Z_\pi$  present equal asymptotic distributions as the ADF t-statistic.

The Kwiatkowski-Phillips-Schmidt-Shin (KPSS, 1992) test offers an alternative method for assessing the stationarity of time series. Unlike the ADF or PP tests, the KPSS test reverses the hypothesis framework. In the KPSS approach, the null hypothesis assumes that  $Y_t$  is I(0), meaning the time series is stationary. This positions the KPSS test as a stationarity test compared to a unit root test. The KPSS test utilises the following model for its analysis:

$$Y_t = \eta D_t + \mu_t + u_t \quad (4.10)$$

$$\mu_t = \mu_{t-1} + \varepsilon_t, \varepsilon_t \sim WM(0, \sigma_\varepsilon^2)$$

Where  $D_t$  represents the deterministic components,  $u_t$  may be heteroskedastic and I(0).  $\mu_t$  is represented by a random walk with innovation variance  $\sigma_\varepsilon^2$ . To satisfy the null hypothesis that  $Y_t$  is I(0),  $\mu_t$  must be constant. Hence,  $\sigma_\varepsilon^2 = 0$ .

The t-statistic of the KPSS test is the Lagrange multiplier (LM) for testing  $\sigma_\varepsilon^2 = 0$  against the alternative that  $\sigma_\varepsilon^2 > 0$  and is presented as follows:

$$KPSS = \left( T^{-2} \sum_{t=1}^T [\widehat{S}_t^2] \right) / \widehat{\lambda}^2 \quad (4.11)$$

Where  $\widehat{S}_t = \sum_{j=1}^t \widehat{u}_j$ ,  $\widehat{u}_t$  is the residual of a regression of  $Y_t$  on  $D_t$ , and  $\widehat{\lambda}^2$  is a consistent estimate of the long-run variance of  $u_t$  through  $\widehat{u}_t$ .

The KPSS test is particularly useful when the ADF or PP tests do not provide strong evidence to reject their null hypothesis of a unit root. Under such circumstances, the KPSS test can be instrumental in determining if there's insufficient evidence to reject its null hypothesis of stationarity. As a result, the KPSS test is best employed as a supplementary unit root test, complementing the findings from the ADF and PP tests.

### **4.3.2 Stationarity and Unit Root Tests with Structural Break**

Accurately addressing structural breaks in time series data is important, as ignoring them can lead to biased outcomes in unit root tests. Neglecting to account for structural breaks in time series data can result in not rejecting unit root null hypotheses and falsely classifying the time series as non-stationary. Chiappini, Jégourel and Raymond (2019) noted that the natural gas market has significantly changed over the past decade due to the 2008 financial crisis and the US shale gas revolution, potentially causing structural breaks. These events have significantly decoupled oil and gas prices. In this subsection, we will provide a concise overview of the commonly used unit root tests that enable the detection of a single structural break. We will focus on two structural break detection methodologies: the models proposed by Perron (1989, 1997) and the Zivot and Andrews (1992) model.

#### **4.3.2.1 Perron's Model for Unit Root Tests with Structural Break**

Perron (1989) argues that the conventional ADF unit root test may mistakenly identify a trend-stationary time series as non-stationary if a structural break exists in the intercept and/or trend functions. The ADF test may present biased results in favour of the null hypothesis of non-stationarity. To assess this, Perron (1989) enhanced the ADF unit root test by incorporating a broader set of dummy variables in the regression model. This modification allows for considering a structural break in the trend function at a pre-determined date, denoted as  $T_b$  (where  $1 < T_b < T$ ). Essentially, this method tests for a single exogenous break, grounded in underlying asymptotic distribution theory principles (Glynn, Perera and Verma 2007).

Perron (1989) introduces three distinct model settings for testing the existence of a structural break in time series. The first, termed the 'crash' model, is an ADF model enhanced with an intercept dummy variable, allowing for an exogenous break in the level of the series. The second, the 'growth rate' model, accommodates a break in the growth rate of the series, reflected in the slope of the trend function without an abrupt change in the intercept at the break date. The third model merges the features of the first two, permitting breaks in the trend function's level and slope. All three models posit the null hypothesis as a unit root with a break and the alternative hypothesis as a

stationary trend with a break. The specific hypotheses for each of these models are formulated as follows:

○ Null hypotheses

$$\text{Model 1: } Y_t = \mu + dD(Tb)_t + Y_{t-1} + e_t \quad (4.12)$$

$$\text{Model 2: } Y_t = \mu + Y_{t-1} + (\mu_2 - \mu_1)DU_t + e_t \quad (4.13)$$

$$\text{Model 3: } Y_t = \mu + Y_{t-1}dD(Tb)_t + (\mu_2 - \mu_1)DU_t + e_t \quad (4.14)$$

In these equations,  $D(Tb)_t = 1$  if  $t = Tb + 1$ ; if not,  $D(Tb)_t = 0$ . Moreover,  $DU_t = 1$  if  $t = Tb + 1$ ; if not,  $DU_t = 0$ .

○ Alternative hypotheses

$$\text{Model 1: } Y_t = \mu + \beta t + (\mu_2 - \mu_1)DU_t + e_t \quad (4.15)$$

$$\text{Model 2: } Y_t = \mu + \beta t + (\beta_2 - \beta_1)DT_t + e_t \quad (4.16)$$

$$\text{Model 3: } Y_t = \mu + \beta t + (\mu_2 - \mu_1)DU_t + (\beta_2 - \beta_1)DT_t + e_t \quad (4.17)$$

In these alternative hypothesis equations,  $DT = t - Tb$  if  $t > Tb$ ; if not,  $DT = 0$ .

In all models, only one exogenous structural break is allowed. Model 1, under the null hypothesis of a unit root, accommodates a one-time structural change at  $Tb$  in the time series level. The alternative hypothesis suggests trend stationarity with a single structural break impacting only the trend function's intercept. Model 2, known as the 'growth rate model,' involves a shift in the drift parameter ( $\mu$ ) under the unit root null hypothesis. This shift happens at the structural break date  $Tb$ , transitioning from  $\mu_1$  to  $\mu_2$ . The alternative hypothesis in model 2 implies a change in the trend function's slope without a significant drift in the time series level. Model 3, as delineated in (4.9), uniquely incorporates the potential for a sudden change in the time series' level and slope.

Building upon Perron's initial concept, Perron and Vogelsang (1992) expanded the previous approach by creating two distinct test statistics for different structural breaks. The first, the Additive Outliers (AO) model, focuses on sudden shifts in the mean of

the time series, enhancing the original 'crash' model. The second, the Innovation Outliers (IO) model, considers more gradual changes in the time series function. Later, Perron (1997) refined the IO model by differentiating between two types: IO1, which involves gradual changes in the intercept, and IO2, which includes alterations in the intercept and slope of the trend function. Perron (1997) concluded this line of work by proposing three distinct models to test for unit roots in the presence of an exogenous structural break, each represented by a specific equation. The equations representing these three models are as follows:

$$\text{AO Model: } y_t = \mu + \beta t + \gamma(DT) + \alpha y_{t-1} + \sum_{i=1}^k ci\Delta y_{t-1} + e_t \quad (4.18)$$

$$\begin{aligned} \text{IO1 Model:} \\ y_t = \mu + \beta t + \theta(DU) + \delta D(T_b) + \alpha y_{t-1} + \sum_{i=1}^k ci\Delta y_{t-1} \\ + e_t \end{aligned} \quad (4.19)$$

$$\begin{aligned} \text{IO2 Model: } y_t = \mu + \beta t + \theta(DU) + \gamma(DT) + \delta D(T_b) + \alpha y_{t-1} \\ + \sum_{i=1}^k ci\Delta y_{t-1} + e_t \end{aligned} \quad (4.20)$$

Where  $T_b$  is the period chosen for the structural break,  $DU$  represents the intercept dummy,  $DT$  represents the slope dummy, and  $D(T_b)$  represents the crash dummy. In addition,  $DU = DT = D(T_b) = 1$  if  $t > T_b$ ; if not,  $DU = DT = D(T_b) = 0$ . The test is based on the t-statistic for testing  $\alpha = 1$ .

The appropriate model selection for testing the unit root follows a top-down approach, aligning with the standard criteria outlined by Shrestha and Chowdhury (2005). This method initially employs the IO2 model as the most comprehensive option, accommodating breaks in the trend function's level and slope. This model includes the time trend ( $t$ ) and the structural break date  $D(T_b)$ , as key components. After applying the IO2 model, the statistical significance of the t-statistic for  $t$  and  $D(T_b)$  is examined. If both are not statistically significant, the analysis then shifts to assess the significance of  $DU$  and  $DT$ . A lack of significance in these four variables indicates an absence of a structural break in the time series at the specified period  $T_b$ .

However, if all four variables in the regression are statically significant, then the model IO2 is the appropriate one to choose. Another possibility is that only  $t$  and  $D(T_b)$  could be statistically significant in the set of break variables, with  $DU$  being significant testing for  $DU$  and  $DT$  only. In this case, the model allows for a structural break in the intercept ( $DU$ ). Therefore, the selected model would be the IO1 with time trend  $t$  and structural break date  $T_b$ . On the other hand, the AO model should be selected if only  $DT$  is significant when testing for  $DU$  and  $DT$ , therefore assuming a structural break only in the trend function ( $DU$ ) slope.

#### **4.3.2.2 Zivot and Andrews' Model for Unit Root Tests with Structural Break**

A key limitation of Perron's (1989) method lies in the predetermined, or exogenous, nature of the break point, chosen based on the observed behaviour of the time series. This predetermined selection can lead to issues, as Christiano (1992). In response to Christiano's critique, researchers explored methods to determine the structural break date endogenously, aiming to mitigate bias in unit root tests. The Zivot and Andrews (1992) (ZA) endogenous structural break test is a notable development in this area. This sequential method utilizes a unique dummy variable for each potential break date, allowing for a more intrinsic evaluation of the structural break.

In the ZA test, the selection of the break date is by the minimum t-statistic (negative) value from the ADF test of unit root, meaning that the break will be considered when there is the least evidence of a unit root null hypothesis (Glynn, Perera and Verma 2007). This means that the methodology does not determine the break point by inputting an estimated break date (exogenously) but rather by determining the break point endogenously through a recursive method (i.e. estimating regression models for all possible dates). The test evaluates the null hypothesis of a unit root by checking the t-statistics based on a critical value greater than Perron's (1997) critical value, thus making it harder to reject the null hypothesis in the ZA model. It chooses the break point according to the minimum t-statistic on  $\alpha = 1$  for a break point at period  $1 < T_b < T$ . The model then tests the null hypothesis by analysing the smallest t-value of the series, comparing it with some specific critical values estimated by Zivot and Andrews (1992).

The test first identifies the smallest t-value at a significant level and then compares it with the critical value. If the smallest t-value is greater than the critical value, the null hypothesis of the unit root is rejected, and the stationarity of the time series is confirmed. However, the time series can be considered non-stationary if the smallest t-value is smaller than the critical value. Zivot and Andrews (1992) developed three equations for their methodology: models A, B, and C.

A crucial consideration in all ZA models is including the endogenous variable lagged by one period. Model A consists of a regression containing the intercept dummy variable ( $DU$ ), equal to 1 in periods after the break point and zero in the periods before the break. Moreover, model A contains the endogenous variable lagged by one period, the first difference of the endogenous variable lagged by one period, and a linear trend function. Model B consists of a regression containing the trend dummy ( $DT$ ), a variable that allows for a change in the trend function and produces a new linear trend. Finally, Model C consists of a regression containing the trend dummy ( $DT$ ) and intercept dummy variable ( $DU$ ) to investigate a break in the intercept and trend functions. Formally, the models are as follows:

$$\text{Model A: } y_t = \mu + \beta t + \theta DU(T_b) + \alpha y_{t-1} + \sum_{j=1}^k c_j \Delta y_{t-1} + e_t \quad (4.21)$$

$$\text{Model B: } y_t = \mu + \beta t + \gamma DT(T_b) + \alpha y_{t-1} + \sum_{j=1}^k c_j \Delta y_{t-1} + e_t \quad (4.22)$$

$$\text{Model C: } y_t = \mu + \beta t + \theta DU(T_b) + \gamma DT(T_b) + \alpha y_{t-1} + \sum_{j=1}^k c_j \Delta y_{t-1} + e_t \quad (4.23)$$

Where the dummy variables are  $DT$  which allows for a break in the trend function, and  $DU$  which allows for a break in the intercept at the break time ( $T_b$ ).  $DU = 1$  if  $t > t - T_b$ , and equals to zero otherwise.  $DT = 1$  if  $t > T_b$ , and equals to zero otherwise. Furthermore, the rejection of the null hypothesis depends on the coefficient of  $\alpha$ . If the coefficient of  $\alpha$  is statistically significant in the regression, then the model rejects the null hypothesis.

#### **4.4 Implementation of the Unit Root Tests to the Thesis Dataset**

As outlined in Chapter 3, the market integration analysis will encompass seven time series. This set comprises six natural gas prices alongside one crude oil benchmark price. Within the natural gas price series, Europe's TTF and Asia's ALNG have a relatively shorter sampling period from July 2010 to February 2020, in contrast to the longer durations covered by the other five time series in the dataset.

This thesis uses two primary data samples for the econometric models to examine the relationships between the time series. The first sample encompasses four conventional natural gas prices and the Brent crude oil price, covering the period from January 2001 (2001M01) to February 2020 (2020M02). The second sample, with a shorter duration, includes all seven price variables, which incorporate the TTF and ALNG time series spanning from July 2010 (2010M07) to February 2020 (2020M02). Consequently, unit root testing will be conducted on these two distinct dataset representations. The outcomes of these tests will lay the groundwork for further econometric modelling that necessitates unit root verification.

This section presents the empirical findings from the unit root tests conducted on the thesis data. Figure 4.5 provides a graphical representation of all seven price time series over time. The time series plots are followed by a presentation of the descriptive statistics for each of the two data samples under consideration. Subsequently, the outcomes of the standard unit root tests and the Zivot and Andrews (1992) endogenous structural break test (ZA) are detailed.

#### 4.4.1 Graphical Representation of Thesis Data



Figure 4.5 Graphical Representation of Time Series in Natural Logs.



#### 4.4.2 Descriptive Statistics of Thesis Data

Tables 4.3 and 4.3 display the descriptive statistical outcomes for the two datasets. Table 4.1 focuses on the dataset covering January 2001 to February 2020, which includes 230 monthly observations for each of the five time series. On the other hand, Table 4.3 encompasses the dataset from July 2010 to February 2020, featuring 116 monthly observations for each of the seven time series.

Table 4.3 Descriptive Statistics of the Dataset with Sample Period between 2001M01 and 2020M02.

	<b>LnHH</b>	<b>LnNBP</b>	<b>LnRUS</b>	<b>LnJPN</b>	<b>LnOIL</b>
Mean	1.419	1.803	1.918	2.157	2.346
Median	1.371	1.862	2.012	2.168	2.401
Maximum	2.597	2.929	2.768	2.891	3.148
Minimum	0.548	0.577	1.033	1.379	1.189
Std. Dev.	0.436	0.474	0.439	0.437	0.489
Skewness	0.405	-0.364	-0.222	-0.011	-0.425
Kurtosis	2.545	2.502	1.946	1.895	2.312
Jarque-Bera	8.271	7.472	12.527	11.712	11.445
Probability	0.0160	0.0238	0.0019	0.0029	0.0033
Sum	326.425	414.649	441.188	496.146	539.4945
Sum Sq. Dev.	43.491	51.408	44.170	43.698	54.848
Observations	230	230	230	230	230

Table 4.4 Descriptive Statistics of the Dataset with Sample Period between 2010M07 and 2020M02.

	<b>LnHH</b>	<b>LnNBP</b>	<b>LnRUS</b>	<b>LnTTF</b>	<b>LnJPN</b>	<b>LnALNG</b>	<b>LnOIL</b>
Mean	1.138	1.958	2.023	1.930	2.434	2.237	2.573
Median	1.094	2.013	2.113	1.978	2.395	2.257	2.556
Maximum	1.792	2.445	2.556	2.780	2.891	2.944	3.092
Minimum	0.548	1.037	1.068	0.967	1.763	1.099	1.684
Std. Dev.	0.248	0.327	0.381	0.361	0.317	0.453	0.357
Skewness	-0.058	-0.639	-0.452	-0.596	-0.133	-0.198	-0.267
Kurtosis	2.565	2.630	2.008	2.888	1.729	2.006	1.945
Jarque-Bera	0.977	8.560	8.705	6.932	8.151	5.541	6.763
Probability	0.613	0.0138	0.0129	0.0312	0.0170	0.0626	0.0340
Sum	131.966	227.169	234.643	223.870	282.288	259.524	298.469
Sum Sq. Dev.	7.084	12.349	16.721	14.960	11.525	23.649	14.665
Observations	116	116	116	116	116	116	116

An analysis of Tables 4.1 and 4.2 reveals that, on average, producing a unit of energy with oil is generally more expensive than generating the same amount of energy from gas across various markets. The average values indicate that gas traded at HH is the most cost-effective option, while ALNG ranks the most expensive. Notably, the ALNG gas price demonstrates greater volatility compared to oil. This higher fluctuation can be linked to the lack of long-term purchase agreements associated with the ALNG price, highlighting the instability of short-term LNG import prices in Asia. As detailed in Section 3.6 of Chapter 3, long-term contracts often tie gas prices to oil prices using an 'S-curve' mechanism, which establishes limits for significant oil price variations, thus reducing volatility in gas prices.

#### **4.4.3 Empirical Results of the Unit Root Tests Without Structural Break**

This section details each variable's unit root test results within the two data samples, as outlined in the previous section. The findings are presented in two separate tables, featuring three standard unit root tests discussed earlier in this chapter: the augmented Dickey-Fuller (ADF) test, the Phillips and Perron (PP) test, and the Kwiatkowski et al. (KPSS) test.

The ADF and PP tests are initially applied to assess the unit root hypothesis. These tests are conducted on the variables in levels and first differences, selecting the appropriate lag length based on Akaike's Information Criterion (AIC) with a cap of 5 lags. Consistent with established practices, a trend function is incorporated in the regression for level testing but excluded for first difference testing in ADF and PP tests. Both tests use MacKinnon's response surface equation for critical values. The null hypothesis for these tests is the existence of a unit root.

The KPSS test similarly applies to the data in levels and first differences. The trend function is included in level testing and omitted in the first difference testing. The bandwidth for the KPSS test in levels is set at 5, and for the first differences, it is automatically determined using Newey-West selection and the Bartlett-Kernel estimation method. The critical values are derived from Kwiatkowski-Phillips-Schmidt-Shin's equation. Unlike the ADF and PP tests, the KPSS test's null hypothesis is stationarity, meaning rejection of the null indicates non-stationarity.

Table 4.5 presents the t-statistics for the unit root tests applied to each of the dataset's five variables, encompassing monthly observations from January 2001 (2001M01) to February 2020 (2020M02). The results of these unit root tests indicate that all five time series in this dataset achieve stationarity when differenced once. The results show they are integrated in the first order, denoted as I(1).

Table 4.5 Standard Unit Root Testing Results for the 2001-2020 Data Sample.

Time Series	Augmented Dickey-Fuller Tests		Phillips-Perron Tests		Kwiatkowski et al. (KPSS) Tests	
	Levels	First Differences	Levels	First Differences	Levels	First Differences
LnHH	-2.7499	-15.4679***	-3.0866	-15.4586***	0.3420***	0.0393
LnNBP	-3.0386	-7.7831***	-1.9805	-11.022***	0.5213***	0.0978
LnRUS	-0.0709	-6.2435***	-0.7902	-10.4642***	0.7492***	0.2909
LnJPN	-1.8425	-6.4456***	-1.4269	-8.4553***	0.7128***	0.1504
LnOIL	-1.9449	-11.5760***	-2.0246	-11.5052***	0.7191***	0.2132

\*, \*\* and \*\*\* denote significance at better than 10, 5, and 1 percent, respectively.

Table 4.6 displays the t-statistics of the unit root tests for each of the seven variables within the dataset covering monthly observations from July 2010 (2010M07) to February 2020 (2020M02). The results from these tests reveal that all seven time series in this dataset achieve stationarity when differenced once, classified as I(1). However, an exception is noted in the KPSS test results for LnTTF, which indicates evidence of stationarity in levels.

Table 4.6 Standard Unit Root Testing Results for the 2010-2020 Data Sample.

Time Series	Augmented Dickey-Fuller Tests		Phillips-Perron Tests		Kwiatkowski et al. (KPSS) Tests	
	Levels	First Differences	Levels	First Differences	Levels	First Differences
LnHH	-2.8010	-11.4085***	-2.3782	-11.3945***	0.0744	0.0551
LnNBP	-3.1613	-7.6703***	-2.8259	-7.5021***	1.175***	0.0574
LnRUS	-0.1673	-4.8434***	-1.9217	-6.7498***	1.3211***	0.1003
LnJPN	-2.1898	-5.6134***	-1.8820	-5.5364***	0.2130**	0.1606
LnTTF	-2.6047	-10.1376***	-3.4199*	-15.4799***	0.1098	0.0611
LnALNG	-2.7403	-8.1915***	-2.5513	-8.0624***	1.2142***	0.0736
LnOIL	-2.2932	-8.0916***	-2.0712	-7.8862***	1.1162***	0.1224

\*, \*\* and \*\*\* denote significance at better than 10, 5, and 1 percent, respectively.

#### **4.4.4 Empirical Results of the Zivot and Andrews (1992) Unit Root Test with Structural Break**

This section reports the results for the Zivot and Andrews (1992) unit root test on the two datasets in this thesis. The ZA test uses dummy variables to account for changes in the intercept, trend, or both in the time series regression and reports findings through three distinct models: Model A, Model B, and Model C, as previously outlined. Model A focuses on a single break in the intercept, Model B on a break in the trend, and Model C on both level and trend changes.

Ben-David and Pepell (1997) suggest that most economic time series exhibit a trend, advocating for the inclusion of a trend function in regression analyses. Given this, either Model B or Model C could be more suitable for estimating the break date. However, contemporary literature in econometrics often favours a single dummy variable approach for structural breaks, marked as one from the break date onwards and zero prior (as seen in works by Kisswani, 2021; Pata and Caglar, 2021; Murshed, 2020; Alsamara et al., 2019). This approach is represented in Equation 4.22. Therefore, Model B, which aligns with this methodology and focuses on a shift in trend, is deemed appropriate for identifying the break date and assessing stationarity in the time series across the two sample periods.

The ZA test is also utilised to verify the order of integration of each time series. Many econometric models analysing time series relationships are limited to handling variables that are either integrated of order zero ( $I(0)$ ) or order one ( $I(1)$ ). The inclusion of a second-order integrated ( $I(2)$ ) time series can lead to issues like highly positive residual correlations, an unexpectedly high coefficient of determination, and the risk of spurious regression. Haldrup (2002) suggests that, in cases where  $I(2)$  time series are identified, appropriate data treatment is necessary to mitigate these issues. One solution is to divide the time series into two segments, each characterised as  $I(1)$ .

The ZA test is applied to their first differences to determine if the variables are  $I(2)$ . If the results indicate trend-stationarity with a structural break, it confirms that the series is  $I(1)$ , characterised by a single structural break. Conversely, if the test shows non-stationarity in the first-differenced series, it implies the need to segment the series into two  $I(1)$  parts, each defined by a distinct structural break.

Similar to the previous unit root tests, the ZA test is conducted on two datasets. Table 4.7 displays the t-statistics for the three ZA models applied to the five variables within the dataset covering monthly observations from January 2001 (2001M01) to February 2020 (2020M02). According to Zivot and Andrews (1992), the most appropriate selection of a structural break date in a time series is when the t-statistics are significant at a maximum 5 per cent level. This significance level is indicative of the trend-stationarity of the variable.

Table 4.7 Zivot and Andrews (1992) Unit Root Testing Results for the 2001-2020 Data Sample in Levels.

	Model A		Model B		Model C	
	t-stat	Break	t-stat	Break	t-stat	Break
LnHH	-	-	-4.92***	2008M06	-5.66***	2008M08
LnNBP	-4.58*	2004M08	-4.56**	2005M11	-4.72	2004M08
LnRUS	-2.23	2015M01	-3.06	2008M05	-3.02	2007M10
LnJPN	-5.67***	2014M12	-3.87	2011M08	-5.52**	2014M12
LnOIL	-4.33	2014M07	-3.58	2007M11	-4.24	2014M07

\*, \*\* and \*\*\* denote significance at better than 10, 5, and 1 percent, respectively.

Model B determines the stationarity and identifies a structural break for each variable. Referring to Table 4.7, LnRUS, LnJPN, and LnOIL are the three variables that do not reject the null hypothesis of a unit root, aligning with earlier interpretations of non-stationarity at levels in the standard unit root tests without a structural break. Conversely, Model B indicates trend-stationarity for LnHH and LnNBP, rejecting the null hypothesis. The break points for these variables are identified in June 2008 (2008M06) and November 2005 (2005M11), respectively. This finding contradicts their previously determined non-stationarity, concluding that these two variables are integrated of order zero (I(0)) with a structural break in levels.

To verify whether the remaining three time series in the dataset are integrated of order one (I(1)), Table 4.8 presents the outcomes of the three ZA models applied to all natural gas and Brent crude oil price variables, but in their first differences. Moreover, rejecting the null hypothesis in these ZA tests will guide the optimal selection of the break point for these three variables.

Table 4.8 Zivot and Andrews (1992) Unit Root Testing Results for the 2001-2020 Data Sample in First Differences.

	Model A		Model B		Model C	
	t-stat	Break	t-stat	Break	t-stat	Break
$\Delta \text{LnHH}$	-15.62***	2008M07	-	-	-15.82***	2005M11
$\Delta \text{LnNBP}$	-8.32***	2009M08	-8.16***	2017M03	-8.45***	2005M12
$\Delta \text{LnRUS}$	-6.62***	2004M04	-6.72***	2004M08	-7.14***	2016M09
$\Delta \text{LnJPN}$	-8.61***	2016M06	-8.71***	2014M12	-8.71***	2014M12
$\Delta \text{LnOIL}$	-11.81***	2016M02	-11.88***	2016M02	-11.88***	2016M02

\*, \*\* and \*\*\* denote significance at better than 10, 5, and 1 percent, respectively.

According to Table 4.8, the five time series in the dataset demonstrate trend-stationarity in their first differences. This indicates that, aside from the LnHH and LnNBP variables, which are  $I(0)$ , the other three time series are indeed  $I(1)$ . These findings confirm the presence of both  $I(0)$  and  $I(1)$  variables in the dataset covering the period from January 2001 (2001M01) to February 2020 (2020M02). This distinction is crucial for selecting appropriate econometric models to examine the level relationships among natural gas prices. Additionally, the results verify that no time series in the dataset are non-stationary in their first differences, confirming the absence of  $I(2)$  variables.

Turning to the dataset spanning from July 2010 (2010M07) to February 2020 (2020M02), Table 4.9 presents the t-statistics for the three ZA models as applied to the seven variables, evaluated at their level values.

Table 4.9 Zivot and Andrews (1992) Unit Root Testing Results for the 2010-2020 Data Sample in Levels.

	Model A		Model B		Model C	
	t-stat	Break	t-stat	Break	t-stat	Break
LnHH	-3.391	2014M12	-2.911	2018M09	-3.441	2014M12
LnNBP	-3.632	2017M07	-3.521	2018M09	-4.158	2018M03
LnTTF	-2.652	2016M10	-2.762	2018M09	-3.792	2017M11
LnRUS	-2.517	2015M01	-2.083	2018M09	-2.768	2018M04
LnJPN	-4.933**	2015M02	-2.799	2016M06	-4.740	2015M02
LnALNG	-3.531	2014M10	-2.943	2018M09	-3.665	2017M09
LnOIL	-4.549	2014M10	-3.314	2016M01	-4.712	2014M10

\*, \*\* and \*\*\* denote significance at better than 10, 5, and 1 percent, respectively.

Table 4.9 shows that all variables in the dataset remain non-stationary at level values, consistent with previous unit root test findings. Table 4.10 then examines these variables in their first differences using the ZA models to determine their integration order and potential break points.

Table 4.10 Zivot and Andrews (1992) Unit Root Testing Results for the 2010-2020 Data Sample in First Differences.

	Model A		Model B		Model C	
	t-stat	Break	t-stat	Break	t-stat	Break
$\Delta \text{LnHH}$	-11.791***	2012M05	-11.449***	2018M08	-11.943***	2012M05
$\Delta \text{LnNBP}$	-6.617***	2016M09	-6.287***	2018M08	-6.825***	2016M09
$\Delta \text{LnTTF}$	-10.716***	2016M03	-10.337***	2018M03	-10.745***	2016M09
$\Delta \text{LnRUS}$	-6.134***	2016M09	-5.279***	2018M08	-6.403***	2016M09
$\Delta \text{LnJPN}$	-6.983***	2016M06	-5.977***	2015M03	-6.943***	2016M06
$\Delta \text{LnALNG}$	-8.792***	2016M03	-8.434***	2018M07	-8.791***	2016M03
$\Delta \text{LnOIL}$	-9.353***	2016M02	-8.249***	2014M11	-9.279***	2016M02

\*, \*\* and \*\*\* denote significance at better than 10, 5, and 1 percent, respectively.

Table 4.10 reveals that model B rejects the null hypothesis of non-stationarity for all time series when assessed in their first differences, indicating that these variables are trend-stationary with a structural break. This finding aligns with prior standard unit root test results, confirming that all variables in this dataset are I(1) and discarding the presence of I(2) variables. Thus, modelling the time series in both datasets with a single structural break is appropriate.

Building on the assumption that the rejection of the null hypothesis in the ZA test (model B) confirms the appropriate selection of the break date, Tables 4.11 and 4.12 provide a comprehensive summary of the ZA test outcomes for each time series across both datasets.

Table 4.11 Summary of the Zivot and Andrews (1992) Unit Root Testing Results for the 2001-2020 Data Sample.

	Stationarity in Levels	Break Date
LnHH	I(0)	2008M06
LnNBP	I(0)	2005M11

Table 4.11 Summary of the Zivot and Andrews (1992) Unit Root Testing Results for the 2001-2020 Data Sample, Continued.

	Stationarity in Levels	Break Date
LnRUS	I(1)	2004M08
LnJPN	I(1)	2014M12
LnOIL	I(1)	2016M02

Table 4.12 Summary of the Zivot and Andrews (1992) Unit Root Testing Results for the 2010-2020 Data Sample.

	Stationarity in Levels	Break Date
LnHH	I(1)	2018M08
LnNBP	I(1)	2018M08
LnTTF	I(1)	2018M03
LnRUS	I(1)	2018M08
LnJPN	I(1)	2015M03
LnALNG	I(1)	2018M07
LnOIL	I(1)	2014M11

## 4.5 Concluding Remarks

Chapter 4 provides a detailed exploration of the time series properties of the thesis' data, first introducing their graphical representation based on the sample periods and their descriptive statistics. Furthermore, our comprehensive analysis, employing a range of unit root tests, enabled an understanding of the stationarity and integration order of the natural gas and oil price variables. The augmented Dickey-Fuller, Phillips-Perron, and Kwiatkowski-Phillips-Schmidt-Shin tests, along with the Zivot-Andrews test, were used to assess the stationarity nature of the time series and also in the context of structural breaks.

The results obtained have significant implications for future econometric modelling in this thesis. The confirmation of stationarity and structural breaks guide the selection of appropriate econometric techniques in subsequent chapters. This analysis enhances the reliability of our market integration studies and contributes to a more nuanced understanding of the natural gas market dynamics. This chapter lays the groundwork



for the advanced modelling techniques employed in the following chapters, ensuring that the interpretations and conclusions are accurate and relevant.

# **CHAPTER 5**

## **ANALYSING NATURAL GAS PRICE CONVERGENCE: THE PHILLIPS-SUL RELATIVE CONVERGENCE TEST**

### **5.1 Introduction**

This chapter outlines the methodology, results and discussions in my recent paper, Loureiro, Inchauspe and Aguilera (2023). It performs a general assessment of long-run convergence between the various prices and identifies within convergence groups. The testing methodology keeps structural assumptions at a minimum and does not rely on unit root or structural break assumptions. The latter will be considered in the upcoming structural time series models, whose results will be interpreted in conjunction with the findings of this chapter.

Chapters 2 and 3 highlighted the significant regional price differences in natural gas, contrasting with oil prices, and explored the literature review on the debate over their convergence over time. This topic has gained considerable attention in media and specialised reports, driven by the rapid evolution of the industry and LNG trade. For instance, a 2017 The Wall Street Journal article announced the arrival of a global natural gas market (The Wall Street Journal 2017). A 2020 report from the Center for Strategic & International Studies acknowledged some price convergence post-2014 but questioned its sustainability (Center for Strategic & International Studies 2020). Existing academic studies provide varied, often outdated insights on price convergence. This points to the need for a comprehensive re-evaluation of the issue with up-to-date data for all major regional prices. Addressing this gap, this chapter presents a thorough and up-to-date global evaluation of the convergence among key regional gas prices.

Natural gas markets exhibit diverse regional structures, leading to complex economic reasoning for potential price convergence. A notable theory, originating in Neumann (2009), suggests LNG trade as a key factor in reducing price disparities across the US,

Europe, and Asia Pacific, gaining more relevance with the recent surge in LNG exports from Australia, the US, and Qatar. In contrast, another hypothesis, particularly pertinent to Asia and Europe, proposes that natural gas prices converge due to their linkage with international oil prices (Li, Joyeux, and Ripple 2014). However, the rise of spot price trading hubs in these regions and a decreasing dependency on oil indexation have challenged this view (Fulwood 2019, Liao and Skykes 2019). Hence, the expected results are unclear. Additionally, a third perspective posits that enhanced market integration, through improved distribution and storage and reduced trade barriers, should narrow price gaps; this has been observed within European markets (Dukhanina, Massol, and Lévêque 2019; Chiappini, Jégourel, and Raymond 2019; Broadstock, Li, and Wang 2020) and between the US and Canada (Olsen, Mjadel, and Bessler 2015).

In this chapter, the testing approach enhances traditional methods used in related research. Traditional testing often involves bivariate regression models with a time-varying slope coefficient determined by a Kalman filter. This method assumes that the slope coefficient should approach one if the Law of One Price is valid. Advanced studies use these estimates to develop a non-linear error-correction term with a time-varying coefficient. However, these conventional methods face limitations: they are pairwise and hence, problematic when multiple convergence pairs with common prices exist and rely heavily on parametric assumptions for the error correction term. This chapter applies the Phillips and Sul (2007, 2009) relative convergence test to overcome these drawbacks. This test identifies long-run convergence trends in a panel of stacked time series. The methodology was previously applied to gas prices by Li, Joyeux, and Ripple (2014) in a way that the present chapter follows closely to compare results. This chapter also includes traditional Kalman-filter pairwise analysis as a robustness check.

Following what was presented in Chapter 4, the panel we analyse in this thesis comprises six key gas price references. These include four traditional benchmarks: the US Henry Hub (HH), the UK National Balancing Point (NBP), Japan's Average LNG Import Price (JPN), and Russia's Average Export Price (RUS). Additionally, we incorporate spot prices from two newer, influential liquid markets - the Title Transfer Facility (TTF) price in the Netherlands and the Reuters LNG Asia spot price (ALNG).

Our dataset spans from 2001 and deliberately omits the period affected by COVID-19.

This chapter is structured in the following manner: Section 5.2 offers insights and support for testing market arbitrage activities and explores global gas market-integration hypotheses. Section 5.3 details the methodology used, emphasising two properties relevant to this thesis. The core findings from the panel analysis are presented and examined in Section 5.4. Section 5.5 presents an additional analysis for robustness, focusing on pairwise comparisons. An extended discussion of the results is found in Section 5.6. Finally, Section 5.7 concludes the chapter with key takeaways.

## **5.2 Fundamentals of Gas Arbitrage Activities and Key Hypotheses**

For an accurate assessment of arbitrage dynamics, it would be ideal to use netback prices, which represent a gas producer's actual sales revenue, accounting for costs like transportation, storage, and other transaction expenses. However, netback prices are only accessible for a limited number of markets, and compiling a comprehensive dataset of all associated costs for each market is exceedingly challenging.

Given the challenges in data availability for gas price arbitrage, there has been a significant focus in the literature on developing methods to work around this issue. Most empirical studies, including works by Broadstock, Li, and Wang (2020), Chiappini, Jégourel, and Raymond (2019), Dukhanina, Massol, and Lévêque (2019), Heather and Petrovich (2017), Renou-Maissant (2012), Neumann (2009), and Neumann, Siliverstovs, and Hirschhausen (2006), combine various price definitions, such as netback prices, pipeline-traded spot prices, and LNG prices. This thesis aligns with these approaches and focuses on long-run dynamics that can be assumed to be independent of the price definition, which is in line with the literature and the theoretical assumptions outlined next. The baseline theory is illustrated in Figure 5.1 and is deliberately simplified to capture relevant aspects only.

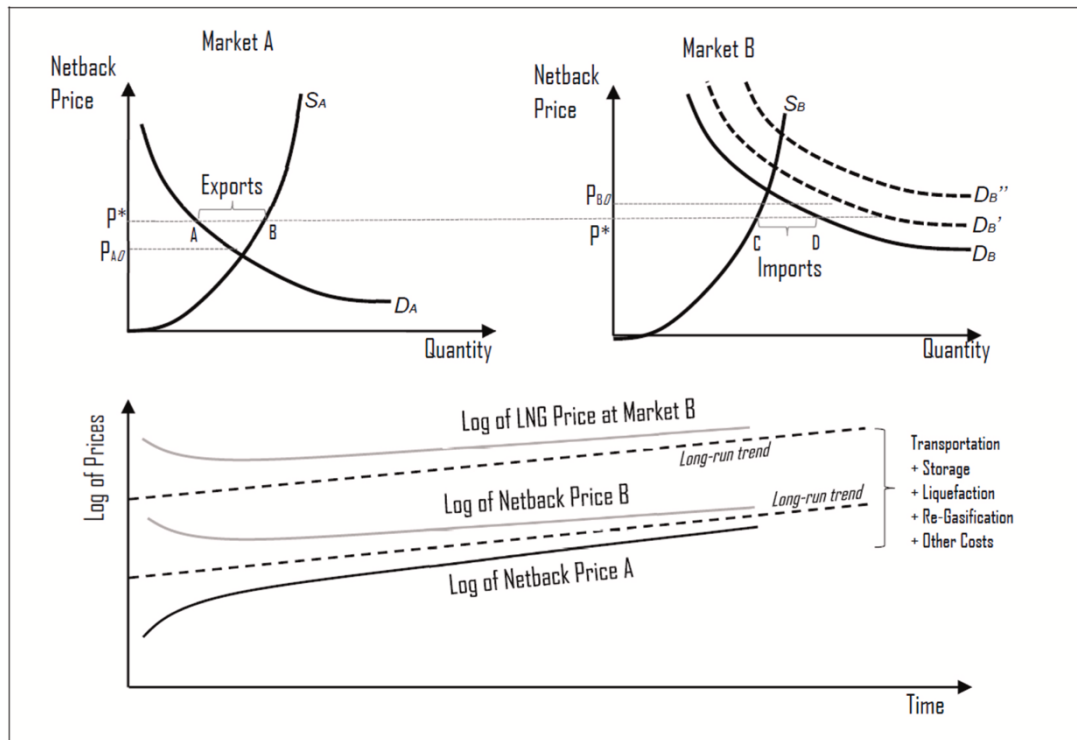


Figure 5.1 Arbitrage in International Markets and Long-run Dynamics Implications.

The model depicted in Fig. 1 illustrates the mechanism of gas arbitrage in international markets and its impact on price dynamics. It features two representative markets at the top: Market A, akin to the USA, and Market B, comparable to Asia. Each market is represented by standard demand and supply curves. Initially, in a state of autarky, Market A's price is lower than Market B's, mirroring the situation before 2016 when US LNG exports were minimal. However, post-2016, the substantial netback price difference spurred investments in infrastructure, enabling increased US LNG exports to Asia.

Theoretically, arbitrage opportunities diminish when the netback price in the US market equals what a producer earns from sales to Asia. This leads to an equilibrium price,  $P^*$ , shown in Figure 5.1.  $P^*$  represents a balance where the excess net demand in Asia ( $\overline{CD}$ ) aligns with the excess net supply from the US ( $\overline{AB}$ ), establishing a state of no-arbitrage equilibrium in the markets.

In an ideal scenario free from infrastructure constraints and other barriers, the dynamic process of gas arbitrage would lead to the convergence of netback gas prices in Markets A and B towards a specific point,  $P^*$ . However, significant infrastructure

limitations exist in the real-world example of the US-Asia gas trade. These restrictions impede the full potential of export capabilities, thus maintaining a gap in netback prices between these markets. In a theoretical framework where suppliers and buyers face no restrictions, there would be a natural incentive to develop necessary infrastructure over time, gradually allowing netback prices in each market to align with  $P^*$ . This convergence tends to be a slow process, and the primary objective of this chapter is to examine its occurrence and persistence in the long term.

Let us now focus on the long-run dynamics of demand and supply growth in gas markets. To incorporate these dynamics, assume that gas demand in Asia increases annually, causing its demand curve to shift rightward each year. Holding other factors constant, this scenario would escalate yearly export demand. Therefore, as Asia's demand surges, the long-run equilibrium netback price,  $P^*$ , to which domestic prices converge, is also expected to rise progressively ( $P^{*t} > P^{*t-1} > P^*$ ). Domestic netback prices do not align with a static equilibrium value but adjust to a continuously ascending trend.

This upward trend in  $P^*$  will not directly mirror the demand growth rate; rather, it will be influenced by the rate in conjunction with the price elasticities of demand and supply in each market. If  $P^*$  grows at a constant rate annually, the logarithm of  $P^*$  would exhibit a linear trend, as illustrated in the lower section of Figure 5.1. For simplicity, we've considered only demand growth here, but the model can be extended to include various assumptions about demand and supply growth. In some instances, we might observe a decline in demand and/or supply, which could lead to a downward trend for  $P^*$ .

In the empirical section of this thesis, we avoid any pre-emptive assumptions about the direction of this underlying trend. Additionally, our chosen empirical method is resilient to temporary price divergences. It is well-suited to accommodate short-term demand and supply shocks, like those experienced during the Global Financial Crisis. This approach is particularly effective for our long-term analysis.

It is important to address the definitions of prices used in our analysis. The focus was primarily on netback gas prices, which are not always directly observable. This led to

exploring relative convergence in this chapter, as opposed to absolute price level convergence.

Relative convergence refers to the alignment of growth rates. The lower section of Figure 5.1 elucidates this concept, showing that the growth rate of LNG prices paid by Market B at its destination aligns with the growth rate of netback spot prices in Market B. However, these prices are anticipated to stay at a manageable level due to costs like transportation, storage, liquefaction, re-gasification, and other transaction fees, which create a vertical gap between the long-run trends depicted in Figure 5.1.

For such convergence to happen, these transaction costs must not display a long-term positive growth trend. If transaction costs were to decrease, which is plausible given the potential economies of scale from trade infrastructure development, this would not impede convergence. Although with a lack of comprehensive cost data, transaction costs seem unlikely to increase over time consistently. Nevertheless, in our robust hypothesis testing, we remain open to the possibility of steady cost increases as a potential explanation for non-convergence outcomes.

This present chapter tests an intuitive hypothesis: the convergence of growth rates among two or more gas prices, each uniquely defined by their associated transaction costs. Should we observe such convergence, it would indicate effective arbitrage, likely driven by the expansion of infrastructure in response to market incentives. On the other hand, if convergence is not evident, there are two alternative explanations to consider. Firstly, the lack of observed convergence might be due to transaction costs, which are influenced by the level of infrastructure investment, exhibiting a long-term upward trend. Secondly, convergence might only occur if the existing infrastructure (and its growth rate) and other barriers, such as regulatory hurdles, are sufficient to facilitate effective arbitrage through trade. While we lean towards the second explanation as being more plausible in most scenarios, our analysis is constrained by data limitations. As a result, our conclusions can only confirm the presence or absence of effective arbitrage but cannot definitively identify the reasons for its failure in certain instances.

The arbitrage process in the gas market can be traced through an analysis of trade flows, consumption, and production data. However, the complexity of identifying

arbitrage operations increases significantly due to multiple trade patterns. This complexity is a key reason why price assessment is often preferred. Another advantage of using price data is its frequency, which results in a larger number of observations compared to gas production and consumption data, thereby enhancing the robustness of the findings.

Although our study employs the Phillips-Sul method, there has yet to be a simulation to determine the minimum number of observations required for this specific approach. Drawing from Toda (1994) findings and recommendations on cointegration analysis, at least 100 observations are generally necessary to achieve robust results. While we do not directly use trade data in our analysis, it remains crucial to consider trade volumes as a contextual factor.

Two major developments have significantly transformed the international gas trading landscape in recent years. Firstly, the period preceding the COVID-19 pandemic witnessed a remarkable surge in LNG trade. Based on BP's 2020 data, US LNG exports experienced exponential growth, escalating from 1.5 bcm in 2010 to 74.5 bcm by 2019. Similarly, Russia's LNG exports expanded from 13.5 bcm in 2010 to 39.4 bcm in 2019, with significant portions delivered to Europe and Asia. Qatar, a major player in the LNG market, saw its exports increase from 77.8 bcm in 2010 to 107.1 bcm in 2019. Australia's LNG exports also showed a remarkable rise, from 25.8 bcm in 2010 to 104.7 bcm in 2019. Notably, Australian exports, primarily consumed within Asia, directly compete with LNG from the US, Qatar, and other Asian sources.

The second key development in recent years is the rise of spot markets in Europe and Asia, representing a shift from traditional oil-indexed pricing systems to those driven by market demand and supply forces. This change in gas price formation offers both benefits and drawbacks. On the positive side, gas prices become less susceptible to the volatile fluctuations of the oil market, including seasonal variations that significantly impact oil storage levels. It could also lead to fewer contract disputes, renegotiations, and arbitrations in the gas sector, as market participants are not bound to long-term, oil-linked agreements spanning decades. However, the transition from oil-indexed pricing to spot pricing is challenging. While economists and analysts generally favour spot pricing for its reflection of the product's fundamental value, natural gas in most economies no longer directly competes with oil. Oil is mainly used in transportation,



whereas natural gas is primarily for stationary applications like power generation and industrial production.

Despite these changes, there are still advocates for oil indexation. It often results in higher prices for consumers, but it offers a degree of security regarding volume and predictability for future imports. Oil prices are volatile and unpredictable, but oil-indexed contracts often include protective measures like price ceilings and floors, offering some stability. Oil-indexed pricing ensures demand security for gas producers, encouraging investment in capital-intensive infrastructure, especially for LNG sellers concerned about demand risks. Furthermore, the gas market can experience significant instability, with periods of oversupply or shortage. Oil indexation can provide a measure of price stability, especially during relatively stable oil market periods. Over the past decade, a notable trend has been adopting a hybrid pricing model, combining oil-indexation and spot gas pricing elements. In this hybrid model, the gas price is determined by a formula where a portion is based on oil-indexation and the remainder on a spot price benchmark like HH.

Considering the insights above, the rise in LNG trade flows, coupled with markets functioning in a more price-flexible setting, could create sufficient arbitrage opportunities. This idea may lead to the relative convergence of regional gas prices globally, though not necessarily to their absolute convergence. Figure 5.2 provides additional visualisation that supports this hypothesis.

**Average spot gas prices (US\$/MMBtu) in 2019:**

Henry Hub (North America): \$ 2.57

Title Transfer Facility (Europe): \$ 4.14

National Balancing Point (Europe): \$ 4.69

Reuters LNG Asia Delivered Ex-Ship Spot Price (Asia): \$ 5.41

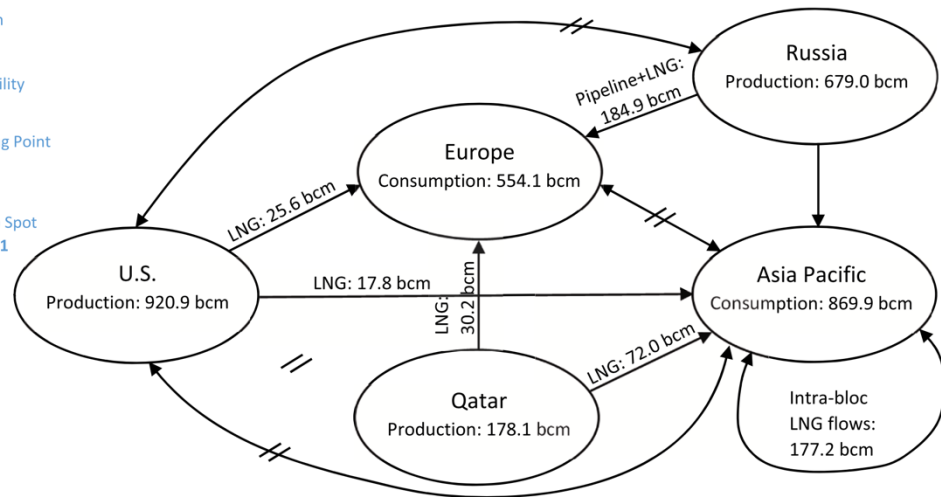


Figure 5.2 Schematic Representation of the Price-Levelling Arbitrage Hypothesis.

Source: Data from BP (2020) for 2019.

This analysis supports using the prices listed in Table 5.1 for our evaluation. All price data was converted into natural logarithms of their value in US dollars per energy unit (MMbtu). The Brent oil price, a leading global benchmark, is also included to assess the oil-gas connection in the three main markets. Following Hodrick and Prescott (1997), we applied the Hodrick-Prescott filter to the data series to align with common practices in studies using the Phillips-Sul methodology (2007, 2009). This process helps filter out short-term fluctuations, enhancing the effectiveness of the long-term convergence test. The summary statistics of this analysis can be found in Table 5.2.

Table 5.1<sup>4</sup> Variables Description and Sourcing.

Data Sourcing	
Price Variable	Original Source
U.S. Henry-Hub Price ( $P_{HH}$ )	U.S. Energy Information Administration
British National Balancing Point ( $P_{NBP}$ )	Refinitiv. Reuters Instrument Code: TRGBNBD1
Japan's Average LNG Import Price ( $P_{JPN}$ )	Refinitiv. Reuters Instrument Code: LNG-TOT-JP

<sup>4</sup> Table 5.1 continues on the next page.

Table 5.1 Variables Description and Sourcing, Continued.

Price Variable	Original Source
Russia's Export/Spot Weighted Average ( $P_{RUS}$ )	World Bank Open Database
Reuters LNG Asia Delivered Ex-Ship Spot Price ( $P_{AS}$ )	Refinitiv. Reuters Instrument Code: LNG-AS
Dutch Title Transfer Facility Day-Ahead ( $P_{TTF}$ )	Refinitiv. Reuters Instrument Code: TRNLTTFD1
Brent Crude Oil Price ( $P_{BRENT}$ )	U.S. Energy Information Administration

Table 5.2 Descriptive Statistics for Monthly Regional Price Variables.

Descriptive Statistics							
	Sample: 01:2001-02:2020					Sample: 07:2010-02:2020	
	$P_{HH}$	$P_{NBP}$	$P_{JAP}$	$P_{RUS}$	$P_{BRENT}$	$P_{AS}$	$P_{TTF}$
Mean	1.419238	1.802822	2.157158	1.918210	2.345628	2.237279	1.929912
Median	1.280852	1.921270	2.204283	1.928491	2.428987	2.171551	1.889165
Maximum	2.033104	2.262484	2.791791	2.419452	2.961698	2.735219	2.258278
Minimum	0.884456	0.981227	1.484790	1.243779	1.384435	1.615430	1.328654
Std. Dev.	0.354551	0.356089	0.404491	0.384240	0.438752	0.348527	0.256293
Jarque-Bera Prob.	0.000025	0.000000	0.003718	0.000081	0.000030	0.004719	0.048803
Observations	230	230	230	230	230	116	116

The sample period was carefully chosen using the best monthly data available, excluding the COVID-19 era, to avoid skewed results. This exclusion is crucial because the pandemic led to irregular demand and supply shifts in natural gas, resulting in significant price differences and storage capacity problems unique to that period and not indicative of long-term trends. During the pandemic, oil-linked gas prices frequently fell to their lowest contract levels, aligning with the 2020 oil price crash. This crash included instances where West Texas Intermediate oil futures contracts had negative values due to storage constraints. To ensure comprehensive analysis, convergence tests were applied to the entire dataset and specific segments identified by structural changes and data availability.

### 5.3 Relative Convergence Testing and Clustering Methodology

In this section, we describe our methodology, which is grounded in the approach developed by Phillips and Sul (2007, 2009). The central concept of this test is to identify price time series, denoted as  $P_{it}$ <sup>5</sup>, from a panel encompassing  $N$  distinct prices expanded by  $i = \{1, \dots, N\}$  over a period  $t = 1, \dots, T$ . These series are analysed for convergence towards a common long-run trend, represented as  $\mu_t$ . It is hypothesised that each time series  $P_{it}$  includes a unique long-term stochastic trend  $x_{it}$  and a stochastic, potentially non-linear, time-varying factor  $a_{it}$  that accounts for all other variations. The model can be expressed as follows:

$$P_{it} = \left( \frac{a_{it} + x_{it}}{\mu_t} \right) \mu_t = \delta_{it} \mu_t \quad (5.1)$$

This introduces the challenge that the transitional dynamic coefficient  $\delta_{it}$  and the long-run trend  $\mu_t$  are not directly observable. To address this, we focus on deriving the relative transition coefficient  $h_{it}$ , which can be calculated from the available data. The formula for  $h_{it}$  is defined as follows:

$$h_{it} = \frac{P_{it}}{N - 1 \sum_{i=1}^N P_{it}} = \frac{\delta_{it}}{N - 1 \sum_{i=1}^N \delta_{it}} \quad (5.2)$$

Equation (5.2) facilitates the testing of the convergence hypothesis by examining whether  $\delta_{it}$  approaches a constant value  $\delta$  as time  $t$  approaches infinity. Similarly, this is equivalent to checking if the relative transition coefficient converges towards 1 as  $t$  goes to infinity. To implement this, we adopt semi-parametric approximations for the mean-square transition differential, denoted as  $H_t$ , and for  $\delta_{it}$ . These approximations are used to implement the test effectively.

$$H_t = \frac{1}{N} \sum_{i=1}^N (h_{it} - 1)^2 \cong \frac{A}{L(t)^2 t^{2\alpha}} \quad (5.3)$$

---

<sup>5</sup> For a price time series  $p_t$  expanding over  $t = \{1, \dots, T\}$ , the Hodrick and Prescott (1997) filtered value  $P_t$  is obtain as  $Argmin_{p_t} \{ \sum_{t=1}^T (p_t - P_t)^2 + \lambda [(p_{t+1} - P_t) - (p_t - P_{t-1})]^2 \}$ . Following the authors' recommendation, we adopt the penalty coefficient  $\lambda = 14,400$  for monthly observations.

$$\delta_{it} \cong \delta_i + \frac{\sigma_i}{L(t)t^\theta} \xi_{it} \quad \text{with } \delta_{it} \rightarrow \delta_i = \delta \text{ as } t \rightarrow \infty \quad (5.4)$$

In Equation (5.3)  $A > 0$  and  $\sigma_i > 0$ . The variable  $\xi_{it}$  is assumed to be independently and identically distributed (*iid*) with a mean of 0 and a standard deviation of 1. The parameter  $\alpha$  represents the speed of adjustment, which is positive in the context of convergence. To facilitate the examination of convergence, even in scenarios where  $\alpha$  equals 0, we introduce the slow-changing function  $L(t) = \log(t)$ . This formulation clearly articulates the null hypothesis of relative convergence and its corresponding alternative hypothesis as<sup>6</sup>:

$$H_0: \delta_i = \delta, \text{ and } \alpha \geq 0 \quad (5.5)$$

$$H_1: \{\delta_i = \delta \text{ with } \alpha < 0\} \text{ or } \{\delta_i \neq \delta \text{ with } \alpha \geq 0\} \quad (5.6)$$

In their study, Phillips and Sul (2009) present a method for deriving a rejection rule based on the HAC *t*-statistic of  $b$  in OLS regression, as outlined in Equation (5.7). They explain that as  $b$  converges in probability towards  $2\alpha$ , it follows an asymptotic normal distribution. This convergence and distribution characteristic forms the basis for their proposed rejection rule.

$$\log\left(\frac{H_1}{H_t}\right) - 2 \log L(t) = a + b \log t + u_t \quad (5.7)$$

The term  $2 \log L(t)$  in the model functions as a penalty factor in the test, with an initial trimming of 25% of the observations. The test rejects the null hypothesis of relative convergence when the calculated t-statistic,  $t_{\hat{b}} < t_{0.05; T-k-1} \cong -1.65$ .

Initially, this test was applied to a comprehensive panel dataset of  $N$  prices. However, it can also be effectively utilised for subsets of prices or clusters, which leads to the potential identification of multiple convergence clusters, denoted as  $\{G_1, \dots, G_c\}$ , with each cluster having its distinct  $\mu_c$  and  $\delta_c$  values. To identify these clusters, Phillips and Sul (2007, 2009) propose a clustering algorithm as presented next. This

---

<sup>6</sup> That is,  $\lim_{t \rightarrow \infty} \delta_{it} = \delta$  iff  $\delta_{it} = \delta$  and  $\alpha \geq 0$ ; and  $\lim_{t \rightarrow \infty} \delta_{it} \neq \delta$  iff  $\delta_{it} \neq \delta$  and/or  $\alpha < 0$ .

approach is more efficient than “brute force's” exhaustive examination of all possible cluster combinations “.

- Step 1.** Arrange the series  $\{P_{it}\}$  from the panel dataset in descending order based on their final values, beginning with the series that has the highest final value.
- Step 2.** Form a core convergence cluster with the first  $k$  highest ranked  $\{P_{it}\}$  series in the panel. Determine the optimal cluster size  $k^*$  after running regression (5.8) for all possible cluster formations starting from the top and applying this criterion:  $k^* = \text{ArgMax}_k \{t_{\widehat{b}_k}\}$  subject to  $\text{Min}_k \{t_{\widehat{b}_k}\} > -1.65$ .
- Step 3.** Sieve the data. Add one price series  $I$  at a time to the core group starting from the top. Keep the member if  $t_{\widehat{b}_k} > 0$ .
- Step 4.** Repeat steps 2 and 3 in a loop with the remaining prices to identify all possible convergence groups in the data set.
- Step 5.** Conduct a convergence test on all contiguous groups identified through the procedure, known as the "club merging test".

In the context of this chapter's application, it is important to emphasize two notable characteristics of this procedure.

**Property 1.** *Define relative convergence as  $\lim_{t \rightarrow \infty} E[P_{it}/P_{jt}] = 1$ , level convergence as  $\lim_{t \rightarrow \infty} E[P_{it} - P_{jt}] = 0$ , and growth convergence as  $\lim_{t \rightarrow \infty} E[\Delta P_{it} - \Delta P_{jt}] = 0$ . Consequently, (i) The type of convergence observed in this test is relative, not level, which is ideal for including price differences due to transportation, storage, liquefaction, and re-gasification costs; and (ii) Relative convergence suggests the presence of growth convergence.*

This concept can be better understood with an example. Suppose we have  $\mu_t = t$ ,  $P_{1t} = 5 + 0.05\mu_t + u_t$ , and  $P_{2t} = 2 + 0.05\mu_t + v_t$ , where  $u_t$  and  $v_t$  are zero mean i.i.d. shocks. From this, it can be deduced that  $\lim_{t \rightarrow \infty} E[P_{1t}/P_{2t}] = 1$ , indicating that the expected ratio of  $P_{1t}$  to  $P_{2t}$  approaches 1 as  $t$  goes to infinity. Similarly,  $\lim_{t \rightarrow \infty} E[P_{1t} - P_{2t}] = 3$ , showing a consistent difference of 3 in the long term. Furthermore,  $\lim_{t \rightarrow \infty} E[\Delta P_{1t} - \Delta P_{2t}] = 0$  suggests that the difference in their increments over time converges to zero, implying both relative and growth convergence. Visually,

imagine the series expanding over time along two parallel lines that converge in relative values and growth rates, yet they never intersect (converge in levels), maintaining a constant differential.

**Property 2.** *A pair of prices sets,  $\{P_{it}, P_{jt}\}$ , demonstrating price convergence might not always pass a cointegration test. However, they can be regarded as asymptotically cointegrated. Moreover, these sets may also display brief periods of transitional divergence.*

From Equation (5.1), we can express the sequence as  $P_{it} - P_{jt} = (\delta_{it} - \delta_{jt}) \mu_t$ , and focus on scenarios where  $\mu_t$  is non-stationary while  $\delta_{it}$  and  $\delta_{jt}$  converge to  $\delta$ . When adjustment speeds are relatively slow,  $P_{it}$  and  $P_{jt}$  might not form a stationary linear combination that confirms cointegration. Nevertheless, these series can still be considered asymptotically cointegrated. This concept is particularly relevant for analyzing markets gradually evolving towards a stable long-term relationship, even if they have not fully stabilized yet. Additionally, the theory acknowledges that short-term transitional divergence may occur but does not significantly impact the overall convergence assessment; a point further elaborated in Phillips and Sul (2007).

## **5.4 Phillips and Sul (2007, 2009) Convergence Test Results**

### **5.4.1 Whole sample convergence clustering**

Our analysis commenced with a convergence test on the entire dataset, covering the period from January 2001 to February 2020, with the process and results detailed in Table 5.3. The data was organized by the logged US dollar prices per million British thermal units, revealing a significant correlation between the Brent oil price and Japan's imported gas price, indicative of Japan's long-term oil-indexed gas contracts.

Further analysis highlighted a trend between Russia's average gas price and the UK's National Balancing Point gas price, reinforcing the known pattern of interconnected European gas prices, including the UK. However, these trends are individual and do not merge into a single long-term direction. In a unique case, the Henry Hub price follows its distinct pattern, not aligning with the other long-term trends identified in the dataset.

Table 5.3 Convergence Test Results for 01:2001-02:2020.

Clustering Algorithm Implementation			
Numerical values are $t_{\bar{b}_k}$ statistics for each step in the algorithm			
Step 1: Ordering	Order the members of the panel according to their final values: $P_{1T} > P_{2T} > \dots > P_{8T}$		
Step 2: Form core groups	Iteration I	Iteration II (Step 4)	Iteration III
Cumulatively moving down.			
Optimal cut-off point criterion:	$P_{BRENT}$	(Base)	-
	$P_{JPN}$	3.87*	-
$k^* =$	$P_{NBP}$	-23.18	(Base)
$\text{ArgMax}_k\{t_{\bar{b}_k}\}$ subject to	$P_{RUS}$	-32.72	6.00*
$\text{Min}_k\{t_{\bar{b}_k}\} > -1.65$	$P_{HH}$	-55.39	-39.34
Core groups	Core group	$P_{BRENT} + P_{JPN}$	$P_{NBP} + P_{RUS}$
Step 3: Sieving			
Add one member at a time to Core Group. Keep member iff $t_{\bar{b}_k} > c^*$ , where $c^* = 0$ is adopted.	$P_{BRENT}$	-	-
	$P_{JPN}$	-	-
	$P_{NBP}$	-23.18	-
	$P_{RUS}$	-30.06	-
	$P_{HH}$	-52.82	-39.34
Final Cluster ("Club") Formations	Club 1: $P_{BRENT} + P_{JPN}$		
	Club 2: $P_{NBP} + P_{RUS}$		
	Club 3: $P_{HH}$		
Club-merging $t_{\bar{b}_k}$ (Step 5)	Club 1 + Club 2: -32.72 (Non-Convergent)		

In Figure 5.3, the relative convergence coefficients reveal shifting patterns in gas prices. Considering data until mid-2006, all prices moved together, as indicated by coefficients nearing 1. However, data after 2006 show a change, clarifying some past misunderstandings in research on the topic. Specifically, Brent and Japan's gas prices followed the same trend until mid-2015, heavily influencing overall convergence assessments. Yet, there is a noticeable divergence after 2015, possibly due to the rise of gas-on-gas competition and new market entrants. Meanwhile, the convergence between the UK's NBP and Russia's gas prices strengthened from 2016 onwards, which aligns with earlier findings of European market integration and might need to be noticed by studies that include the most recent data.



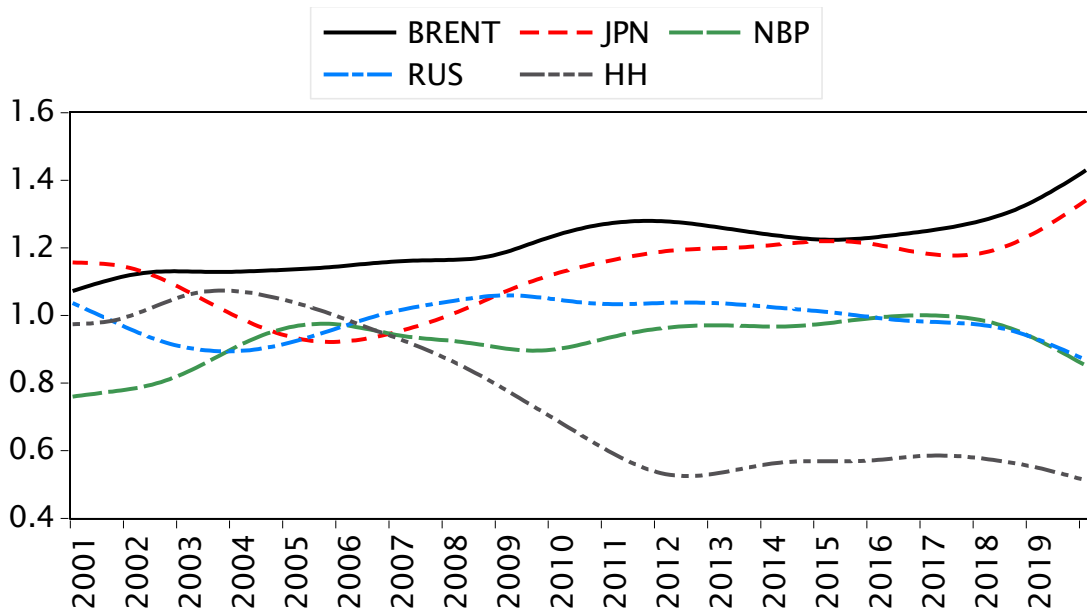


Figure 5.3 Relative Convergence Coefficients for the 2001-2019 Panel Data of 6 Stacked Prices.

The findings presented require careful consideration due to a significant structural break that coincides with the global financial crisis (GFC)—this period marked historical highs in oil and gas prices. The Zivot and Andrews test, detailed in Chapter 4, confirmed the existence of this breakpoint. A notable finding from the test is the early detection of a break in the Henry Hub price series around August 2008, marked by a highly significant t-statistic of -5.66, significant at the 1% level. This significant break point justifies further analysis of the data in subsamples, as shown in the following sections.

### 5.4.2 The Pre-GFC Sub-period

Figure 5.3 suggests that the period from 2001 to 2008 was characterized by a trend of apparent convergence in gas prices, impacting studies that excluded later data. Analysis of this sub-sample (presented in Table 5.4) indicates a weak convergence among all gas prices towards a single trend, as per the club formation test results. This finding is even more pronounced when excluding the last two years of the sample.

This period's analysis clarifies certain findings in previous studies. For example, it supports Nick and Tischler (2014) observation of converging prices in the US and UK, a pattern not evident in the complete dataset. It also aligns with Li, Joyeux, and Ripple

(2014) results of convergence between the NBP and Japan's gas price, initially attributed to the oil-link hypothesis. However, this hypothesis does not hold with more recent data, indicating that the impact of oil prices may be limited to Japan's gas price.

Table 5.4 also demonstrates that no single gas price converges towards the oil price, which suggests that oil indexation in contracts does not automatically lead to convergence in growth rates with oil prices for two reasons. Firstly, long-term contracts often include a profit margin on oil prices that varies with market conditions. Secondly, these contracts usually have clauses for extreme oil price scenarios (S-curve), activated by the oil price fluctuations during this time.

Table 5.4 Convergence Test Results for 01:2001-08:2008.

<b>Clustering Algorithm Implementation</b>			
Numerical values are $t_{b_k}$ statistics for each step in the algorithm			
Step 1: Ordering	Order the members of the panel according to their final values: $P_{1T} > P_{2T} > \dots > P_{8T}$		
Step 2: Form core groups	Iteration I	Iteration II (Step 4)	Iteration III
Cumulatively moving down.			
Optimal cut-off point criterion:	$P_{BRENT}$	(Base)	-
	$P_{JPN}$	-21.07	(Base)
$k^* =$	$P_{RUS}$	-31.83	1.20
$\text{ArgMax}_k\{t_{b_k}\}$ subject to	$P_{NBP}$	-23.78	-0.28
$\text{Min}_k\{t_{b_k}\} > -1.65$	$P_{HH}$	-20.34	-0.41*
Core groups	Core group	$P_{BRENT}$	$P_{JPN} + P_{RUS} + P_{NBP} + P_{HH}$
Step 3: Sieving			
Add one member at a time to Core	$P_{BRENT}$	-	-
Group. Keep member iff $t_{b_k} > c^*$ ,	$P_{JPN}$	-21.07	-
where $c^* = 0$ is adopted.	$P_{NBP}$	-12.00	-
	$P_{RUS}$	-3.94	-0.28
	$P_{HH}$	-25.06	-3.12
Final Cluster ("Club") Formations	Club 1: $P_{BRENT}$		
	Club 2: $P_{JPN} + P_{RUS} + P_{NBP} + P_{HH}$		

### 5.4.3 The 2010-2020 Sub-period with New Gas Price Benchmarks

This analysis incorporates data from two emerging benchmark spot markets: the Title Transfer Facility (TTF) in the Netherlands and the Asian Spot (AS) gas prices. Data availability limits the chosen timeframe for the study, deliberately omitting the periods affected by the global financial crisis (GFC) and the COVID-19 pandemic.

Surprisingly, this sub-period disrupts many previously identified trends of convergence. According to Table 5.5, there is no convergence in natural gas price outside of Europe. The data supports that gas prices in the UK's National Balancing Point (NBP), the Netherlands' TTF, and Russia converge. This aligns with earlier research, suggesting that incorporating additional European gas prices into the study would likely reinforce this trend of convergence. In contrast, the rest of the prices in the study display distinct trends, illustrated in Figure 5.4, effectively challenging the hypothesis that liquefied natural gas (LNG) arbitrage leads to price convergence.

Moreover, the AS spot price in Asia does not show convergence with Japan's LNG import price. Similarly, Japan's LNG import price does not converge with oil prices.

Table 5.5<sup>7</sup> Convergence Test Results for 01:2010-02:2020.

Clustering Algorithm Implementation					
Numerical values are $t_{b_k}$ statistics for each step in the algorithm					
Step 1: Ordering	Order the members of the panel according to their final values: $P_{1T} > P_{2T} > \dots > P_{8T}$				
Step 2: Form core groups		Iteration I	Iteration II (Step 4)	Iteration III (Step 4)	Iteration IV (Step 4)
Cumulatively moving down. Optimal cut-off point criterion:	$P_{BRENT}$	(Base)	-	-	-
	$P_{JPN}$	-5.23	(Base)	-	-
$k^* =$	$P_{AS}$	-26.64	-28.30	(Base)	-
ArgMax $_k\{t_{b_k}\}$ subject to	$P_{RUS}$	-28.17	-25.58	-17.31	(Base)
Min $_k\{t_{b_k}\} > -1.65$	$P_{NBP}$	-21.79	-17.79	-7.91	5.34
	$P_{TTF}$	-18.77	-14.82	-7.83	0.43*
	$P_{HH}$	-29.54	-28.96	-26.35	-28.67
Core groups	Core group	$P_{BRENT}$	$P_{JPN}$	$P_{AS}$	$P_{RUS}$ + $P_{NBP}$ + $P_{TTF}$

<sup>7</sup> Table 5.5 continues on the next page.

Table 5.5 Convergence Test Results for 01:2010-02:2020, Continued.

Clustering Algorithm Implementation					
Numerical values are $t_{\widehat{D}_k}$ statistics for each step in the algorithm					
Step 3: Sieving					
	$P_{BRENT}$	-	-	-	-
Add one member at a time to Core	$P_{JPN}$	-5.23	-	-	-
Group. Keep member iff $t_{\widehat{D}_k} > c^*$ ,	$P_{AS}$	-26.24	28.30	-	-
where $c^* = 0$ is adopted.	$P_{RUS}$	-29.80	-30.46	-17.31	-
	$P_{NBP}$	-18.88	-3.60	-3.16	-
	$P_{TTF}$	-17.31	-13.94	-6.33	0.43
	$P_{HH}$	-37.99	-32.71	-24.58	-28.67
Final Cluster ("Club") Formations					
	Club 1: $P_{BRENT}$				
	Club 2: $P_{JPN}$				
	Club 3: $P_{AS}$				
	Club 4: $P_{RUS} + P_{NBP} + P_{TTF}$				
	Club 5: $P_{HH}$				
Club-merging $t_{\widehat{D}_k}$					
(Step 5)	Club 4 + Club 1: -19.28 (Not Convergent)				
	Club 4 + Club 2: -16.33 (Not Convergent)				
	Club 4 + Club 3: -7.83 (Not Convergent)				
	Club 4 + Club 5: -16.33 (Not Convergent)				

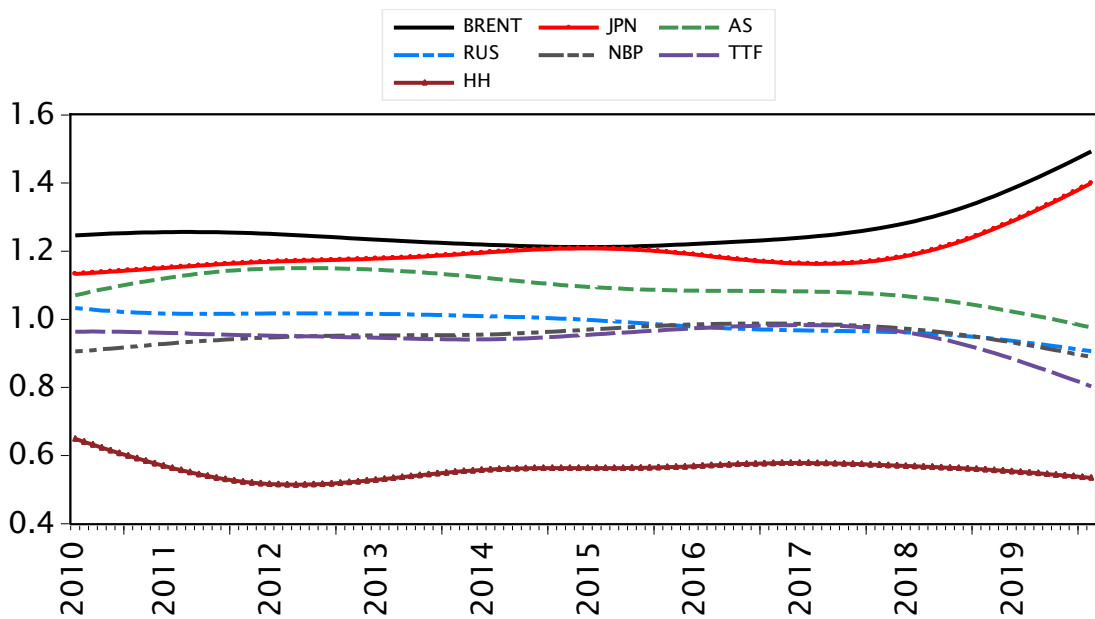


Figure 5.4 Relative Convergence Coefficients for the 2010-2019 Panel Data of 8 Stacked Prices.

## 5.5 Robustness Check: Kalman Filter (Law of One Price Estimates)

Building upon the previous findings, this section examines the data using the perspective of the Kalman-filter LOOP methodology, a method extensively acknowledged in previous studies. This analysis employs a pairwise approach based on stricter assumptions, resulting in outcomes that are more limited in scope.

The main idea is to estimate the constant  $\gamma_{ij}$  and the time-varying coefficient  $\beta_{ij,t}$  in the following pairwise model:

$$P_{i,t} = \gamma_{ij} + \beta_{ij,t}P_{j,t} + \varepsilon_{it}, \varepsilon_{it} \sim nid(0, H_t) \quad (5.8)$$

This model is designed to assess the validity of different versions of the Law of One Price (LOOP). The LOOP holds true without transaction costs if  $\gamma_{ij} = 0$  and  $\beta_{ij,t} = 1$ . However, when accounting for transportation, liquefaction, gasification, and storage costs, represented by  $(a_{ij} + \varepsilon_{it})$ , a less strict form of the LOOP focuses on verifying whether  $\beta_{ij,t} = 1$ . This condition suggests relative and growth convergence, meaning, as time approaches infinity, the expected ratio of prices approaches to 1 ( $\lim_{t \rightarrow \infty} E[P_{it}/P_{jt}] = 1$ ), and the difference in their growth rates approaches to 0 ( $\lim_{t \rightarrow \infty} E[\Delta P_{it} - \Delta P_{jt}] = 0$ ), respectively. However, it does not imply that the absolute level of prices will converge ( $\lim_{t \rightarrow \infty} E[P_{it} - P_{jt}] = \gamma_{ij} + \varepsilon_{it}$ ). The parameter  $\beta_{ij,t}$  indicates a co-movement or non-linear correlation between prices. When beta is not equal to 1, it does not have a specific interpretation.

The insights from Equation (5.8) can be applied to model cointegration. A basic method involves a two-step process where  $e_{i,t}$ , defined as  $e_{i,t} = [P_{i,t-1} - \gamma_{ij} - \beta_{ij,t}P_{j,t-1}]$ , acts as a non-linear error-correction term, assuming it is stationary. However, this thesis opts not to use this approach, as it does not add substantial new information.

To calculate the time-varying coefficient  $\beta_{ij,t}$  over time, we use the Kalman Filter technique (Kalman 1960, Kalman and Bucky 1961). We reformulate Equation (5.8) into a standard state-space model, utilizing vectors  $\pi_t = [\gamma_{ij,t} \ \beta_{ij,t}]$  and  $z_t = [1 \ P_{i,t}]$ :

$$P_{i,t} = z_t' \pi_t + \varepsilon_{it}, \varepsilon_{it} \sim nid(0, H_t) \quad (5.9)$$

$$\pi_t = \pi_{t-1} + \eta_{it}, \eta_{it} \sim nid(0, Q_t) \quad (5.10)$$

In this formulation,  $\pi_t$  represents the state variable, conceptualized as a random walk to facilitate inference on time-varying elements. Equation (5.10) acts as the measurement equation. During the *prediction step*, estimates for  $\pi_t$  and its covariance matrix  $V_t$  are derived using historical data up to time  $(t - 1)$ , starting from initial values  $\pi_0$  and  $V_0$ , as detailed in Harvey (1989).

$$\hat{\pi}_{t|t-1} = \hat{\pi}_{t-1} \quad (5.11)$$

$$V_{t|t-1} = V_{t-1} + Q_t \quad (5.12)$$

$$V_{t|t-1} = E[(\pi_{t-1} - \hat{\pi}_{t-1})(\pi_{t-1} - \hat{\pi}_{t-1})'] \quad (5.13)$$

Following this, the filtering step occurs, during which the predictor for the state variable at time  $t$  is updated to incorporate information available up to that point, as described below:

$$\hat{\pi}_t = \hat{\pi}_{t|t-1} + K_t(P_{j,t} - z_t' \hat{\pi}_{t|t-1}) \quad (5.14)$$

$$V_t = V_{t|t-1} - K_t z_t' V_{t|t-1} \quad (5.15)$$

In this context,  $K_t$ , representing the Kalman Gain, is calculated as  $K_t = V_{t|t-1} z_t' (z_t' V_{t|t-1} z_t' + H_t)^{-1}$ . The prediction error for one step ahead is defined as  $v_t = P_{j,t} - z_t' \hat{\pi}_{t|t-1}$ . To determine the parameter estimates, the log-likelihood function is maximised, a method supported by the works of Kalman (1960) and Harvey (1989):

$$\log L = \sum_{t=1}^T \log(\hat{\pi}_t' V_{t|t-1} \hat{\pi}_t + H_t) + T \log \frac{1}{T} \sum_{t=1}^T \frac{(P_{j,t} - z_t' \hat{\pi}_{t|t-1})^2}{(\hat{\pi}_t' V_{t|t-1} \hat{\pi}_t + H_t)} \quad (5.16)$$

The insights derived from the estimates produced by the Kalman Filter are relevant to the Phillips-Sul approach. These can be compared by integrating Equations (5.1) and (5.8):

$$P_{it} = \underbrace{\left(\frac{a_{it} + x_{it}}{\mu_t}\right)\mu_t}_{\text{Phillips-Sul}} = \delta_{it}\mu_t \cong \underbrace{\gamma_{ij} + \beta_{ij,t}P_{j,t} + \varepsilon_{it}}_{\text{Weak LOOP via Kalman Filter}} \quad (5.17)$$

The comparison between the Phillips-Sul method and the Kalman Filter approach highlights two primary differences. Firstly, the Phillips-Sul method posits that the price  $P_{it}$  adheres to a long-term trend  $\mu_t$ , which could either be another price  $P_{j,t}$  as seen in the Kalman Filter method or an average of various prices within the panel. Secondly, a notable distinction arises if we assume  $\mu_t = P_{j,t}$  under the Phillips-Sul method in that the Kalman Filter suggests a linear functional form. This presupposition is more stringent than Phillips-Sul's allowance for potential non-linearity in  $a_{it}$  and  $x_{it}$ , which are neither directly observed nor identified. Instead, Phillips-Sul utilizes the metric  $\delta_{it}$  for statistical inference regarding long-term trends within the data, which can exist amidst temporary divergences. Despite these constraints, the Kalman Filter's Weak LOOP approach proves instrumental in examining shifts in transaction costs ( $a_{ij} + \varepsilon_{it}$ ), a topic we explore in Section 5.6.

Figure 5.5 presents the estimates from the Kalman Filter analysis of  $\beta_{ij,t}$  across various pairs of traditional gas price benchmarks and the oil price. These estimates offer a dynamic and non-linear correlation, suggesting co-movement towards unity, where the slope determines the degree of convergence. The findings corroborate earlier research, showing a tendency for the NBP (UK's National Balancing Point) and Russian gas prices to align closely. In contrast, the Henry Hub shows a weak correlation with other benchmarks. Prices for Japan, Brent crude, NBP, and Russian gas all move together within the 0.8 to 1 range, indicating co-movement. However, these time-varying estimates do not reveal the number of distinct long-term convergence trends among these series, marking a limitation of this method. This gap was addressed through formal analysis in the preceding section, establishing that only the NBP and Russian prices share a common convergence trend.

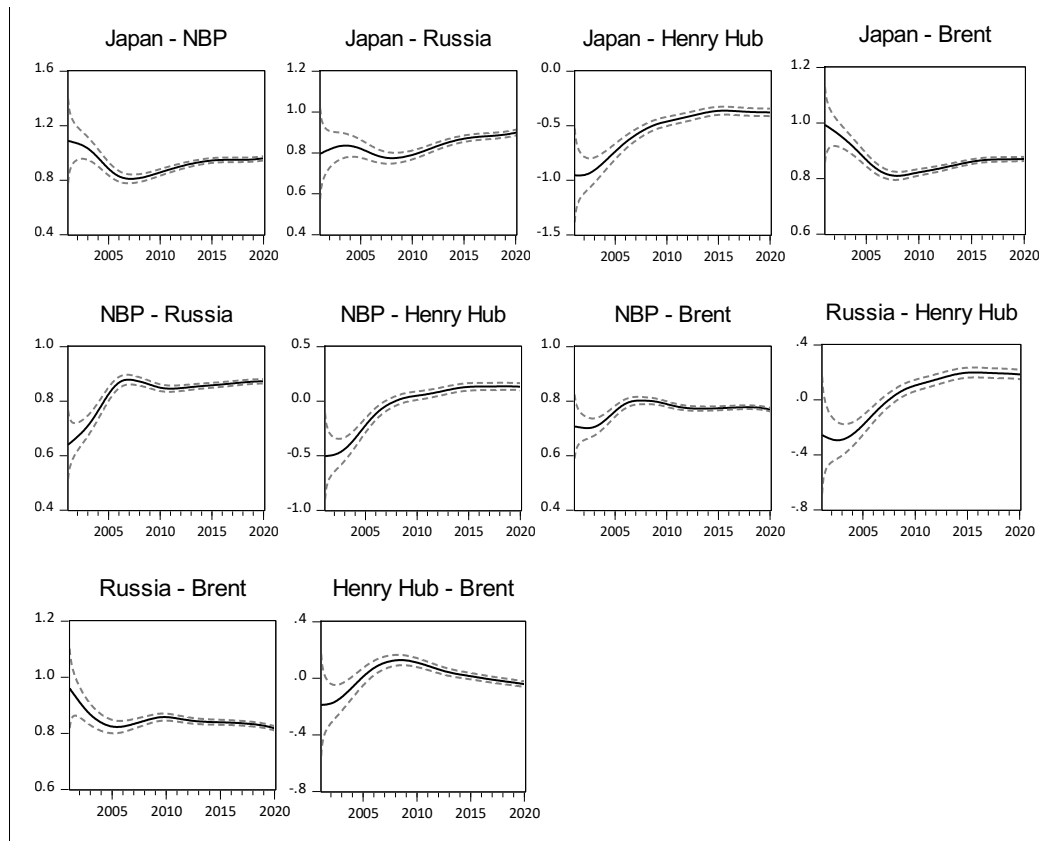


Figure 5.5 Traditional Gas Price Benchmarks and Oil Price: Kalman-Filter Estimates with 2-Standard Deviation Prediction Bands.

Figure 5.6 presents the dynamic co-movement estimates for emerging benchmarks, specifically AS (Asian Spot) and TTF (Title Transfer Facility in Europe). It shows that the AS price does not align with Japan's LNG import price or any others, supporting the idea that regional demand and supply dynamics predominantly influence spot LNG trading in Asia. The TTF price also does not demonstrate clear convergence with any of the prices presented in Figure 5.5. A notable limitation of this pairwise method is its inability to identify a unified trend towards which all European prices might converge, which was established in Section 5.4. The analysis highlights a period of divergence among European prices during 2018-2020, suggesting an area of interest for future research. This divergence underscores gas price benchmarks' complex and evolving nature, indicating the need for continuous monitoring and analysis.



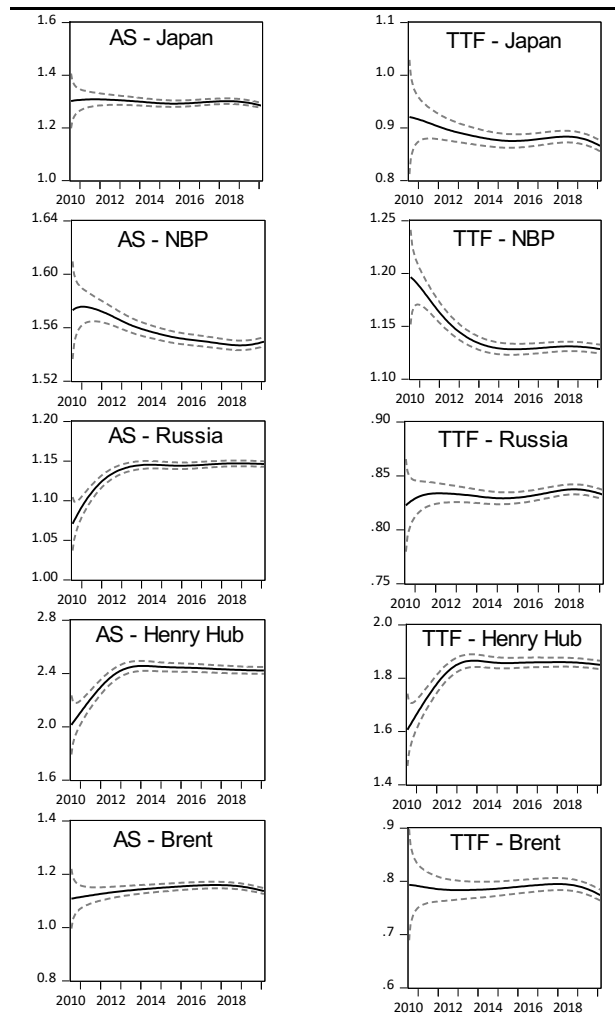


Figure 5.6 The New Gas Price Benchmarks (AS and TTF): Kalman-Filter Estimates with 2-Standard Deviation Prediction Bands.

## 5.6 Discussion of the Results

The key observation that there is no integration across transoceanic gas markets necessitates a retrospective analysis of prior convergence trends. These past trends can be partially attributed to oil-indexation and to similar regional demand and supply dynamics. For instance, the historical linkage between Japan's gas prices and Brent oil prices is a case in point. During 2010-2015, Japanese gas prices rose sharply in tandem with oil prices, which were around \$100 a barrel, and then fell significantly after 2015 as oil prices declined. Although spot trading has become more prevalent in Asia, replacing government-regulated prices, the proportion of oil-indexed pricing in the region's market has grown compared to two decades ago.

The post-2015 convergence of European and Asian gas prices can be linked to a surge in LNG production, leading to a global oversupply and lower prices. In Europe, gas-on-gas competition, particularly at the TTF hub, has become a significant pricing factor, moving away from the oil-indexation that dominated earlier. This shift was further supported by the construction of numerous LNG regasification facilities across Europe, enhancing the continent's access to global gas markets and reducing its dependence on Russian gas, which was traditionally priced against oil products. Consequently, the oil-linked pricing formulas in LNG and pipeline sales contracts have been largely replaced by hub-based pricing, such as the Dutch TTF and the UK NBP. In the US, gas prices have remained relatively low and disconnected from other markets due to the boom in unconventional gas production during the 2010s. Domestic production exceeded consumption, but it was not until the late 2010s that the US began exporting significant amounts of LNG, marking a shift in its role in the global gas market.

This thesis also explores whether the costs associated with transportation, storage, liquefaction, and gasification can account for the observed lack of growth convergence in transoceanic gas prices. While a detailed analysis goes beyond the scope of this thesis, the preliminary conclusion is that it is highly unlikely. A scenario where transaction costs steadily increase over the long term, which could explain the lack of convergence, contradicts available evidence. Temporary cost shocks might lead to short-term divergence but are insufficient to justify a sustained lack of convergence over the long term. This perspective is reinforced by research from Oglend, Kleppe, and Osmundsen (2016), which developed a model to estimate gas economic profit spreads, incorporating cost factors like freight, distance, LIBOR rates (as a proxy for expected returns), and the NBER recession indicator (reflecting global economic conditions), while treating liquefaction, regasification, and storage costs as fixed constants in the model. Both linear and quadratic models for cost functions were tested. The findings underscore that variations in price spreads between markets like the US-Europe, Europe-Asia, and US-Asia are largely influenced by factors other than transaction costs, which tended to remain stable over time.

## 5.7 Concluding Remarks

The findings in this chapter provide a critical insight for energy analysts, researchers, industry stakeholders, and policymakers: there is no formal evidence supporting the existence of price convergence across transoceanic natural gas markets. Convergence is only found within European gas prices, rejecting the hypothesis that LNG trade serves as a global price equaliser. The investigation replicates temporary convergence patterns previously reported, which were tied to now-changed market conditions. This result is confirmed by incorporating two new benchmark prices for natural gas and deliberately excluding the volatile COVID-19 period. Additionally, the results show that considering the period after the Global Financial Crisis, natural gas prices do not align with oil prices, indicating a divergence of the oil-gas price linkage.

The findings suggest that despite moving away from oil price dependence, the current volume of LNG trade needs to be increased to act as an effective mechanism for price arbitrage. This leads to the anticipation that increased LNG trade volumes may ultimately foster price convergence in the future.

# CHAPTER 6

## REVIEW OF ECONOMETRIC METHODS

### 6.1 Introduction

Economic datasets are often presented in dynamic multivariate time series, where contemporary or sequential relationships may exist in the short-run or the long-run. Additionally, variables can be dependent and/or independent within such frameworks. This thesis aims to examine how natural gas prices across various regional markets are linked. The main challenge in modelling the relationship between time series in a multivariate analysis is finding the correct representation of the model, as highlighted by Brandt and Williams (2007). This chapter will present various methodological approaches to modelling multiple time series, discussing the implementation and implications of each.

To facilitate model development, the literature in this area employs four key related methods for analysing multivariate time series data, namely: Autoregressive Integrated Moving Average (ARIMA) models, Simultaneous or Structural Equation (SEQ) systems, Error Correction Models (ECM), and Vector Autoregression (VAR) models (Brandt and Williams 2007). Additionally, variants like the Autoregressive Distributed Lag (ARDL) models have been used to accommodate these primary methods for the unique aspects of the multivariate dataset. These econometric approaches share a common mathematical foundation, employing various forms of linear regression, including ordinary least squares, generalised least squares, and multiple least squares. Yet, they are distinguished by their core assumptions and the foundational elements that critically influence the interpretation of their outcomes.

The theory exposition in this chapter is centred around VAR modelling and includes an extensive discussion of single-equation ECMs and ARDL models as special cases. Similar procedures are used in all cases for lag selection and to make causality inferences.

The chapter reviews the following econometric methods. In Section 6.2, a brief introduction to the primary methods is presented. Section 6.3 introduces the main concepts of the VAR analysis. Section 6.4 presents a modification of the VAR model, the vector error correction model (VECM), and the concept of cointegration among time series. Sections 6.5 and 6.6 investigate a different method to assess cointegration between time series using the linear and asymmetric ARDL model approach. At the end of the chapter, Section 6.7, we present notable indicators of global economic activity and market fundamentals relevant to modelling oil prices in recent literature. These indicators can also enhance the econometric models used in this thesis, with natural gas prices as endogenous variables. Finally, Section 6.8 presents concluding comments.

## **6.2 Introduction to Multivariate Time Series Models**

### **6.2.1 Simultaneous Equation Method**

The Simultaneous Equation (SEQ) modelling approach involves transforming a singular economic theory into a set of equations, laying the groundwork for building a multi-equation time series model. This process builds on the theory's delineation of relationships between variables, categorizing them as either exogenous or endogenous within the system. Leamer (1985) states that an exogenous variable,  $x$ , influences another variable,  $y$ , without its value being affected by any alterations in  $y$ . Essentially, exogenous variables are predetermined and external to the equation system. Conversely, endogenous variables are integral to and dependent within the system, and the equation framework determines their values. A single structural system of equations represents the relationship between the variables (Brandt and Williams 2007).

Using an example from Brooks (2008), where SEQ is used to model the quantity of new houses sold, the set of structural equations are defined as:

$$Q = \alpha + \beta P + \gamma S + u \quad (6.1)$$

$$Q = \lambda + \mu P + kS + v \quad (6.2)$$

where;  $Q$  is the number of new houses,  $P$  is the average price of new houses,  $S$  is the price of older houses,  $T$  is a variable containing the state of housebuilding technology, and  $u$  and  $v$  are the residuals.

In Equations (6.1) and (6.2), the price depends on quantity, and so does quantity on price. Therefore,  $Q$  and  $P$  are considered endogenous, while  $S$  and  $T$  are exogenous to the system of equations. To estimate the parameters in the equations by an Ordinary Least Squares model (OLS), they must be represented in the reduced form, where all exogenous variables are on the right-hand side of the equation. The reduced form of Equations (6.1) and (6.2) are:

$$Q = \frac{\mu\alpha - \beta\lambda}{\mu - \beta} - \frac{\beta k}{\mu - \beta} T - \frac{\mu\gamma}{\mu - \beta} S + \frac{\mu u - \beta v}{\mu - \beta} \quad (6.3)$$

$$P = \frac{\lambda - \alpha}{\beta - \mu} + \frac{k}{\beta - \mu} T - \frac{\gamma}{\beta - \mu} S + \frac{v - u}{\beta - \mu} \quad (6.4)$$

If we rewrite Equations (6.3) and (6.4) as:

$$Q = \pi_{10} - \pi_{11}T - \pi_{12}S + \varepsilon_1 \quad (6.5)$$

$$P = \pi_{20} - \pi_{21}T - \pi_{22}S + \varepsilon_2 \quad (6.6)$$

where the  $\pi_i$  coefficients are the combinations of the original coefficients in (6.5) and (6.6). Rather than obtaining the demand and supply equations by re-parametrisation (something that can only be done in exactly identified models), authors typically rely on a Two-Stage-Least-Square method in which predicted values from the OLS reduced-form equations are substituted into the demand and supply equations. A GMM can also be used to obtain parameter estimates.

A notable challenge in constructing Simultaneous Equation (SEQ) models is the need for users to select variables and their lags to ensure model identification. This process often employs hypothesis testing to determine which variables to exclude, a practice that can introduce bias into the final model. Sims (1980) highlighted a critical concern: such exclusion restrictions might lack theoretical support and fail to be corroborated by empirical evidence, potentially causing a misrepresentation of the SEQ model's dynamics.

Generally, vector autoregressors (VAR) are preferred over SEQ to model simultaneous economic systems. This is because the former provides a much richer collection of dynamic insights and superior forecasting power in general.

### **6.2.2 ARIMA Method**

The ARIMA model, also known as the Box-Jenkins approach, proposed by Box and Jenkins (1970), analyses multiple time series by treating them as individual univariate entities influenced by exogenous variables within the regression framework. The core objective of the ARIMA methodology is to identify the most suitable and streamlined model for each time series, ensuring the residuals are uncorrelated, which involves distinguishing between endogenous and exogenous variables within the model's structure. A key advantage of ARIMA, compared to the Simultaneous Equation (SEQ) method, is its superior forecasting capability, attributed to its emphasis on the principle of parsimony.

The ARIMA model is primarily univariate; hence, it attributes the behaviour of most time series to its historical data, overlooking the potential influence of other variables. This approach prioritizes the analysis of individual time series without adequately addressing the broader interactions among them. Furthermore, a significant drawback arises from the model's requirement for one equation per time series, necessitating perfect independence among these equations to accurately estimate parameters. This condition can be challenging to meet in practice, affecting the model's efficiency in parameter estimation.

### **6.2.3 Error Correction Model**

Brandt and Williams (2007) describe Error Correction Models (ECM) as an extension of ARIMA and SEQ approaches, leveraging the autoregressive distributed lag (ADL) framework for analysing two or more time series that may share long-term relationships. Unlike ARIMA, which focuses on a time series past values, ECM explicitly accounts for the dynamic interplay between multiple series, capturing both their long-term equilibrium and short-term fluctuations. The process begins with testing for stationarity using unit root tests to ensure the time series are suitable for analysis. Upon confirming stationarity, the modeller employs specialised techniques

to explore the long-term relationships between the series. Subsequently, regression analysis is applied to model the short-term deviations around this established long-term trend, offering insights into how the series interact over time in the short and long run.

The ECM can be applied as a vector error correction model (VECM) in a multivariate system. In this case, the model may show a variety of long-run and short-run dynamics across all different time series. The ECM can also be applied to a single regression equation; typically, this is done in Autoregressive Distributed Lag (ARDL) models.

#### **6.2.4 Vector Autoregression Method**

The Vector Autoregression (VAR) model combines elements of SEQ and ARIMA models. It is the most encompassing framework for analysing multivariate time series, prioritising the relationships and dynamics among series over their generative structures. Originating from the work of Sims (1972, 1980), the VAR approach models each variable as influenced by its historical values and those of every other variable in the analysis. Sims (1980) emphasised that the efficacy of VAR models hinges on carefully selecting relevant variables and determining appropriate lag lengths based on empirical data. Unlike the SEQ models, which necessitate the exclusion of certain variables and lags for identification, the VAR model avoids potential bias from omitted variables by including all variables and their lags across all equations. While this inclusive strategy helps prevent bias, it also leads to a more complex model with more parameters to estimate, potentially affecting the model's efficiency.

In the upcoming sections, we aim to offer a comprehensive examination of the Vector Autoregression (VAR) model and its restricted version, the Vector Error Correction Model (VECM). These multivariate time series frameworks are foundational for uncovering causal relationships between time series data.

### **6.3 Vector Autoregression Model**

Sims (1980) argues that in simultaneous relationships among a group of time series, the traditional separation of variables into endogenous and exogenous categories becomes irrelevant; instead, all variables should be treated as endogenous. Brandt and Williams (2007) describe the Vector Autoregression (VAR) model as an



"interdependent reduced form dynamic model". A basic illustration of this concept is the bivariate VAR model, which involves two variables,  $y_{1t}$  and  $y_{2t}$ , as detailed by Brooks (2008):

$$y_{1t} = \beta_{10} + \beta_{11}y_{1t-1} + \dots + \beta_{1k}y_{1t-k} + \alpha_{11}y_{2t-1} + \dots + \alpha_{1k}y_{2t-k} + u_{1t} \quad (6.7)$$

$$y_{2t} = \beta_{20} + \beta_{21}y_{2t-1} + \dots + \beta_{2k}y_{2t-k} + \alpha_{21}y_{1t-1} + \dots + \alpha_{2k}y_{1t-k} + u_{2t} \quad (6.8)$$

The Vector Autoregression (VAR) model. It is a flexible system that includes multiple time series variables, with each variable's equation incorporating its lagged values and those of all other variables in the model. This results in each equation having  $f(l + 1)$  coefficients, where  $l$  is the number of lags and  $f$  is the number of variables. Brandt and Williams (2007) describe how the VAR model efficiently encapsulates these dynamics in a compact mathematical form:

$$y_t = c + \sum_{l=1}^p y_{t-l}\beta_l + e_t \quad (6.9)$$

where  $y_{t-l}$  is the  $1 \times f$  vector of  $l$ th lagged variables,  $\beta_l$  is the  $f \times f$  matrices of the coefficients for the  $k$ th lag,  $c$  is the  $1 \times f$  vector of the  $\beta_{i0}$  intercepts, and  $e_t$  is the  $1 \times f$  vector of residuals.

As a non-specific general framework, the Vector Autoregression (VAR) model has versatile applications in economic modelling. Brandt and Williams (2007) identify three primary uses of the VAR model: exploring causality among endogenous variables, analysing the dynamic effects of changes in one variable on others within the system, and assessing how changes within each variable and across other variables contribute to its overall variance. These different interpretative approaches provide a multifaceted understanding of the interactions within the VAR model and will be examined further in this section. Before exploring these applications, an essential step in VAR model estimation involves selecting an optimal lag length for the endogenous variables in the regression system, ensuring the model's efficacy and accuracy.

### 6.3.1 Optimal VAR Lag Length Selection

To accurately select the optimal number of lags for each time series within a VAR model, Brandt and Williams (2007) suggest employing two hypothesis tests. The first test is the likelihood ratio test, which involves comparing the likelihood values derived from VAR models with different lag lengths, specifically comparing a model with  $k_1$  lags against another with fewer lags,  $k_0$ , where  $k_1 > k_0$ . This comparison allows for an assessment of how additional lags affect the model's likelihood, aiding in determining the most appropriate lag length. The likelihood comparison between models with  $k_1 > k_0$  lags can be expressed as:

$$(T - 1 - f k_1)(\log|\widehat{\Sigma}_0| - \log|\widehat{\Sigma}_1|). \quad (6.10)$$

where  $\widehat{\Sigma}_i$  is the error covariance of the VAR containing  $k_i$  lags. The null and alternative hypothesis of the likelihood ratio test or  $\chi^2$  test are:

- $H_0$ : The VAR model has  $k_0$  lags.
- $H_A$ : The VAR model has  $k_1$  lags.

The  $\chi^2$  is based on the number of restrictions equal to the degrees of freedom; thus, the t-statistics is asymptotically distributed (Brooks 2008). The test is distributed  $\chi^2$  with  $f^2(k_1 - k_0)$  degrees of freedom. If the restricted model is accepted in the hypothesis test, the number of lags excluded from the unrestricted model ( $k_1$  lags) will be equal to  $f^2(k_1 - k_0)$ . The disadvantage of the likelihood ratio test is that it requires the errors from each equation to be normally distributed, which is rare in most economic data.

The second approach for determining the optimal number of lags in a VAR model involves using information criteria (IC). This method, which is also crucial for establishing the optimal lag length in other econometric tests like the ADF and PP unit root tests (as discussed in Chapter 4), offers a significant advantage by not necessitating the normality assumption required by the likelihood ratio test. It works by incorporating a penalty for each additional lag in the model. The methodology is quantified through three distinct information criteria equations, each tailored to

balance the trade-off between model complexity and fit. The equations that represent the three different information criteria (IC) are:

- Akaike's IC:  $MAIC = \log|\widehat{\Sigma}| + \frac{2k'}{T}$  (6.11)

- Bayesian's IC:  $MBIC = \log|\widehat{\Sigma}| + \frac{k'}{T} \log(T)$  (6.12)

- Hannan-Quinn IC:  $MHQIC = \log|\widehat{\Sigma}| + \frac{2k'}{T} \log(\log(T))$  (6.13)

where  $\widehat{\Sigma}$  is the error covariance matrix of the VAR,  $T$  is the number of observations, and  $k'$  is the number of regressors in all equations, also presented as  $f(kf + 1)$  for a VAR system with  $k$  lags and  $f$  variables. The value for each IC is given for each number of lags (0, 1, ...,  $k$ ). The optimal lag length is selected from the minimum value of the given IC. As presented in the literature, the most common IC used is Akaike's Information Criterion (AIC). Therefore, when applying the IC method to select the optimal lag length required for any econometric method in this thesis, the AIC will be used.

### 6.3.2 Granger Causality Tests

Understanding how lagged variables influence the dependent variable is crucial in analysing an unrestricted Vector Autoregression (VAR) model. This concept explores the extent to which a lagged variable  $A_t$  can predict future values of another variable  $B_t$ . Pioneering work by Granger (1969) and Sims (1972) first investigated these interrelationships within multiple time series. Granger's contribution led to what is now known as the Granger causality test, which assesses the sequential predictive power of one time series over another within a multiple time series framework. A bivariate VAR model is typically employed to define and apply this test formally:

$$A_t = \kappa_0 + \sum_{i=1}^n \kappa_i A_{t-i} + \sum_{i=1}^n \lambda_i B_{t-i} + u_{1t} \quad (6.14)$$

$$B_t = \lambda_0 + \sum_{i=1}^n \mu_i A_{t-i} + \sum_{i=1}^n \nu_i B_{t-i} + u_{2t} \quad (6.15)$$

From Equation (6.14), if the coefficients of the lagged values of  $B_t$  are nonzero ( $\lambda_i \neq 0$ , for  $i = 1, 2, \dots, n$ ), then  $B_t$  Granger causes  $A_t$ . From Equation (6.15), the logic is the same. If the coefficients of the lagged values of  $A_t$  are nonzero ( $\mu_i \neq 0$ , for  $i = 1, 2, \dots, n$ ), then it is said that  $A_t$  Granger causes  $B_t$ . In this case, the Granger causality test investigates if past values of the variable  $A_t$  ( $A_{t-1}, A_{t-2}, \dots, A_{t-n}$ ) impacts the contemporaneous value of  $B_t$ . The test consists of an  $F$ -test or a  $\chi^2$ -test for a joint hypothesis that a regressor does not cause changes in the dependent variable, considering one equation in the VAR system. Therefore, from Equation (6.14), the null hypothesis of the Granger causality test can be described as:

- $H_0$ :  $B_t$  does not Granger cause  $A_t$  if  $\lambda_1 = \lambda_2 = \lambda_3 = \dots = \lambda_n = 0$ .
- $H_A$ :  $B_t$  Granger causes  $A_t$  if  $\lambda_1 \neq 0$ ,  $\lambda_2 \neq 0$ , ..., or  $\lambda_n \neq 0$ .

Brooks (2008) elaborates on the concept of Granger causality, explaining that if the lagged values of  $B_t$  significantly affect  $A_t$  without the reverse occurring, it indicates a one-way causality from  $B_t$  to  $A_t$ , making  $B_t$  strongly exogenous as per Equation (6.14). The reverse scenario, detailed in Equation (6.15), also holds. However, if lagged values from both time series significantly influence each other, this suggests a two-way or bidirectional causality. Conversely, if neither set of lagged values has a significant impact, then variables  $B_t$  and  $A_t$  are deemed independent, indicating no causal relationship between them.

Brooks (2008) states that Granger causality should not be interpreted as a literal causal relationship where one variable directly influences another. Rather, it is a statistical concept used in VAR models to identify correlations between past and present observations of time series. The essence of Granger causality lies in determining the statistical significance of lagged variables in predicting changes in dependent variables within the model.

### 6.3.3 Impulse Responses in a VAR Model

A key aspect of analysing VAR models involves examining the dynamic interactions between variables, specifically how changes in one variable might positively or negatively influence others and the duration of these effects. To assess these dynamics, analysts employ impulse response functions and variance decompositions, which are

instrumental in understanding the temporal effects of variables on each other within the VAR framework.

Impulse responses in a VAR model measure how dependent variables react to shocks in each of the model's regressors (Brooks 2008). The approach involves introducing a one-unit shock to the error term of each dependent variable in each equation and observing the effect on the entire VAR system over time. In a VAR model with  $y$  variables, this results in  $y^2$  possible impulse responses across all equations. By conceptualizing the VAR as a vector moving average (VMA), it is understood that these shocks will gradually decrease as the system returns to stability. Brandt and Williams (2007) note that viewing the VAR through a VMA lens helps clarify its structural dynamics, allowing for an analysis of how it reacts to external shocks in the error terms. Brooks (2008) illustrates this with an example of a bivariate VAR(1) model, which includes just one lag.

$$y_t = A_1 y_{t-1} + u_t \quad (6.16)$$

where  $A_1 = \begin{bmatrix} 0.5 & 0.3 \\ 0.0 & 0.2 \end{bmatrix}$

The general form of the bivariate VAR in (5.16) can also be expressed using vectors and matrices as follows:

$$\begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} = \begin{bmatrix} 0.5 & 0.3 \\ 0.0 & 0.2 \end{bmatrix} \begin{bmatrix} y_{1t-1} \\ y_{2t-1} \end{bmatrix} + \begin{bmatrix} u_{1t} \\ u_{2t} \end{bmatrix} \quad (6.17)$$

Following the example from Brooks (2008), if a unit shock is applied in  $y_{1t}$  at time  $t = 0$ , the impacts in the VAR system at time  $t = 0, 1, 2$  can be described as follows:

$$y_0 = \begin{bmatrix} u_{10} \\ u_{20} \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad (6.18)$$

$$y_1 = A_1 y_0 = \begin{bmatrix} 0.5 & 0.3 \\ 0.0 & 0.2 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0.5 \\ 0 \end{bmatrix} \quad (6.19)$$

$$y_2 = A_1 y_1 = \begin{bmatrix} 0.5 & 0.3 \\ 0.0 & 0.2 \end{bmatrix} \begin{bmatrix} 0.5 \\ 0 \end{bmatrix} = \begin{bmatrix} 0.25 \\ 0 \end{bmatrix} \quad (6.20)$$

Thus, introducing a unit shock to  $y_{1t}$  allows for assessing the impulse response functions for both  $y_{1t}$  and  $y_{2t}$ . The coefficient of  $y_{1t-1}$  in the equation for  $y_{2t}$ , derived

from matrix  $A_1$ , is zero, indicating that a shock to  $y_{1t}$  will invariably not affect  $y_{2t}$ . Similarly, applying a unit shock to  $y_{2t}$  at  $t = 0$  will result in specific impulse response functions:

$$y_0 = \begin{bmatrix} u_{10} \\ u_{20} \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad (6.21)$$

$$y_1 = A_1 y_0 = \begin{bmatrix} 0.5 & 0.3 \\ 0.0 & 0.2 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0.3 \\ 0.2 \end{bmatrix} \quad (6.22)$$

$$y_2 = A_1 y_1 = \begin{bmatrix} 0.5 & 0.3 \\ 0.0 & 0.2 \end{bmatrix} \begin{bmatrix} 0.3 \\ 0.2 \end{bmatrix} = \begin{bmatrix} 0.21 \\ 0.04 \end{bmatrix} \quad (6.23)$$

The examples illustrate that computing impulse response functions in a simple two-variable VAR model are relatively straightforward. However, this method can also be applied to more complex VAR models with more variables and longer lags, where the effects of external shocks can be more complicated. In these scenarios, computational analysis becomes essential to achieve precise outcomes (Brooks 2008).

These simple examples assume no correlation exists between shocks to the variables introduced through the error terms. When shocks are assumed to be contemporaneously correlated, then a structural VAR representation with contemporaneous relationships will emerge. Several techniques are available to identify and estimate the model. This thesis uses a Cholesky-decomposition technique that restricts the matrix of contemporaneous relationships to be upper triangular, based on assumptions. In other words, with a Cholesky-decomposition approach, the modeller has to choose one independent shock, a second shock that will depend on the latter only, a third shock that will depend on the last two, etc. This decision can be informed by economic theory, causality results, and forecast error variance decomposition.

### 6.3.4 Decomposition of Forecast Error Variance

Forecast variance decompositions offer an alternative method for examining the dynamic relationships between variables in a VAR model. This approach quantifies the extent to which fluctuations in each variable within the VAR system can be attributed to changes in the other variables over a specified timeframe. In a VAR system, a shock to one variable can affect others due to the system's inherent structural

dynamics. Variance decompositions measure the contribution of changes in each explanatory variable to the variance of forecast errors for an endogenous variable at future steps ( $s=1, 2, \dots$ ), using the VMA representation to calculate these forecast errors. An example from Brandt and Williams (2007) demonstrates how a VAR(p) model is transformed into a VMA for this purpose.

$$\begin{aligned}
y_t &= c + y_{t-1}B_1 + y_{t-2}B_2 + \dots + y_{t-p}B_p + e_t, \\
y_t - y_{t-1}B_1 - y_{t-2}B_2 - \dots - y_{t-p}B_p &= c + e_t, \\
y_t(I - LB_1 - L^2B_2 - \dots - L^pB_p) &= c + e_t, \\
y_t - d &= e_t(I - LB_1 - L^2B_2 - \dots - L^pB_p)^{-1}, \\
y_t - d &= e_t(I + LC_1 + L^2C_2 - \dots)
\end{aligned} \tag{6.24}$$

where  $L^k$  is the lag operator that transforms any variables in its lagged  $k$  periods:  $L^k y_t = y_{t-k}$ . Furthermore,  $d$  is the VAR constant term ( $c$ ) divided by the AR lag polynomial. The VMA coefficients  $C_i$  in (6.24) are obtained by recursive equations as follows:

$$\begin{aligned}
C_1 &= B_1 \\
C_2 &= B_1C_1 + B_2 \\
C_3 &= B_1C_2 + B_2C_1 + B_3
\end{aligned} \tag{6.25}$$

Considering the VMA form of a VAR(p) model as outlined in equation (6.24), the forecast errors for the VAR system at time period  $s$  are calculated as follows:

$$y_{t+s} - \hat{y}_{t+s} = e_{t+s} + C_1e_{t+s-1} + C_2e_{t+s-2} + \dots + C_{s-1}e_{t+1} \tag{6.26}$$

In Equation (6.26), the left side shows the discrepancy between the actual observed values of the endogenous variables at time  $t + s$  and their predicted values based on VAR estimations. On the right side, Equation (6.26) depicts the forecast errors expressed in the VMA format, tracing from the current time  $T = s$  back to  $T = s - 1$  (Brandt and Williams 2007). This setup illustrates how the VAR model's innovations

are accounted for by their historical values in the VMA context. Additionally, the forecast error variance in Equation (6.26) is defined as follows:

$$\begin{aligned} V(y_{t+s} - \hat{y}_{t+s}) &= E[(y_{t+s} - \hat{y}_{t+s})'(y_{t+s} - \hat{y}_{t+s})] \\ &= \Sigma + C_1 \Sigma C_1' + C_2 \Sigma C_2' + \dots + C_{s-1} \Sigma C_{s-1}' \end{aligned} \quad (6.27)$$

where  $\Sigma = E[e_t' e_t]$  is the covariance of the forecast errors in time  $t$ .

The outcomes of forecasting error decomposition are typically displayed in tables or charts, illustrating the proportion of a variable's forecast error variance attributed to its shocks and those from other variables within the VAR system. It is important to recognise the inherent accuracy limitations of impulse responses and variance decompositions, making the use of confidence intervals around these results indispensable (Runkle 1987).

## 6.4 Cointegration Techniques

This section discusses two cointegration methods: The Engle-Granger method, which applies to single-equation settings, and the Johansen method, which applies to VARs.

Granger and Newbold (1974) highlight that trends in non-stationary data can result in incorrect interpretations. Addressing this issue, the ECM approach is designed to account for these trends and reduce the risk of spurious regression outcomes. The Error Correction Model (ECM) is based on a principle that was formally first studied in Engle and Granger (1987). This article states that a stationary linear relationship may be obtained from a linear combination between two non-stationary time series that share a common trend. Brandt and Williams (2007) outline three key benefits of this approach. Firstly, the ECM facilitates the study of a shared trend and interactions between two or more time series, thereby minimising the chances of encountering spurious regression issues. Secondly, the ECM provides insights into the time series' short-term and long-term dynamics within a single regression framework. Lastly, it allows for assessing causal relationships (Granger causality) among the time series, even when they are integrated or exhibit trends. An example of the ECM includes a model with two I(1) variables and illustrates how these concepts are applied.



$$Y_t = \beta_1 Z_t + \beta_2 Y_{t-1} + \beta_3 Z_{t-1} + \epsilon_t \quad (6.28)$$

where  $\frac{\beta_1 + \beta_3}{1 - \beta_2}$  is the total multiplier that accounts for changes in  $Z_t$ . Engle and Granger (1987) explain that if a linear regression involving  $Z_t$  and  $Y_t$  results in a stationary series, these variables are deemed cointegrated. However, it is not feasible to estimate the overall multiplier directly using Ordinary Least Squares (OLS) due to the nonstationary nature of the residuals in Equation (6.28). To address this issue, the model is adjusted by incorporating the first differences of the variables, thereby transforming it into an error correction model as described below:

$$Y_t = Y_{t-1} + Z_t \beta_1^* - Z_{t-1} \beta_1^* + \beta_2^* Y_{t-1} - \beta_2^* Y_{t-2} + \beta_3^* Z_{t-1} - \beta_3^* Z_{t-2} + u_t, \quad (6.29)$$

$$(1 - L)Y_t = \beta_1^* (1 - L)Z_t + \beta_2^* (1 - L)Y_{t-1} + \beta_3^* (1 - L)Z_{t-1} + u_t \quad (6.30)$$

$$\Delta Y_t = \Delta Z_t \beta_1^* + \beta_2^* \Delta Y_{t-1} + \beta_3^* \Delta Z_{t-1} + u_t$$

One drawback of employing Equation (6.29) is the inability to calculate the overall multiplier directly from Equation (6.28). Brandt and Williams (2007) show that using the Error Correction Model (ECM), Equation (6.30) can be transformed into a stationary form as follows:

$$(1 - L)Y_t = \beta_1 Z_t + (\beta_2 - 1)Y_{t-1} + \beta_3 Z_{t-1} + v_t,$$

$$(1 - L)Y_t = \beta_1 (1 - L)Z_t + (\beta_2 - 1)Y_{t-1} + (\beta_1 + \beta_3)Z_{t-1} + d_t,$$

$$\Delta Y_t = \beta_1 \Delta Z_t + (\beta_2 - 1) \left[ Y_{t-1} + \left( \frac{\beta_1 + \beta_3}{\beta_2 - 1} \right) Z_{t-1} \right] + g_t, \quad (6.31)$$

$$\Delta Y_t = \beta_1 \Delta Z_t + (\beta_2 - 1)u_{t-1} + g_t$$

The representation of the ECM in Equation (6.31), facilitates the estimation of the long-term multipliers,  $\beta_1$  and  $\beta_3$ , as well as the short-term dynamics through the coefficient  $\beta_2$ . Additionally, the residual  $g_t$  is stationary, which mitigates the issue of spurious regression. In incorporating a second equation for  $Z_t$  from the prior example,

the ECM is integrated into a VAR framework. To illustrate this, we consider  $y_t$  as a 1 x 2 vector comprising  $Y_t$  and  $Z_t$ , as outlined below:

$$y_t = \sum_{l=1}^p y_{t-l} A_l + u_t \quad (6.32)$$

As in Equation (6.29), we subtract  $y_{t-1}$  from both sides of Equation (6.32) as follows:

$$y_t - y_{t-1} = -y_{t-1} + \sum_{l=1}^p y_{t-l} A_l + u_t, \quad (6.33)$$

$$\Delta y_t = y_{t-1} \Pi + \Delta y_{t-1} \Gamma_1 + \dots + \Delta y_{t-p+1} \Gamma_{p-1} + u_t.$$

where;

$$\Pi = -(I_m - A_1 - \dots - A_p),$$

$$\Gamma_i = -(A_{i+1} + \dots + A_p), \quad i = 1, \dots, p-1,$$

$\Delta y_t$  represents the first difference of  $y_t$ ,  $\Pi$  represents the long-run equilibrium, and  $\Gamma_i$  the short-run dynamics. Upon identifying a cointegrated relationship within the model, this relationship is captured in the cointegration term  $\Pi$ . The coefficients representing long-term and short-term effects in equation (6.33) are derived from the coefficients of the Vector Autoregression (VAR) model as follows:

$$A_1 = \Gamma_1 + \Pi + I_m,$$

$$A_i = \Gamma_i - \Gamma_{i-1}, \quad i = 2, \dots, p-1,$$

$$A_p = -\Gamma_{p-1}.$$

A set of reduced rank regression models or canonical correlation analysis estimates the model. This method evaluates the number of cointegrating relationships between the time series by estimating the rank (i.e. number of linearity independent relationships) in the matrix  $\Pi = \lambda\beta$ . The short-run relationship is obtained by estimating the coefficients  $\Gamma_i$  (Johansen 1995). Brandt and Williams (2007) argue that the VECM models are suitable for multiple time series that follow the same trend over time.

If cointegration is confirmed, Equation (6.33) can be represented as an ECM. For example, we present a bivariate VECM ( $p_{1,t}$  and  $p_{2,t}$ ) with 2 optimal lags as follows:

$$\Delta p_{1,t} = \lambda_1(p_{1,t-1} - \phi_0 - \phi_1 p_{2,t-1}) + \theta_1 \Delta p_{1,t-1} + \theta_2 \Delta p_{2,t-1} + \varepsilon_{t_1} \quad (6.34)$$

$$\Delta p_{2,t} = \lambda_2(p_{1,t-1} - \phi_0 - \phi_1 p_{2,t-1}) + \theta_3 \Delta p_{1,t-1} + \theta_4 \Delta p_{2,t-1} + \varepsilon_{t_2}$$

The error-correction term, represented by  $p_{1,t-1} - \phi_0 - \phi_1 p_{2,t-1}$ , along with the coefficients  $\lambda_i$ , which denote the speed of adjustment, are key components in understanding how two variables adjust over time. The coefficients  $\theta_i$  describe the short-term interactions between these variables. Stigler and Sherwin (1985) highlighted that the closer  $\lambda_i$  is to one, the quicker the adjustment to price changes. Engle and Granger (1987) noted that although the error correction term approaches zero in the long run, indicating a return to equilibrium, any deviation from the long-term trend by either variable will trigger an adjustment back to equilibrium at the rate of  $\lambda_i$ . For the adjustment towards long-term equilibrium to be corrective following an external shock,  $\lambda_i$  must be negative, ensuring that the time series reverts to its long-term trend.

In summary, this section introduced two principal methodologies for examining cointegration among time series. The first approach, developed by Engle and Granger (1987), employs a two-step residual-based procedure to evaluate the null hypothesis that there is no cointegration. The second method, proposed by Johansen (1995), utilizes a system of equations through reduced rank regression to investigate cointegration presence. A common limitation of both approaches is their applicability solely to variables integrated of order one, meaning they are non-stationary at their levels. This requirement necessitates preliminary tests for the stationarity of the time series, which can introduce additional uncertainty and potential bias into the findings (Pesaran, Shin and Smith 2001). Furthermore, as highlighted in Chapter 4, overlooking structural breaks in standard unit root tests can distort the outcomes, leading to biased conclusions.

## 6.5 ARDL Cointegration Model

Chapter 4 revealed that within the dataset spanning from January 2001 to February 2020, there exists a combination of both stationary (I(0)) and non-stationary (I(1)) variables. For this reason, the cointegration methodology proposed by Pesaran, Shin and Smith (2001) will be used. This method is capable of testing relationships between

time series without the need for the regressors to be exclusively I(0), exclusively I(1), or mutually cointegrated. Specifically, this method uses a ARDL (Autoregressive Distributed Lag) approach for investigating cointegration. It also offers several additional benefits, which will be elaborated later. Before exploring the specifics of the cointegration technique, an overview of the fundamental principles of an ARDL model is provided.

In essence, an ARDL model incorporates both current and lagged values of independent variables (defining it as a distributed-lag model) as well as lagged values of the dependent variable (making it autoregressive) (Gujarati 2003). In a simplified bivariate model where  $p_1$  is the dependent variable and  $p_2$  is the independent variable, an ARDL( $p, q$ ) framework includes  $p$  lags of  $p_1$  and  $q$  lags of  $p_2$  in its explanatory variables. The representation of the model is as follows:

$$p_{1,t} = \beta_0 + \beta_1 p_{1,t-1} + \beta_2 p_{1,t-2} + \cdots + \beta_p p_{1,t-p} + \delta_1 p_{2,t-1} + \delta_2 p_{2,t-2} + \cdots + \delta_q p_{2,t-q} + \varepsilon_t \quad (6.35)$$

In a compact form, the ARDL( $p, q$ ) presented in (6.35) can be written as follows:

$$p_{1,t} = \beta_0 + \sum_{i=1}^p \beta_i p_{1,t-i} + \sum_{j=0}^q \delta_j p_{2,t-j} + \varepsilon_t \quad (6.36)$$

The ARDL model uniquely combines autoregressive elements, where its past values influence the dependent variable  $p_1$ , and distributed lag aspects, where  $p_1$  is affected by past (and potentially current) values of an independent variable  $p_2$ . A key strength of the ARDL approach is its capability to handle both I(0) and I(1) series and to accommodate varying lag lengths for different variables, showcasing the model's flexibility in capturing dynamic relationships. To ensure the effectiveness of the ARDL model, especially when applying bounds cointegration tests, it is vital to determine the optimal lag lengths for each variable, which is achieved through the selection of Information Criterion such as the Akaike Information Criterion (AIC), Schwarz Bayesian Criterion (SBC), or Hannan-Quinn Criterion (HQC), where the preferred model minimizes these statistical estimates, thereby optimizing the lag structure for the analysis.

The most significant contribution of Pesaran, Shin and Smith (2001) is introducing the ARDL framework to assess cointegration via the so-called bounds test. If cointegration is found, an error correction term is added to the ARDL model. To illustrate this model, a revised version of Equation (6.35) is proposed as follows:

$$\Delta p_{1,t} = \beta_0 + \sum_{i=1}^p \lambda_i \Delta p_{1,t-i} + \sum_{j=0}^q \delta_j \Delta p_{2,t-j} + \varphi_1 p_{1,t-1} + \varphi_2 p_{2,t-1} + v_t \quad (6.37)$$

where the coefficients  $\varphi_1$  and  $\varphi_2$  represent the long-run relationship, and the coefficients  $\lambda_i$  and  $\delta_i$  corresponds to the short-run dynamics of the model.

The bounds cointegration test consists of an F-test implemented on the joint null hypothesis that the coefficients of the lagged variables  $p_{1,t-1}$  and  $p_{2,t-1}$  in Equation (5.37) are all equal to zero. The,, the null hypothesis of no long-run relationship (cointegration) can be presented as follows:

- $H_0: \varphi_1 = \varphi_2 = 0$  (Null hypothesis, i.e. no long-run relationship or cointegration exists)
- $H_1: \varphi_1 \neq \varphi_2 \neq 0$  (Alternative, i.e. the long-run relationship or cointegration exists)

The F-statistic test used in the ARDL framework is unaffected by whether the model variables are I(0) or I(1). Pesaran, Shin and Smith (2001) provide critical values for the F-statistics, which vary based on the model's number of variables and the inclusion of an intercept and/or trend. These critical values are categorized into two groups: one for models with all I(0) variables, indicating no cointegration, and another for models with all I(1) variables, suggesting the presence of a long-run, cointegrated relationship. A specific range of critical values exists for each model scenario, accommodating all potential classifications of variables as either I(0) or I(1) (Nkoro and Uko 2016).

Upon conducting the bounds tests, if the F-statistics for the variables in Equation (6.37) is greater or lower than the specified critical bounds for I(0) and I(1), the results are conclusive regardless of the variables' classification as I(0), I(1), or mutually cointegrated. An F-statistic above the upper critical bound indicates cointegration, leading to rejecting the null hypothesis ( $H_0$ ). In contrast, an F-statistic below the lower

critical bound suggests no long-run relationship, thus not rejecting  $H_0$ . However, if the F-statistic lies within the bounds, the test's outcome is indeterminate and dependent on the actual integration order of the variables. It is important to note that the presence of any I(2) variables invalidates the F-statistics, as the test assumes variables to be either I(0), I(1), or cointegrated. Although pre-testing for unit roots is not essential, verifying that variables are not I(2) through unit root testing of their first differences before applying the bounds F-test is advisable (Nkoro and Uko 2016).

Pesaran, Shin and Smith (2001) argue that an advantage of the ARDL approach for testing cointegration is that it may represent a single long-run relationship, distinguishing the ARDL method between the dependent and explanatory variables. At first, the ARDL approach considers that only a single reduced-form equation exists between the variables. Another advantage is that the Error Correction Model (ECM) can be modelled from the ARDL through a linear transformation, integrating short-run with long-run dynamics without losing long-run information. As such, the ARDL model may be viewed as a form of unrestricted ECM, as all long-run relationship variables are specified and not restricted.

The ARDL model, which includes an Error Correction Term (ECT) as shown in Equation (6.37), assumes a symmetric effect of explanatory variables on the dependent variable, meaning that both positive and negative variations in the explanatory variables are presumed to have equal and direct impacts on the dependent variable. This assumption is a limitation of the ARDL error correction model outlined by Pesaran, Shin and Smith (2001), given that economic variables often exhibit asymmetric relationships where positive and negative shifts in explanatory variables do not equally affect the dependent variable, as evidenced in studies by Neftçi (1984), Falk (1986), Kisswani, Zaitouni, and Moufakkir (2020), and Kisswani (2021). To enhance the analysis of cointegration between natural gas and oil prices in this thesis, an advanced method, the non-linear ARDL model (NARDL) proposed by Shin, Yu and Greenwood-Nimmo (2014), is considered, addressing the limitations of the traditional ARDL model by accommodating asymmetry. The subsequent section will explore this innovative approach.

## 6.6 A Non-linear ARDL (NARDL) Approach to Cointegration

Shin, Yu and Greenwood-Nimmo (2014) developed a different ARDL bounds cointegration test to assess the asymmetric influence of the explanatory variables in the model by proposing a non-linear ARDL model (NARDL). The asymmetry of explanatory variables is introduced by partial sums, which decompose the variables into positive and negative shocks. Using the explanatory variable  $p_{2,t}$  from Equation (6.35), the positive shock ( $p_{2,t}^+$ ) and negative shock ( $p_{2,t}^-$ ) can be represented as follows:

$$p_{2,t}^+ = \sum_{j=1}^t \Delta p_{2,i}^+ = \sum_{j=1}^t \max(\Delta p_{2,j}, 0) \quad (6.38)$$

$$p_{2,t}^- = \sum_{k=1}^t \Delta p_{2,i}^- = \sum_{k=1}^t \min(\Delta p_{2,k}, 0) \quad (6.39)$$

Using the example of the bivariate model proposed by Equation (6.37), the new ECM model as proposed by Shin, Yu and Greenwood-Nimmo (2014) is as follows:

$$\begin{aligned} \Delta p_{1,t} = & \beta_0 + \sum_{i=1}^p \lambda_i \Delta p_{1,t-i} + \sum_{j=0}^q \delta_{1j} \Delta p_{2,t-j}^+ + \sum_{k=0}^r \delta_{2k} \Delta p_{2,t-k}^- \\ & + \varphi_1 p_{1,t-1} + \varphi_2 p_{2,t-1}^+ + \varphi_3 p_{2,t-1}^- + q_t \end{aligned} \quad (6.40)$$

The ECM model outlined in equation (6.40) facilitates an asymmetric evaluation of the causal impact on the dependent variable  $p_1$  over long-term and short-term periods. Shin, Yu and Greenwood-Nimmo (2014) state that the methodological framework for the bounds cointegration test by Pesaran, Shin and Smith (2001), using the F-statistic, also applies to this equation. In equation (6.40), the coefficients  $\sum_{i=1}^p \lambda_i$ ,  $\sum_{i=0}^q \delta_{1i}$  and  $\sum_{i=0}^r \delta_{2i}$ , related to the changes in explanatory variables, delineate the asymmetric short-term effects on the variable's ( $p_1$ ) volatility. Meanwhile,  $\varphi_1$ ,  $\varphi_2$ , and  $\varphi_3$  capture the asymmetric long-term impacts on the dependent variable. The existence of long-run causal asymmetry between  $p_2$  and  $p_1$  is verified if  $\varphi_2$  significantly diverges from  $\varphi_3$ , and short-run asymmetry is established if  $\sum_{i=0}^q \delta_{1i}$  significantly differ from

$\sum_{i=0}^r \delta_{2i}$ . These differences can be statistically verified through the Wald-test for joint significance.

Therefore, the F-test is applied for the null hypothesis:  $\varphi_1 = \varphi_2 = \varphi_3 = 0$ , against the alternative  $\varphi_1 \neq \varphi_2 \neq \varphi_3 \neq 0$ . As presented previously, if the null is rejected, cointegration is confirmed for the asymmetric model.

## **6.7 Global Economic Conditions Indicator**

Understanding the key drivers of world energy markets is crucial for policymakers and academic researchers, especially given the recent disruptions caused by COVID-19. Academic studies, such as those by Baumeister and Kilian (2012) and Alquist, Kilian, and Vigfusson (2013), have contributed significantly by developing models that improve energy price dynamics modelling by focusing on economic fundamentals. Kilian's (2009) dry-cargo shipping rate measure was a notable early indicator, but its reliability has diminished due to recent market volatility. Researchers now explore alternative indicators, such as world industrial production (Baumeister and Hamilton, 2019; Hamilton, 2019), broader commodity prices (Alquist, Bhattarai, and Coibion, 2020), and global steel production (Ravazzolo and Vespignani, 2020), to better predict energy demand and prices.

Baumeister, Korobilis, and Lee (2022) conducted a remarkable study introducing new variables to improve energy price forecasting, including geopolitical risk, transportation developments, oil price uncertainty, and weather indicators. They discovered that combining these variables with world industrial production enhances the accuracy of global oil price and demand forecasts. Their analysis also includes measures indicating potential upward or downward oil price pressures and predicts the likelihood of oil prices staying within recent ranges over two years. The study's applicability is demonstrated through a risk assessment for shale oil producers during the volatile period of early 2020.



### 6.7.1 The Global Economic Conditions Indicator: A Multi-Dimensional Approach

Baumeister, Korobilis, and Lee (2022) identified 16 indicators to improve forecasting energy prices and demand, guided by four principles: diversity in data categories, economic relevance, broad coverage, and manageability. These categories include:

- **Real Economic Activity:** This involves the World Industrial Production (WIP) index, the Conference Board Leading Economic Index, and the OECD Consumer Confidence Index to gauge global economic conditions and their impact on energy demand.
- **Commodity Prices:** Copper prices, deflated by the U.S. consumer price index, are used as they are key indicators of manufacturing, construction, and future global growth.
- **Financial Indicators:** These include the broad real trade-weighted U.S. dollar index, MSCI world index stock returns, and the excess return on the Fama-French transportation sector portfolio, reflecting trade, financial flows, and energy consumption.
- **Transportation:** Indicators such as vehicle registrations and U.S. total vehicle miles travelled are used to predict future fuel demand.
- **Uncertainty Measures:** The geopolitical risk index and long-run oil price uncertainty, measured by the realized volatility of WTI futures contracts, reflect geopolitical events' impact on energy markets.
- **Expectations Measures:** The Michigan Survey's index of consumer expectations and an oil price expectations measure, derived from WTI futures prices, indicate future economic and energy market trends.
- **Weather Indicators:** The Oceanic Niño Index (ONI) and the Residential Energy Demand Temperature Index (REDTI) are used to account for the impact of weather phenomena like El Niño on energy use and demand.
- **Energy-related Indicators:** The measure of energy production and electricity distribution in the EU28 reflects the overall intensity of economic activity and direct energy demand.

The authors created a global economic conditions indicator (GECON) using the first principal component from the 16 variables. This indicator is normalized to a mean of zero and a standard deviation of one. A 3-month moving average is applied to emphasize persistent economic trends. This indicator effectively tracks major economic events, showing severe downturns during the 2008-09 financial crisis and early 2020 due to the coronavirus pandemic, and periods of strong growth in the late 1980s and mid-2000s. It also highlights an improvement in 2013 and sluggish growth from 2015 to 2017.

Baumeister, Korobilis, and Lee (2022) concluded that the most accurate model employs the newly developed global economic conditions indicator in forecasting the real Brent price and fuel consumption together. They demonstrate how this model's real-time forecasts for price and consumption can generate metrics that offer policymakers and markets a quantitative assessment of anticipated oil price pressures and future energy demand trends.

### **6.7.2 Incorporating the GECON Indicator into Econometric Models of this Thesis**

Baumeister, Korobilis, and Lee (2022) demonstrated that the GECON indicator significantly enhances econometric models for forecasting oil prices and consumption. Given that natural gas is a major byproduct of oil, and its prices are still somewhat correlated with oil prices, we will incorporate the GECON indicator into the econometric models presented in Chapters 7, 8, and 9. These models will serve as a robustness check. Specifically, we will compare models that include only natural gas and crude oil prices and their causality and cointegration outcomes with similar models that also incorporate the GECON indicator as an exogenous variable. The objective is to determine whether the results are consistent across both sets of models, thereby verifying their reliability.

## **6.8 Concluding Remarks**

This chapter provides a detailed overview of the methodologies employed in subsequent chapters to explore the links between global natural gas prices and the extent of interdependence between natural gas and oil prices. It sets the stage for the

econometric models to be detailed in Chapters 7, 8 and 9 by initially explaining how economic models articulate the interactions among multiple time series. Particular emphasis is placed on two critical methods: the VAR and the ARDL approaches, with or without cointegration.

Given the dataset (2001M01 to 2020M02) contains a mix of  $I(0)$  and  $I(1)$  time series, the bounds cointegration test by Pesaran, Shin and Smith (2001) is identified as the most suitable econometric approach to examine the relationship between natural gas prices. This test offers several advantages over traditional cointegration methods like those by Engle and Granger (1987) and Johansen (1995). Key advantages include its capability to analyse short-term and long-term causality between time series without requiring the regressors to be exclusively  $I(0)$  or  $I(1)$ , its utilization of a simplified single-equation form, its effectiveness in addressing endogeneity issues, and its ability to determine the optimal lag length for variables. Importantly, the ARDL framework allows for simultaneous evaluation of the short-run and long-run impacts of explanatory variables on the dependent variables, aligning closely with the objectives of this thesis.

Additionally, the bounds cointegration test has been enhanced by Shin, Yu and Greenwood-Nimmo (2014) through the introduction of the NARDL model, which allows for the examination of asymmetric causal relationships between time series. This innovative approach distinguishes between the effects of positive and negative shocks, analysing whether the impacts of such changes differ in the short and long term across variables. It is important as it may detect cointegrating relationships and causal links that may not be detectable in the linear (misspecified) ARDL models.

In the next chapter, the traditional ARDL method for the bounds cointegration test will be employed in bivariate models across two datasets. Chapter 8 will further explore the NARDL method in bivariate models to examine and substantiate the standard ARDL bounds cointegration test findings, specifically focusing on asymmetries in the causal relationships between natural gas and crude oil prices. Consistent with the approach outlined in Chapter 4, a structural break analysis for each dependent variable will be incorporated to ensure the accuracy and reliability of the results. The analysis of bivariate models will be complemented with a more general VAR model analysis that will help discard spurious causality in the former models.

# CHAPTER 7

## BIVARIATE LINEAR ARDL MODELS

### 7.1 Introduction

The last chapter outlined the theoretical foundation for the econometric methods used in this thesis, highlighting a distinction between short- and long-run analysis. It also identified the autoregressive distributed lag (ARDL) model, favoured for handling time series of mixed integration orders, as the best approach for analysing natural gas prices, noting its ability to produce unbiased estimates and to incorporate structural breaks to improve model accuracy.

This chapter analyses price relationships between different regional natural gas markets and the Brent crude oil prices, employing the ARDL method and cointegration bounds test as per Pesaran, Shin and Smith (2001) to uncover short-run and long-run causality. It focuses on linear analysis within the ARDL models, highlighting symmetric adjustments to shocks in the variables relative to their long-term equilibrium. The results are examined for their statistical significance, particularly focusing on detecting cointegration among natural gas prices for Europe, North America, and Asia, including some newly established spot price benchmarks. Additionally, this chapter uses the ARDL cointegration models to explore price discovery dynamics, identifying leading and lagging markets.

It is worth noting that the lack of growth convergence observed in Chapter 5 may or may not be consistent with an error-correction term that includes a drift, restricted constant or a deterministic trend. In this sense, the former test is more restrictive than the latter.

### 7.2 Empirical Results of the Bivariate Linear ARDL Models

This section outlines the findings from the bounds cointegration test and the calculation of short-run and long-run coefficients' estimates for the variables in each bivariate ARDL model, applied to two distinct datasets: Sample 1 and Sample 2. The rationale behind using these two samples and their unit root test results were discussed

in Chapter 4. The Sample 1 notation and break dates are detailed in Table 7.1. Sample 2 is likewise shown in shown in Table 7.2.

Table 7.1 Summary of Sample 1 Dataset (Jan 2001 – Feb 2020).

Markets	North America	Europe		Asia	Crude Oil
Variables	LnHH	LnNBP	LnRUS	LnJPN	LnOIL
Structural Break	2004M06	2005M11	2004M08	2014M12	2016M02

Table 7.2 Summary of Sample 2 Dataset (Jul 2010 – Feb 2020).

Markets	North America	Europe			Asia		Crude Oil
Variables	LnHH	LnNBP	LnRUS	LnTTF	LnJPN	LnALNG	LnOIL
Structural Break	2018M08	2018M08	2018M08	2018M03	2015M03	2018M07	2014M11

When choosing the best lag length for VAR or ARDL models, we consider three information criteria: Akaike's Information Criterion (AIC), Schwarz Bayesian Criterion (SIC), and Hannan-Quinn Criterion. The optimal lag is selected based on the lowest value within each criterion. Generally, AIC is favoured for its superior predictive accuracy, aiming to provide the best forecast using historical data and capturing complex data relationships. On the other hand, SIC seeks to find the most streamlined model. Akaike's criterion is preferred for examining the causal links in natural gas price time series, allowing for selecting the best lag with up to four lags for both dependent and explanatory variables in the ARDL models, focusing on achieving the most accurate predictive performance.

After defining the optimal lag length, we calculate each bivariate model's bounds test statistic to determine if the variables are cointegrated. If cointegration exists, we estimate the Error Correction Model (ECM), from which we explore the long-run and short-run dynamics for further analysis. In cases where cointegration is absent, we estimate a new ARDL model using the first differences of the time series to explore short-run causality. This step involves testing whether the lagged coefficients of the explanatory variables are significantly different from zero. It is important to note that in assessing short-run causality, we do not consider the contemporary coefficient

estimates of the explanatory variables in the ARDL model. For clarity, the ARDL methodology is outlined in Figure 7.1.

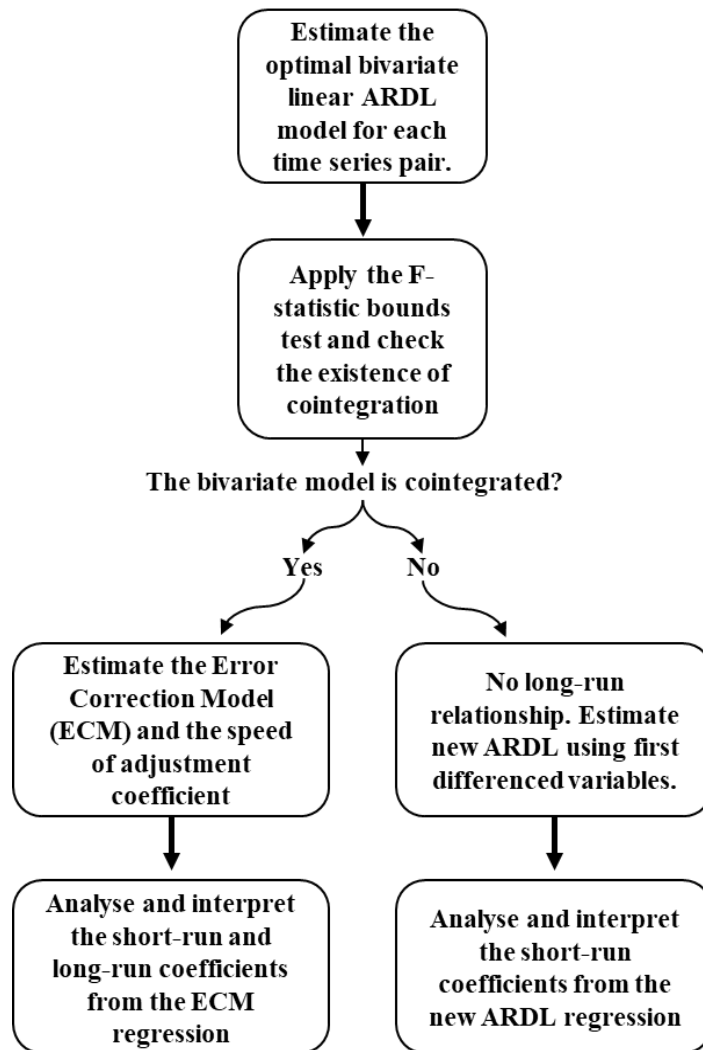


Figure 7.1 Flowchart of the ARDL Approach.

In the following subsection, we discuss the findings for Sample 1. The outcomes related to Sample 2 are detailed in Section 7.2.3.

### 7.2.1 Unit Root and Structural Break Analysis

Before exploring the ARDL method, we conducted ADF, PP, and KPSS unit root tests, with findings in Chapter 4, Tables 4.3 (Sample 1) and 4.4 (Sample 2). These tests indicate that variables from both datasets are first-order integrated, or  $I(1)$ . To refine our analysis, we utilized the ZA unit root test, which accounts for a structural break, to assess the true integration order of the variables that might misleadingly seem to be

I(1) but could actually be I(0) with a break, and similarly for variables that might appear to be I(2) instead of I(1) with a break. The ZA test results for both levels and the first differences of Sample 1 variables are in Tables 4.5 and 4.6, while for Sample 2, they are in Tables 4.7 and 4.8 of Chapter 4. These outcomes suggest a mixture of I(0) and I(1) variables in Sample 1 and confirm all Sample 2 variables as I(1), showing no evidence of I(2) variables when accounting for a structural break. When such breaks are significant, we incorporate appropriate dummy variables into the ARDL model to adjust for them.

### 7.2.2 Bivariate Linear ARDL Models Applied to Sample 1

The ARDL method for conducting the bounds cointegration test is outlined in Equation 6.37, found in Chapter 6, and re-written as follows:

$$\Delta p_{1,t} = \beta_0 + \sum_{i=1}^p \lambda_i \Delta p_{1,t-i} + \sum_{j=0}^q \delta_j \Delta p_{2,t-j} + \varphi_1 p_{1,t-1} + \varphi_2 p_{2,t-1} + v_t \quad (7.1)$$

In the model,  $p_1$  is the dependent endogenous variable, while  $p_2$  is the explanatory variable. Additionally, lagged values of  $p_1$  are included on the right-hand side of the equation as explanatory variables. The coefficients  $\lambda_i$  and  $\delta_i$  capture the short-run dynamics, which follow the first-differenced explanatory variables. The optimal lag lengths,  $p$  and  $q$ , for each explanatory variable in their first differences are determined using the Akaike Information Criterion (AIC). The long-term relationship is depicted through the coefficients  $\varphi_1$  and  $\varphi_2$ . The F-distributed bounds cointegration test, which tests the null hypothesis that the long-run coefficients jointly equal zero, is used to ascertain cointegration. Rejecting this null hypothesis indicates cointegration, i.e., a long-term relationship exists. Upon establishing cointegration, we adopt the error correction model representation of Equation 7.1 to analyse the relationship further.

$$\Delta p_{1,t} = \beta_0 + \sum_{i=1}^p \lambda_i \Delta p_{1,t-i} + \sum_{j=0}^q \delta_j \Delta p_{2,t-j} + \alpha ECT_{t-1} + v_t \quad (7.2)$$

The error correction term ( $\alpha ECT_{t-1}$ ) replaces the ARDL long-run terms ( $\varphi_1 p_{1,t-1} + \varphi_2 p_{2,t-1}$ ). The long-run influence of the explanatory variable is denoted by the

coefficient  $\varphi_2$  normalised on  $\varphi_1$ ;  $\frac{\varphi_2}{\varphi_1}$ . The coefficient  $\alpha$  represents the speed of adjustment of the bivariate model to long-run equilibrium, while the long-run coefficient  $\frac{\varphi_2}{\varphi_1}$  represents the long-run impact of the explanatory variable  $p_2$  over  $p_1$ . To ensure convergence toward long-run equilibrium, the coefficient estimate  $\alpha$  must be negative and significant; otherwise, the model will be considered unstable and explosive. If  $\alpha$  is significant, then  $p_2$  Granger-causes  $p_1$  in the long-run. The significance of  $\alpha$  is assessed by its t-statistic value. The error correction term is represented as follows:

$$ECT_{t-1} = p_{1,t-1} - (\beta + \beta_{expl}p_{2,t-1}) \quad (7.3)$$

where  $\beta_{expl}$  represents the loading factor of the explanatory variable in the bivariate model. From Equations 7.1, 7.2 and 7.3 we conclude that  $-\frac{\varphi_2}{\varphi_1}$  is equal to  $\alpha^*(-\beta_{expl})$ .

Table 7.3 reports the F-statistics for the linear bivariate ARDL cointegration bounds test (Pesaran, Shin and Smith 2001) applied to Sample 1. If cointegration is confirmed, the speed of adjustment  $\hat{\alpha}$ , and the long-run term of the explanatory variable  $-\frac{\hat{\varphi}_2}{\hat{\varphi}_1}$  (denoted by  $\hat{\varphi}$  in Table 6.3) are also reported.

Table 7.3<sup>8</sup> Linear ARDL Bivariate Models Applied to Sample 1 – Cointegration and ECM Long-Run Estimates.

Bivariate Model (Dependant – Explanatory)	ARDL Model Definition (AIC)	Bounds Test: F-statistics	I(0) 5% Critical Value	I(1) 5% Critical Value	Coint. (Y/N)	Speed of Adjustment $\hat{\alpha}$ (t-stats)	Loading Factor $\hat{\beta}_{expl}$ (t-stats)	$\hat{\alpha}$ $*(-\hat{\beta}_{expl}) = \hat{\varphi}$ (t-stats)
LnHH – LnNBP	ARDL(1,2)	2.58	4.94	5.73	N	NA	NA	NA
LnHH – LnRUS	ARDL(1,1)	4.09	4.94	5.73	N	NA	NA	NA
LnHH – LnJPN	ARDL(1,2)	4.80	4.94	5.73	N	NA	NA	NA
LnHH – LnOIL	ARDL(1,3)	2.46	4.94	5.73	N	NA	NA	NA
LnNBP – LnHH	ARDL(2,0)	5.08	4.94	5.73	N	NA	NA	NA
LnNBP – LnRUS	ARDL(2,1)	13.10***	4.94	5.73	Y	-0.189*** (-5.130)	0.801*** (5.601)	0.151*** (3.598)
LnNBP – LnJPN	ARDL(2,1)	8.48***	4.94	5.73	Y	-0.124*** (-4.127)	0.875*** (3.375)	0.609*** (2.926)
LnNBP – LnOIL	ARDL(2,2)	24.57***	4.94	5.73	Y	-0.202*** (-7.026)	1.156*** (7.952)	0.234*** (6.420)

<sup>8</sup> Table 7.3 continues on the next page.



Table 7.3 Linear ARDL Bivariate Models Applied to Sample 1 – Cointegration and ECM Long-Run Estimates, Continued.

Bivariate Model (Dependant – Explanatory)	ARDL Model Definition (AIC)	Bounds Test: F-statistics	I(0) 5% Critical Value	I(1) 5% Critical Value	Coint. (Y/N)	Speed of Adjustment $\hat{\alpha}$ (t-stats)	Loading Factor $\hat{\beta}_{expl}$ (t-stats)	$\hat{\alpha}$ $*(-\hat{\beta}_{expl}) = \hat{\phi}$ (t-stats)
LnRUS – LnHH	ARDL(4,2)	3.70	4.94	5.73	N	NA	NA	NA
LnRUS – LnNBP	ARDL(4,4)	5.74**	4.94	5.73	Y	-0.064*** (-3.393)	1.075*** (4.781)	0.068*** (3.161)
LnRUS – LnJPN	ARDL(2,3)	1.50	4.94	5.73	N	NA	NA	NA
LnRUS – LnOIL	ARDL(4,0)	11.17***	4.94	5.73	Y	-0.077*** (-4.737)	1.009*** (5.231)	0.078*** (4.273)
LnJPN – LnHH	ARDL(2,3)	1.07	4.94	5.73	N	NA	NA	NA
LnJPN – LnNBP	ARDL(2,0)	6.88**	4.94	5.73	Y	-0.030*** (-3.718)	0.863*** (4.609)	0.026*** (3.265)
LnJPN – LnRUS	ARDL(2,2)	1.80	4.94	5.73	N	NA	NA	NA
LnJPN – LnOIL	ARDL(3,2)	34.92***	4.94	5.73	Y	-0.108*** (-8.376)	0.902*** (20.448)	0.098*** (8.174)
LnOIL – LnHH	ARDL(2,1)	2.49	4.94	5.73	N	NA	NA	NA
LnOIL – LnNBP	ARDL(2,1)	2.58	4.94	5.73	N	NA	NA	NA
LnOIL – LnRUS	ARDL(2,0)	3.64	4.94	5.73	N	NA	NA	NA
LnOIL – LnJPN	ARDL(2,1)	2.05	4.94	5.73	N	NA	NA	NA

Notes: The I(0) and I(1) Bound Test critical values are reported from Pesaran, Shin, and Smith (2001), Critical values: Case III – constant and no trend.  $\hat{\alpha}$  is the speed of adjustment coefficient. \*\*\*, \*\* represent statistical significance at the 1% and 5% levels, respectively. The ARDL model definition ARDL(x,y) using the AIC approach represents that the lag length of the dependent variable is x and that the lag length of the explanatory variable is y.

Analysing Table 7.3 reveals that within Sample 1, 7 out of 20 ARDL bivariate models exhibit cointegration, indicated by their F-statistic values exceeding the 5% critical value for the I(1) Bound. Notably, the NBP gas price is the most responsive in the long run, influenced by three variables, while the RUS and JPN gas prices react to two variables. The long-term impact of the OIL price series is observed in the European gas prices (NBP and RUS) and the Asian JPN price, although OIL itself is unaffected by any other series in Sample 1. Such a pattern positions the OIL price as a leading factor, likely driving the cointegration process among natural gas prices across Europe and Asia. Conversely, the HH gas price shows no long-term adjustment to other series, indicating the US natural gas market's independence and disconnection from other gas and oil prices. This aligns with research by Kim et al. (2020), Li, Joyeux, and Ripple (2014), and Siliverstovs et al. (2005), which identified integration between European and Asian gas markets and characterised Henry Hub as a distinct market with unique pricing dynamics. Additionally, Li, Joyeux, and Ripple (2014) and Siliverstovs et al.

(2005) highlighted that the cointegration between European and Asian gas markets is supported by long-term contracts that are indexed to oil prices.

As discussed in Section 6.7, we will replicate the bounds cointegration analysis performed on Sample 1, this time incorporating the Global Economic Condition (GECON) Indicator as an exogenous variable in all bivariate models. Baumeister, Korobilis, and Lee (2022) demonstrated that the GECON indicator enhances the forecasting accuracy of econometric models for oil prices. Thus, we expect that including it in the models presented in Table 7.3 will provide a significant robustness check. Consequently, Table 7.4 presents the cointegration analysis for these bivariate ARDL models, mirroring the format of Table 7.3.

Table 7.4<sup>9</sup> Linear ARDL Models with GECON Indicator as Exogenous Variable Applied to Sample 1 – Cointegration and ECM Long-Run Estimates.

Bivariate Model (Dependant – Explanatory)	ARDL Model Definition (AIC)	Bounds Test: F-statistics	I(0) 5% Critical Value	I(1) 5% Critical Value	Coint. (Y/N)	Speed of Adjustment $\hat{\alpha}$ (t-stats)	Loading Factor $\hat{\beta}_{expt}$ (t-stats)
LnHH – LnNBP	ARDL(1,2)	2.44	4.94	5.73	N	NA	NA
LnHH – LnRUS	ARDL(1,1)	3.51	4.94	5.73	N	NA	NA
LnHH – LnJPN	ARDL(1,1)	3.87	4.94	5.73	N	NA	NA
LnHH – LnOIL	ARDL(1,3)	2.31	4.94	5.73	N	NA	NA
LnNBP – LnHH	ARDL(2,0)	4.91	4.94	5.73	N	NA	NA
LnNBP – LnRUS	ARDL(2,1)	13.79***	4.94	5.73	Y	-0.195*** (-5.265)	0.845*** (6.020)
LnNBP – LnJPN	ARDL(2,1)	8.23***	4.94	5.73	Y	-0.123*** (-4.066)	0.885*** (3.357)
LnNBP – LnOIL	ARDL(3,2)	25.97***	4.94	5.73	Y	-0.224*** (-7.223)	1.106*** (8.420)
LnRUS – LnHH	ARDL(4,2)	3.55	4.94	5.73	N	NA	NA
LnRUS – LnNBP	ARDL(4,4)	5.58	4.94	5.73	N	NA	NA
LnRUS – LnJPN	ARDL(2,3)	1.27	4.94	5.73	N	NA	NA
LnRUS – LnOIL	ARDL(4,0)	11.56***	4.94	5.73	Y	-0.093*** (-4.420)	1.004*** (6.349)
LnJPN – LnHH	ARDL(3,3)	1.12	4.94	5.73	N	NA	NA
LnJPN – LnNBP	ARDL(2,0)	6.50**	4.94	5.73	Y	-0.028*** (-3.612)	0.885*** (4.514)
LnJPN – LnRUS	ARDL(2,1)	3.59	4.94	5.73	N	NA	NA
LnJPN – LnOIL	ARDL(2,4)	33.24***	4.94	5.73	Y	-0.136*** (-8.172)	0.897*** (25.851)

<sup>9</sup> Table 7.4 continues on the next page.

Table 7.4 Linear ARDL Models with GECON Indicator as Exogenous Variable Applied to Sample 1 – Cointegration and ECM Long-Run Estimates, Continued.

Bivariate Model (Dependant – Explanatory)	ARDL Model Definition (AIC)	Bounds Test: F- statistics	I(0) 5% Critical Value	I(1) 5% Critical Value	Coint. (Y/N)	Speed of Adjustment $\hat{\alpha}$ (t-stats)	Loading Factor $\hat{\beta}_{expl}$ (t-stats)
LnOIL – LnHH	ARDL(3,1)	3.72	4.94	5.73	N	NA	
LnOIL – LnNBP	ARDL(2,4)	5.73	4.94	5.73	N	NA	
LnOIL – LnRUS	ARDL(2,0)	9.58**	4.94	5.73	Y	-0.126*** (-4.387)	0.666*** (4.785)
LnOIL – LnJPN	ARDL(2,1)	2.19	4.94	5.73	N	NA	

Notes: The I(0) and I(1) Bound Test critical values are reported from Pesaran, Shin, and Smith (2001), Critical values: Case III – constant and no trend.  $\hat{\alpha}$  is the speed of adjustment coefficient. \*\*\*, \*\* represent statistical significance at the 1% and 5% levels, respectively. The ARDL model definition ARDL(x,y) using the AIC approach represents that the lag length of the dependent variable is x and that the lag length of the explanatory variable is y.

Table 7.4 largely confirms the results of the models that included only natural gas and oil prices. Specifically, incorporating the GECON indicator as an exogenous variable resulted in 18 out of 20 bivariate model outcomes remaining consistent with those in Table 7.3. Only two models exhibited different cointegration outcomes: the LnRUS - LnNBP bivariate model lost its cointegration when the GECON indicator was included, while the LnOIL – LnRUS ARDL model became cointegrated. Overall, the results of the ARDL models, with and without the GECON indicator, show a 90% consistency.

In summary, Figure 7.2 provides a diagram illustrating the long-run causality directions between the time series in Sample 1, based on the results in Tables 7.3 and 7.4. The arrows indicate the direction of causality. Solid lines represent cointegrated pairs consistent across both ARDL model assessments. Dotted lines denote weakly cointegrated pairs identified only in the models without the GECON indicator (Table 7.3). Meanwhile, dashed lines indicate weakly cointegrated pairs found only in the models with the GECON indicator (Table 7.4).

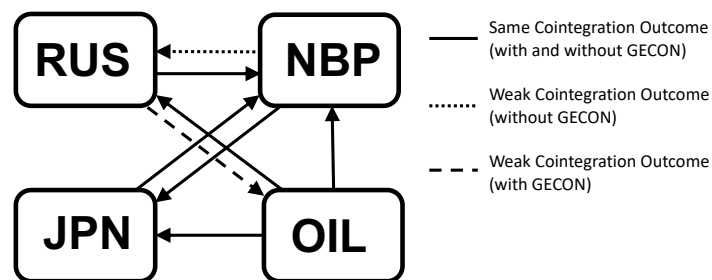


Figure 7.2 Diagram of the Linear Long-Run Causality Directions of Sample 1.

While short-run dynamics are important to assess, our primary focus is on long-run causalities (cointegration). Therefore, we will not perform a robustness check by including the GECON indicator and comparing the short-run causality results. This decision is to avoid excessively extending the results and discussion of this thesis.

The short-run causality of the cointegrated ARDL models can be determined by performing a joint F-test of the explanatory variable's short-run terms (Equation 7.2). The significance of the coefficients of the short-run terms determines the causal impact. Table 7.5 reports the results of the coefficients' estimates, when applicable, as well as the results of the Wald test for joint significance, indicating if there is a short-run causal relationship in the bivariate ECM.

Table 7.5 Linear ARDL Bivariate Models Applied to Sample 1 – ECM Short-Run Estimates.

Bivariate Model (Dependant – Explanatory)	Explanatory variable's coefficients $\hat{\delta}$ (t-stats)	Wald Test: Coefficients Joint Significance - F-statistic	Short-run Causality Relationship	Short-run Causality (Y/N)
$\Delta \text{LnNBP} - \Delta \text{LnRUS}$	No short-run coefficients in ECM	NA	RUS $\rightarrow$ NBP	N
$\Delta \text{LnNBP} - \Delta \text{LnOIL}$	$\Delta \text{LnOIL}_{t-1} = -0.322^{***}$ (-3.080)	9.486 <sup>***</sup>	OIL $\rightarrow$ NBP	Y
$\Delta \text{LnNBP} - \Delta \text{LnJPN}$	No short-run coefficients in ECM	NA	JPN $\rightarrow$ NBP	N
$\Delta \text{LnRUS} - \Delta \text{LnNBP}$	$\Delta \text{LnNBP}_{t-1} = 0.066^{**}$ (2.083) $\Delta \text{LnNBP}_{t-2} = 0.044$ (1.404) $\Delta \text{LnNBP}_{t-3} = -0.075^{**}$ (-2.369)	4.574 <sup>***</sup>	NBP $\rightarrow$ RUS	Y
$\Delta \text{LnRUS} - \Delta \text{LnOIL}$	No short-run coefficients in ECM	NA	OIL $\rightarrow$ RUS	N
$\Delta \text{LnJPN} - \Delta \text{LnNBP}$	No short-run coefficients in ECM	NA	NBP $\rightarrow$ JPN	N
$\Delta \text{LnJPN} - \Delta \text{LnOIL}$	$\Delta \text{LnOIL}_{t-1} = -0.038$ (-1.258)	1.582	OIL $\rightarrow$ JPN	N

Notes:  $\hat{\delta}$  are coefficients of the short-run terms of the bivariate ECM. <sup>\*\*\*</sup>, <sup>\*\*</sup> represent statistical significance at the 1% and 5% levels, respectively.

If cointegration is not confirmed with the F-statistic bounds test, no ECM is included in the bivariate model, but short-run dynamics are still investigated. ARDL models are estimated using time series in the first differences for these cases. Again, the AIC is used to determine the optimal lag lengths. Formally, these ARDL models are as follows.

$$\Delta p_{1,t} = \gamma_0 + \sum_{i=1}^r \eta_i \Delta p_{1,t-i} + \sum_{j=0}^s \theta_j \Delta p_{2,t-j} + u_t \quad (7.4)$$

where  $\eta_i$  are the short-run terms of the dependent variable,  $\theta_i$  are the short-run terms of the explanatory variable, and  $r$  and  $s$  are the optimal lag lengths determined by

AIC. Similar to the cointegrated ARDL model, the short-run causality is determined by performing a joint F-test of the short-run terms  $\theta_i$  of the explanatory variable. Results for the estimates of  $\theta_i$ , and the Wald test for joint significance (where applicable), are presented in Table 7.6.

Table 7.6 AIC-Augmented Linear ARDL Bivariate Models Applied to Sample 1 – Short-Run Estimates.

Bivariate Model (Dependant – Explanatory)	ARDL Model Definition (AIC)	Explanatory variable's coefficients $\hat{\theta}$ (t-stats)	Wald Test: Coefficients Joint Significance - F-statistic	Short-run Causality Relationship	Short-run Causality (Y/N)
$\Delta \ln \text{HH} - \Delta \ln \text{NBP}$	ARDL(1,1)	$\Delta \ln \text{NBP}_{t-1} = 0.143^{**}$ (2.357)	5.555 <sup>**</sup>	NBP → HH	Y
$\Delta \ln \text{HH} - \Delta \ln \text{RUS}$	ARDL(1,3)	$\Delta \ln \text{RUS}_{t-1} = -0.019$ (-0.136) $\Delta \ln \text{RUS}_{t-2} = 0.363^{***}$ (2.632) $\Delta \ln \text{RUS}_{t-3} = -0.255$ (-1.861)	2.904 <sup>**</sup>	RUS → HH	Y
$\Delta \ln \text{HH} - \Delta \ln \text{JPN}$	ARDL(1,0)	No short-run coefficients	NA	JPN → HH	N
$\Delta \ln \text{HH} - \Delta \ln \text{OIL}$	ARDL(1,2)	$\Delta \ln \text{OIL}_{t-1} = -0.161$ (-1.596) $\Delta \ln \text{OIL}_{t-2} = 0.362^{***}$ (3.780)	7.375 <sup>***</sup>	OIL → HH	Y
$\Delta \ln \text{NBP} - \Delta \ln \text{HH}$	ARDL(1,0)	No short-run coefficients	NA	HH → NBP	N
$\Delta \ln \text{RUS} - \Delta \ln \text{HH}$	ARDL(4,1)	$\Delta \ln \text{HH}_{t-1} = 0.101^{***}$ (3.048)	9.290 <sup>***</sup>	HH → RUS	Y
$\Delta \ln \text{RUS} - \Delta \ln \text{JPN}$	ARDL(4,3)	$\Delta \ln \text{JPN}_{t-1} = -0.098$ (-0.817) $\Delta \ln \text{JPN}_{t-2} = 0.313^{***}$ (2.618) $\Delta \ln \text{JPN}_{t-3} = 0.170$ (1.510)	4.793 <sup>***</sup>	JPN → RUS	Y
$\Delta \ln \text{JPN} - \Delta \ln \text{HH}$	ARDL(4,2)	$\Delta \ln \text{HH}_{t-1} = 0.046^{**}$ (2.369) $\Delta \ln \text{HH}_{t-2} = 0.056^{***}$ (2.862)	6.688 <sup>***</sup>	HH → JPN	Y
$\Delta \ln \text{JPN} - \Delta \ln \text{RUS}$	ARDL(4,3)	$\Delta \ln \text{RUS}_{t-1} = 0.097^{**}$ (2.341) $\Delta \ln \text{RUS}_{t-2} = 0.049$ (1.163) $\Delta \ln \text{RUS}_{t-3} = 0.065$ (1.563)	3.541 <sup>**</sup>	RUS → JPN	Y
$\Delta \ln \text{OIL} - \Delta \ln \text{HH}$	ARDL(1,0)	No short-run coefficients	NA	HH → OIL	N
$\Delta \ln \text{OIL} - \Delta \ln \text{NBP}$	ARDL(1,0)	No short-run coefficients	NA	NBP → OIL	N
$\Delta \ln \text{OIL} - \Delta \ln \text{RUS}$	ARDL(1,1)	$\Delta \ln \text{RUS}_{t-1} = -0.125$ (-1.400)	1.960	RUS → OIL	N
$\Delta \ln \text{OIL} - \Delta \ln \text{JPN}$	ARDL(1,0)	No short-run coefficients	NA	JPN → OIL	N

Notes:  $\hat{\theta}$  are the coefficients of the ARDL models. \*\*\*, \*\* represent statistical significance at the 1% and 5% levels, respectively. The ARDL model definition ARDL(x,y) using the AIC approach represents that the lag length of the dependent variable is x and that the lag length of the explanatory variable is y.

The analysis of short-term dynamics in the energy market, using data from Tables 7.5 and 7.6, reveals that Henry Hub (HH) and Russia (RUS) gas prices are highly responsive in the short term, influenced by European gas and oil prices. Out of 20 studied bivariate ARDL models, 9 exhibit short-term causal relationships. The Japanese (JPN) gas price also shows significant short-term sensitivity, mainly affected by RUS and HH prices, while the UK's National Balancing Point (NBP) gas price is the least responsive, only influenced by oil prices. Oil prices in turn impact NBP and

HH prices but remain unaffected by other series within the sample. The analysis suggests that the increase in LNG exports from the U.S. to Europe and the significant role of crude oil are key drivers in short-term price formations, especially for HH, despite its lack of long-term relationships with other variables.

Following the representation in Figure 7.2, Figure 7.3 displays the directions of short-term causality among the time series, as outlined in Tables 7.5 and 7.6 findings.

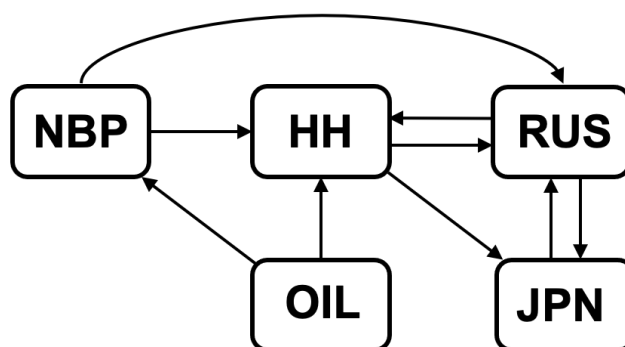


Figure 7.3 Diagram of the Linear Short-Run Causality Directions of Sample 1.

The next section presents the results of the linear ARDL bounds cointegration test applied to Sample 2.

### 7.2.3 Bivariate Linear ARDL Models Applied to Sample 2

Following the methodology proposed in the previous subsection, Table 7.7 reports the F-statistic value for the linear bivariate ARDL cointegration bounds test applied to Sample 2. If cointegration is confirmed, the speed of adjustment and the long-run term estimate are also reported.

Table 7.7<sup>10</sup> Linear ARDL Bivariate Models Applied to Sample 2 – Cointegration and ECM Long-Run Estimates.

Bivariate Model (Dependant – Explanatory)	ARDL Model Definition (AIC)	Bounds Test: F-statistics	I(0) 5% Critical Value	I(1) 5% Critical Value	Coint. (Y/N)	Speed of Adjustment $\hat{\alpha}$ (t-stats)	Loading Factor $\hat{\beta}_{expl_{t-1}}$ (t-stats)	$\hat{\alpha}^*$ $(-\beta_{expl_{t-1}})$ = $\hat{\phi}$ (t-stats)
LnHH – LnNBP	ARDL(1,0)	9.21***	4.94	5.73	Y	-0.206*** (-4.312)	0.649*** (3.811)	0.133*** (3.332)

<sup>10</sup> Table 7.7 continues on the next pages.

Table 7.7 Linear ARDL Bivariate Models Applied to Sample 2 – Cointegration and ECM Long-Run Estimates, Continued.

Bivariate Model (Dependant – Explanatory)	ARDL Model Definition (AIC)	Bounds Test: F-statistics	I(0) 5% Critical Value	I(1) 5% Critical Value	Coint. (Y/N)	Speed of Adjustment $\hat{\alpha}$ (t-stats)	Loading Factor $\hat{\beta}_{expl_{t-1}}$ (t-stats)	$\hat{\alpha}^*$ $(-\beta_{expl_{t-1}})$ $=$ $\hat{\varphi}$ (t-stats)
LnHH – LnRUS	ARDL(1,3)	5.54	4.94	5.73	N	NA	NA	NA
LnHH – LnTTF	ARDL(1,1)	5.62	4.94	5.73	N	NA	NA	NA
LnHH – LnJPN	ARDL(1,1)	4.63	4.94	5.73	N	NA	NA	NA
LnHH – LnALNG	ARDL(1,0)	7.58***	4.94	5.73	Y	-0.191*** (-3.912)	0.434*** (3.275)	0.083*** (2.833)
LnHH – LnOIL	ARDL(1,4)	4.13	4.94	5.73	N	NA	NA	NA
LnNBP – LnHH	ARDL(2,0)	3.92	4.94	5.73	N	NA	NA	NA
LnNBP – LnRUS	ARDL(4,1)	16.80***	4.94	5.73	Y	-0.653*** (-5.824)	0.747*** (21.766)	0.488*** (5.265)
LnNBP – LnTTF	ARDL(2,2)	12.49***	4.94	5.73	Y	-0.207*** (-4.071)	0.846*** (13.080)	0.337*** (4.335)
LnNBP – LnJPN	ARDL(2,0)	4.70	4.94	5.73	N	NA	NA	NA
LnNBP – LnALNG	ARDL(2,1)	8.9***	4.94	5.73	Y	-0.318*** (-4.212)	0.630*** (10.059)	0.201*** (3.659)
LnNBP – LnOIL	ARDL(2,3)	8.21***	4.94	5.73	Y	-0.202*** (-7.026)	0.699*** (5.888)	0.145*** (3.434)
LnRUS – LnHH	ARDL(2,3)	1.54	4.94	5.73	N	NA	NA	NA
LnRUS – LnNBP	ARDL(1,4)	9.33***	4.94	5.73	Y	-0.244*** (-4.340)	1.269*** (18.53)	-0.244*** (-4.272)
LnRUS – LnTTF	ARDL(2,2)	3.02	4.94	5.73	N	NA	NA	NA
LnRUS – LnJPN	ARDL(2,3)	2.75	4.94	5.73	N	NA	NA	NA
LnRUS – LnALNG	ARDL(3,4)	5.71	4.94	5.73	N	NA	NA	NA
LnRUS – LnOIL	ARDL(4,0)	7.85***	4.94	5.73	Y	-0.172*** (-3.982)	0.957*** (8.881)	0.165*** (3.805)
LnTTF – LnHH	ARDL(3,2)	1.76	4.94	5.73	N	NA	NA	NA
LnTTF – LnNBP	ARDL(1,2)	36.63***	4.94	5.73	Y	-0.697*** (-8.598)	1.063*** (18.795)	0.741*** (7.915)
LnTTF – LnRUS	ARDL(3,4)	6.26**	4.94	5.73	Y	-0.371*** (-3.556)	0.728*** (6.619)	0.270** (2.604)
LnTTF – LnJPN	ARDL(3,1)	2.14	4.94	5.73	N	NA	NA	NA
LnTTF – LnALNG	ARDL(4,0)	11.53***	4.94	5.73	Y	-0.381*** (-4.825)	0.691*** (7.784)	0.264*** (4.353)
LnTTF – LnOIL	ARDL(2,0)	6.36**	4.94	5.73	Y	-0.237*** (-3.584)	0.714*** (4.149)	0.169*** (2.805)
LnJPN – LnHH	ARDL(2,0)	8.48***	4.94	5.73	Y	-0.074*** (-4.137)	0.548 (1.816)	0.041** (2.245)
LnJPN – LnNBP	ARDL(2,1)	9.46***	4.94	5.73	Y	-0.101*** (-4.370)	0.446** (2.558)	0.045** (2.539)

Table 7.7 Linear ARDL Bivariate Models Applied to Sample 2 – Cointegration and ECM Long-Run Estimates, Continued.

Bivariate Model (Dependant – Explanatory)	ARDL Model Definition (AIC)	Bounds Test: F-statistics	I(0) 5% Critical Value	I(1) 5% Critical Value	Coint. (Y/N)	Speed of Adjustment $\hat{\alpha}$ (t-stats)	Loading Factor $\hat{\beta}_{expl_{t-1}}$ (t-stats)	$\hat{\alpha}^*$ $(-\beta_{expl_{t-1}})$ = $\hat{\varphi}$ (t-stats)
LnJPN – LnRUS	ARDL(2,4)	5.80**	4.94	5.73	Y	-0.104*** (-3.422)	0.361** (1.989)	0.038 (1.542)
LnJPN – LnTTF	ARDL(2,0)	7.90***	4.94	5.73	Y	-0.085*** (-3.992)	0.319 (1.896)	0.027** (1.994)
LnJPN – LnALNG	ARDL(2,1)	15.01***	4.94	5.73	Y	-0.124*** (-5.504)	0.471 (4.148)	0.058*** (4.025)
LnJPN – LnOIL	ARDL(2,4)	111.31***	4.94	5.73	Y	-0.477*** (-14.993)	0.829*** (30.088)	0.395*** (14.238)
LnALNG – LnHH	ARDL(2,0)	2.90	4.94	5.73	N	NA	NA	NA
LnALNG – LnNBP	ARDL(1,2)	6.93***	4.94	5.73	Y	-0.236*** (-3.741)	1.313*** (9.565)	0.309*** (3.406)
LnALNG – LnRUS	ARDL(2,4)	10.06***	4.94	5.73	Y	-0.360*** (-4.507)	1.066*** (13.466)	0.383*** (4.087)
LnALNG – LnTTF	ARDL(2,0)	5.16	4.94	5.73	N	NA	NA	NA
LnALNG – LnJPN	ARDL(2,0)	2.93	4.94	5.73	N	NA	NA	NA
LnALNG – LnOIL	ARDL(2,1)	8.93***	4.94	5.73	Y	-0.221*** (-4.246)	1.068*** (7.829)	0.236*** (3.909)
LnOIL – LnHH	ARDL(2,2)	7.88***	4.94	5.73	Y	-0.164*** (-3.988)	0.218 (0.942)	0.036 (0.908)
LnOIL – LnNBP	ARDL(2,1)	7.00**	4.94	5.73	Y	-0.163*** (-3.759)	0.244 (1.236)	0.040 (1.118)
LnOIL – LnRUS	ARDL(2,0)	8.30***	4.94	5.73	Y	-0.185*** (-4.093)	0.238 (1.329)	0.044 (1.171)
LnOIL – LnTTF	ARDL(2,0)	8.64***	4.94	5.73	Y	-0.176*** (-4.176)	0.223 (1.491)	0.039 (1.404)
LnOIL – LnJPN	ARDL(2,0)	7.52**	4.94	5.73	Y	-0.156*** (-3.896)	-0.005 (-0.012)	-0.001 (-0.012)
LnOIL – LnALNG	ARDL(2,2)	7.20**	4.94	5.73	Y	-0.169*** (-3.811)	0.188 (1.054)	0.032 (0.933)

Notes: The I(0) and I(1) Bound Test critical values are reported from Pesaran, Shin, and Smith (2001), Critical values: Case III – constant and no trend.  $\hat{\alpha}$  is the speed of adjustment coefficient. \*\*\*, \*\* represent statistical significance at the 1% and 5% levels, respectively. The ARDL model definition ARDL(x,y) using the AIC approach represents that the lag length of the dependent variable is x and that the lag length of the explanatory variable is y.

Reviewing the findings in Table 7.7 reveals that within Sample 2, 27 out of 42 ARDL bivariate models demonstrate cointegration, indicating a significant level of long-run relationship among the variables. The Japanese (JPN) gas price emerges as the most responsive in the dataset, influenced by all six examined variables. Specifically, the



adjustment speed ( $\hat{\alpha}$ ) to return to long-run equilibrium with the other five gas prices ranges from -0.074 to -0.124, signifying a 7.4% to 12.4% monthly adjustment rate following an external shock. However, when paired with oil (OIL) in a bivariate ECM, JPN's adjustment speed dramatically increases to -0.477, indicating a strong and rapid adjustment to OIL price changes within two months.

Following JPN, the UK's National Balancing Point (NBP) and the Title Transfer Facility (TTF) in Europe are the next most responsive gas prices, each influenced by four variables in the long term. NBP shows a strong long-run correlation with Russian (RUS) gas prices, while TTF's quickest adjustment is with NBP.

The Asian LNG (ALNG) price is the third most responsive, mainly influenced by European gas prices and OIL in the long run. Henry Hub (HH) and RUS gas prices display the least responsiveness, being impacted by two variables each; HH is notably influenced by NBP and ALNG, whereas RUS by NBP and OIL.

A key observation is that despite the long-term influence of OIL on NBP, TTF, RUS, and ALNG prices, their adjustment speeds with other gas prices exceed those with OIL. Remarkably, HH gas price showcases long-run integration with European and Asian gas markets, an aspect not observed in Sample 1. These results suggest an increased integration among intercontinental gas markets over time and a diminishing but still considerable long-term correlation between gas and oil prices, aligning with findings by Chiappini, Jégourel, and Raymond (2019). The exception in this trend is JPN, which remains more significantly affected by OIL due to its historical reliance on oil-indexed long-term contracts for price formation.

An interesting finding in the cointegration analysis from Table 7.7 is that the oil price is affected in the long run by all other five gas prices with similar adjustment speeds ranging from 0.155 to 0.185. Therefore, there is a bidirectional long-run causality between the natural gas prices and OIL price in Sample 2, excluding HH, which is not affected by OIL in the long run.

As outlined in the previous subsection, we conducted a second cointegration assessment using the models presented in Table 7.7, this time incorporating the GECON indicator as an exogenous variable. Again, this additional assessment will

serve as a robustness check to verify the validity of the cointegration outcomes for Sample 2, ensuring that our findings remain relatively consistent and reliable even when external economic conditions are considered. This approach helps to solidify the conclusions drawn from the original models and provides greater confidence in the long-term causal relationships identified. Table 7.8 presents the cointegration analysis for these bivariate ARDL models, following the format of Table 7.7.

Table 7.8<sup>11</sup> Linear ARDL Models with GECON Indicator as Exogenous Variable Applied to Sample 2 – Cointegration and ECM Long-Run Estimates.

Bivariate Model (Dependant – Explanatory)	ARDL Model Definition (AIC)	Bounds Test: F-statistics	I(0) 5% Critical Value	I(1) 5% Critical Value	Coint. (Y/N)	Speed of Adjustment $\hat{\alpha}$ (t-stats)	Loading Factor $\hat{\beta}_{expt_{t-1}}$ (t-stats)
LnHH – LnNBP	ARDL(1,0)	10.96***	4.94	5.73	Y	-0.240*** (-4.703)	0.557*** (3.804)
LnHH – LnRUS	ARDL(1,3)	6.78**	4.94	5.73	Y	-0.212*** (-3.699)	0.299** (2.048)
LnHH – LnTTF	ARDL(1,1)	6.68**	4.94	5.73	Y	-0.197*** (-3.670)	0.392** (2.194)
LnHH – LnJPN	ARDL(1,1)	5.34	4.94	5.73	N	NA	NA
LnHH – LnALNG	ARDL(1,0)	9.06***	4.94	5.73	Y	-0.222*** (-4.277)	0.366*** (3.177)
LnHH – LnOIL	ARDL(1,4)	4.07	4.94	5.73	N	NA	NA
LnNBP – LnHH	ARDL(2,0)	3.95	4.94	5.73	N	NA	NA
LnNBP – LnRUS	ARDL(4,1)	16.71***	4.94	5.73	Y	-0.656*** (-5.809)	0.752*** (21.372)
LnNBP – LnTTF	ARDL(2,2)	13.10***	4.94	5.73	Y	-0.417*** (-5.142)	0.872*** (13.501)
LnNBP – LnJPN	ARDL(2,0)	4.64	4.94	5.73	N	NA	NA
LnNBP – LnALNG	ARDL(2,1)	8.67***	4.94	5.73	Y	-0.207*** (-4.068)	0.720*** (5.440)
LnNBP – LnOIL	ARDL(2,3)	8.21***	4.94	5.73	Y	-0.202*** (-7.026)	0.699*** (5.888)
LnRUS – LnHH	ARDL(2,3)	1.60	4.94	5.73	N	NA	NA
LnRUS – LnNBP	ARDL(1,4)	9.49***	4.94	5.73	Y	-0.249*** (-4.378)	1.256*** (18.26)
LnRUS – LnTTF	ARDL(2,2)	3.16	4.94	5.73	N	NA	NA
LnRUS – LnJPN	ARDL(2,3)	2.68	4.94	5.73	N	NA	NA
LnRUS – LnALNG	ARDL(3,4)	5.59	4.94	5.73	N	NA	NA
LnRUS – LnOIL	ARDL(4,0)	7.77***	4.94	5.73	Y	-0.171*** (-3.961)	0.969*** (8.432)

<sup>11</sup> Table 7.8 continues on the next pages.

Table 7.8 Linear ARDL Models with GECON Indicator as Exogenous Variable Applied to Sample 2 – Cointegration and ECM Long-Run Estimates, Continued.

Bivariate Model (Dependant – Explanatory)	ARDL Model Definition (AIC)	Bounds Test: F-statistics	I(0) 5% Critical Value	I(1) 5% Critical Value	Coint. (Y/N)	Speed of Adjustment $\hat{\alpha}$ (t-stats)	Loading Factor $\hat{\beta}_{expl_{t-1}}$ (t-stats)
LnTTF – LnHH	ARDL(3,2)	2.59	4.94	5.73	N	NA	NA
LnTTF – LnNBP	ARDL(1,2)	41.76***	4.94	5.73	Y	-0.755*** (-9.182)	1.024*** (19.413)
LnTTF – LnRUS	ARDL(3,4)	7.54**	4.94	5.73	Y	-0.404*** (-3.902)	0.699*** (6.698)
LnTTF – LnJPN	ARDL(3,1)	3.07	4.94	5.73	N	NA	NA
LnTTF – LnALNG	ARDL(4,0)	11.69***	4.94	5.73	Y	-0.281*** (-4.233)	0.657*** (7.766)
LnTTF – LnOIL	ARDL(2,0)	5.08	4.94	5.73	N	NA	NA
LnJPN – LnHH	ARDL(2,0)	7.82***	4.94	5.73	Y	-0.075*** (-3.972)	0.572* (1.808)
LnJPN – LnNBP	ARDL(2,1)	8.75***	4.94	5.73	Y	-0.100*** (-4.202)	0.444** (2.512)
LnJPN – LnRUS	ARDL(2,4)	5.77**	4.94	5.73	Y	-0.102*** (-3.132)	0.332** (1.779)
LnJPN – LnTTF	ARDL(2,0)	7.20***	4.94	5.73	Y	-0.085*** (-3.992)	0.320 (1.836)
LnJPN – LnALNG	ARDL(2,1)	14.26***	4.94	5.73	Y	-0.126*** (-5.364)	0.474*** (4.206)
LnJPN – LnOIL	ARDL(2,4)	191.97***	4.94	5.73	Y	-0.505*** (-19.683)	0.831*** (32.223)
LnALNG – LnHH	ARDL(2,0)	2.89	4.94	5.73	N	NA	NA
LnALNG – LnNBP	ARDL(1,2)	6.91***	4.94	5.73	Y	-0.236*** (-3.733)	1.304*** (9.235)
LnALNG – LnRUS	ARDL(2,4)	10.06***	4.94	5.73	Y	-0.360*** (-4.507)	1.066*** (13.466)
LnALNG – LnTTF	ARDL(2,0)	5.12	4.94	5.73	N	NA	NA
LnALNG – LnJPN	ARDL(2,0)	3.11	4.94	5.73	N	NA	NA
LnALNG – LnOIL	ARDL(2,1)	9.30***	4.94	5.73	Y	-0.221*** (-4.334)	1.068*** (7.829)
LnOIL – LnHH	ARDL(2,2)	8.37***	4.94	5.73	Y	-0.167*** (-4.111)	0.046 (0.193)
LnOIL – LnNBP	ARDL(2,1)	7.59**	4.94	5.73	Y	-0.169*** (-3.916)	0.204 (1.048)
LnOIL – LnRUS	ARDL(2,0)	8.90***	4.94	5.73	Y	-0.189*** (-4.238)	0.194 (1.078)
LnOIL – LnTTF	ARDL(2,0)	8.99***	4.94	5.73	Y	-0.178*** (-4.155)	0.216 (1.331)
LnOIL – LnJPN	ARDL(2,0)	9.11**	4.94	5.73	Y	0.089*** (4.289)	3.413 (0.968)

Table 7.8 Linear ARDL Models with GECON Indicator as Exogenous Variable Applied to Sample 2 – Cointegration and ECM Long-Run Estimates, Continued.

Bivariate Model (Dependant – Explanatory)	ARDL Model Definition (AIC)	Bounds Test: F-statistics	I(0) 5% Critical Value	I(1) 5% Critical Value	Coint. (Y/N)	Speed of Adjustment $\hat{\alpha}$ (t-stats)	Loading Factor $\hat{\beta}_{expl_{t-1}}$ (t-stats)
LnOIL – LnALNG	ARDL(2,2)	7.77**	4.94	5.73	Y	-0.171*** (-3.959)	0.139 (0.761)

Notes: The I(0) and I(1) Bound Test critical values are reported from Pesaran, Shin, and Smith (2001), Critical values: Case III – constant and no trend.  $\hat{\alpha}$  is the speed of adjustment coefficient. \*\*\*, \*\* represent statistical significance at the 1% and 5% levels, respectively. The ARDL model definition ARDL(x,y) using the AIC approach represents that the lag length of the dependent variable is x and that the lag length of the explanatory variable is y.

Similar to the findings for Sample 1, Table 7.8 predominantly confirms the cointegration results of the original models presented in Table 7.7 for long-run causality relationships within Sample 2. With the GECON indicator included as an exogenous variable, 39 out of 42 bivariate model cointegration outcomes remain consistent with those in Table 7.7. Only three models exhibit different cointegration outcomes. Overall, the bivariate ARDL models applied to Sample 2, with and without the GECON indicator, demonstrate a 93% consistency in cointegration results.

Similar to the diagram presented in Figure 7.2 for Sample 1, Figure 7.4 illustrates the long-run causality directions between the price time series in Sample 2, as derived from the results in Tables 7.7 and 7.8.

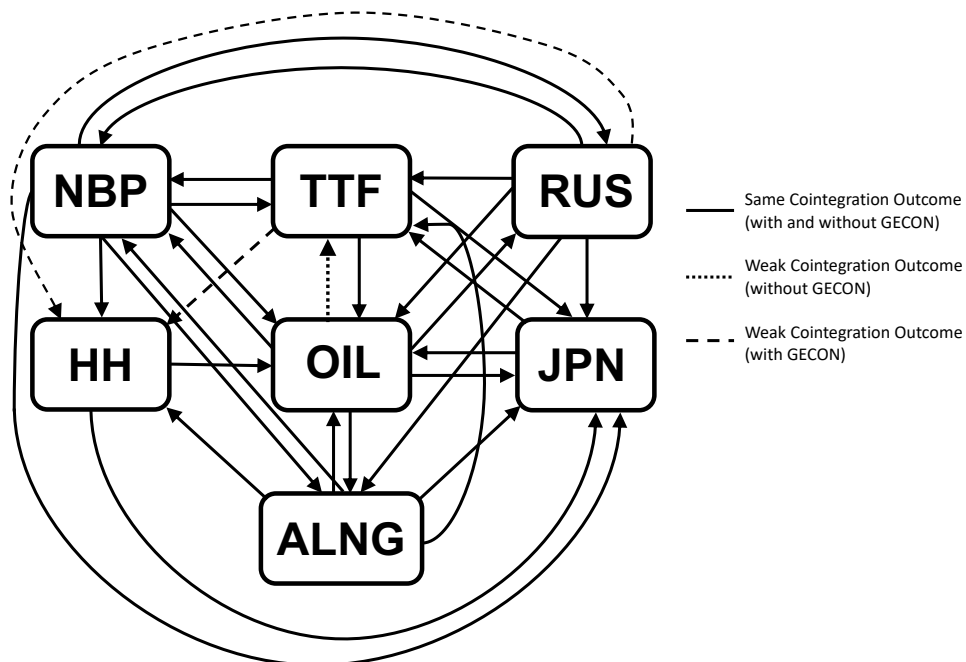


Figure 7.4 Diagram of the Linear Long-Run Causality Directions of Sample 2.

Table 7.9 reports the results of the  $\delta_i$  coefficient's estimates, when applicable, as well as the results of the Wald test for joint significance, indicating if there is a short-run causal relationship in the bivariate ECM applied to Sample 2.

Table 7.9<sup>12</sup> Linear ARDL Bivariate Models Applied to Sample 2 – ECM Short-Run Estimates.

Bivariate Model (Dependant – Explanatory)	First-differenced explanatory coefficients $\delta$ (t-stats)	Wald Test: Coefficients Joint Significance - F-statistic	Short-run Causality Relationship	Short-run Causality (Y/N)
$\Delta \text{LnHH} - \Delta \text{LnNBP}$	No short-run coefficients in ECM	NA	NBP $\rightarrow$ HH	N
$\Delta \text{LnHH} - \Delta \text{LnALNG}$	No short-run coefficients in ECM	NA	LNG $\rightarrow$ HH	N
$\Delta \text{LnNBP} - \Delta \text{LnRUS}$	No short-run coefficients in ECM	NA	RUS $\rightarrow$ NBP	N
$\Delta \text{LnNBP} - \Delta \text{LnTTF}$	$\Delta \text{LnTTF}_{t-1} = -0.114 (-1.918)$	3.679	TTF $\rightarrow$ NBP	N
$\Delta \text{LnNBP} - \Delta \text{LnALNG}$	No short-run coefficients in ECM	NA	LNG $\rightarrow$ NBP	N
$\Delta \text{LnNBP} - \Delta \text{LnOIL}$	$\Delta \text{LnOIL}_{t-1} = -0.245 (-2.080)$	4.326	OIL $\rightarrow$ NBP	N
$\Delta \text{LnRUS} - \Delta \text{LnNBP}$	$\Delta \text{LnNBP}_{t-1} = 0.249^{***} (3.776)$ $\Delta \text{LnNBP}_{t-2} = -0.027 (-0.521)$ $\Delta \text{LnNBP}_{t-3} = -0.112^{**} (-2.146)$	9.303 <sup>***</sup>	NBP $\rightarrow$ RUS	Y
$\Delta \text{LnRUS} - \Delta \text{LnOIL}$	$\Delta \text{LnOIL}_{t-1} = 0.079 (0.772)$ $\Delta \text{LnOIL}_{t-2} = -0.257^{***} (-2.848)$	4.219 <sup>**</sup>	OIL $\rightarrow$ RUS	Y
$\Delta \text{LnTTF} - \Delta \text{LnNBP}$	$\Delta \text{LnNBP}_{t-1} = -0.461^{***} (-2.319)$	5.378 <sup>***</sup>	NBP $\rightarrow$ TTF	Y
$\Delta \text{LnTTF} - \Delta \text{LnRUS}$	$\Delta \text{LnRUS}_{t-1} = 0.242 (1.128)$ $\Delta \text{LnRUS}_{t-2} = 0.303 (1.534)$ $\Delta \text{LnRUS}_{t-3} = 0.301 (1.600)$	2.755	RUS $\rightarrow$ TTF	N
$\Delta \text{LnTTF} - \Delta \text{LnALNG}$	No short-run coefficients in ECM	NA	LNG $\rightarrow$ TTF	N
$\Delta \text{LnTTF} - \Delta \text{LnOIL}$	No short-run coefficients in ECM	NA	OIL $\rightarrow$ TTF	N
$\Delta \text{LnJPN} - \Delta \text{LnHH}$	No short-run coefficients in ECM	NA	HH $\rightarrow$ JPN	N
$\Delta \text{LnJPN} - \Delta \text{LnNBP}$	No short-run coefficients in ECM	NA	NBP $\rightarrow$ JPN	N
$\Delta \text{LnJPN} - \Delta \text{LnRUS}$	$\Delta \text{LnRUS}_{t-1} = -0.007 (-0.132)$ $\Delta \text{LnRUS}_{t-2} = -0.061 (-1.174)$ $\Delta \text{LnRUS}_{t-3} = 0.153^{***} (2.993)$	3.439 <sup>**</sup>	RUS $\rightarrow$ JPN	Y
$\Delta \text{LnJPN} - \Delta \text{LnTTF}$	No short-run coefficients in ECM	NA	TTF $\rightarrow$ JPN	N
$\Delta \text{LnJPN} - \Delta \text{LnALNG}$	No short-run coefficients in ECM	NA	LNG $\rightarrow$ JPN	N
$\Delta \text{LnJPN} - \Delta \text{LnOIL}$	$\Delta \text{LnOIL}_{t-1} = -0.385^{***} (-10.082)$ $\Delta \text{LnOIL}_{t-2} = -0.379^{***} (-10.234)$ $\Delta \text{LnOIL}_{t-3} = -0.284^{***} (-7.017)$	50.185 <sup>***</sup>	OIL $\rightarrow$ JPN	Y
$\Delta \text{LnALNG} - \Delta \text{LnNBP}$	$\Delta \text{LnNBP}_{t-1} = 0.274^{**} (2.582)$	6.667 <sup>**</sup>	NBP $\rightarrow$ LNG	Y
$\Delta \text{LnALNG} - \Delta \text{LnRUS}$	$\Delta \text{LnRUS}_{t-1} = -0.147 (-0.967)$ $\Delta \text{LnRUS}_{t-2} = 0.442^{***} (3.070)$ $\Delta \text{LnRUS}_{t-3} = -0.403^{***} (-2.846)$	4.606 <sup>***</sup>	RUS $\rightarrow$ LNG	Y
$\Delta \text{LnALNG} - \Delta \text{LnOIL}$	No short-run coefficients in ECM	NA	OIL $\rightarrow$ LNG	N
$\Delta \text{LnOIL} - \Delta \text{LnHH}$	$\Delta \text{LnHH}_{t-1} = -0.110 (-1.709)$	2.921	HH $\rightarrow$ OIL	N
$\Delta \text{LnOIL} - \Delta \text{LnNBP}$	No short-run coefficients in ECM	NA	NBP $\rightarrow$ OIL	N

<sup>12</sup> Table 7.9 continues on the next page.

Table 7.9 Linear ARDL Bivariate Models Applied to Sample 2 – ECM Short-Run Estimates, Continued.

Bivariate Model (Dependant – Explanatory)	First-differenced explanatory coefficients $\delta$ (t-stats)	Wald Test: Coefficients Joint Significance - F-statistic	Short-run Causality Relationship	Short-run Causality (Y/N)
$\Delta \text{LnOIL} - \Delta \text{LnRUS}$	No short-run coefficients in ECM	NA	RUS $\rightarrow$ OIL	N
$\Delta \text{LnOIL} - \Delta \text{LnTTF}$	No short-run coefficients in ECM	NA	TTF $\rightarrow$ OIL	N
$\Delta \text{LnOIL} - \Delta \text{LnJPN}$	No short-run coefficients in ECM	NA	JPN $\rightarrow$ OIL	N
$\Delta \text{LnOIL} - \Delta \text{LnALNG}$	$\Delta \text{LnLNG}_{t-1} = -0.095 (-1.570)$	2.465	LNG $\rightarrow$ OIL	N

Notes:  $\hat{\delta}$  are coefficients of the short-run terms of the bivariate ECM. \*\*\*, \*\* represent statistical significance at the 1% and 5% levels, respectively.

As before, for the bivariate models that are not cointegrated, the short-run impacts are assessed by ARDL models using the dependent and explanatory time series in first differences, and the results are reported in Table 7.10.

Table 7.10<sup>13</sup> AIC-Augmented Linear ARDL Bivariate Models Applied to Sample 2 – Short-Run Estimates.

Bivariate Model (Dependant – Explanatory)	ARDL Model Definition (AIC)	First-differenced explanatory coefficients $\theta$ (t-stats)	Wald Test: Coefficients Joint Significance - F-statistic	Short-run Causality Relationship	Short-run Causality (Y/N)
$\Delta \text{LnHH} - \Delta \text{LnRUS}$	ARDL(1,2)	$\Delta \text{LnRUS}_{t-1} = 0.003 (0.019)$ $\Delta \text{LnRUS}_{t-2} = 0.362^{**} (2.388)$	3.173	RUS $\rightarrow$ HH	N
$\Delta \text{LnHH} - \Delta \text{LnTTF}$	ARDL(1,0)	No short-run coefficients	NA	TTF $\rightarrow$ HH	N
$\Delta \text{LnHH} - \Delta \text{LnJPN}$	ARDL(1,0)	No short-run coefficients	NA	JPN $\rightarrow$ HH	N
$\Delta \text{LnHH} - \Delta \text{LnOIL}$	ARDL(1,3)	$\Delta \text{LnLNG}_{t-1} = -0.123 (-0.890)$ $\Delta \text{LnLNG}_{t-2} = 0.370^{***} (2.679)$ $\Delta \text{LnLNG}_{t-3} = 0.227 (1.624)$	4.404 <sup>***</sup>	OIL $\rightarrow$ HH	Y
$\Delta \text{LnNBP} - \Delta \text{LnHH}$	ARDL(4,0)	No short-run coefficients	NA	HH $\rightarrow$ NBP	N
$\Delta \text{LnNBP} - \Delta \text{LnJPN}$	ARDL(4,0)	No short-run coefficients	NA	JPN $\rightarrow$ NBP	N
$\Delta \text{LnRUS} - \Delta \text{LnHH}$	ARDL(1,4)	$\Delta \text{LnHH}_{t-1} = 0.185^{***} (3.125)$ $\Delta \text{LnHH}_{t-2} = -0.082 (-1.345)$ $\Delta \text{LnHH}_{t-3} = -0.034 (-0.578)$ $\Delta \text{LnHH}_{t-4} = 0.146^{**} (2.470)$	4.648 <sup>***</sup>	HH $\rightarrow$ RUS	Y
$\Delta \text{LnRUS} - \Delta \text{LnTTF}$	ARDL(4,1)	$\Delta \text{LnTTF}_{t-1} = 0.151^{***} (3.726)$	13.883 <sup>***</sup>	TTF $\rightarrow$ RUS	Y
$\Delta \text{LnRUS} - \Delta \text{LnJPN}$	ARDL(4,2)	$\Delta \text{LnJPN}_{t-1} = -0.286 (-1.320)$ $\Delta \text{LnJPN}_{t-2} = 0.407^{**} (2.304)$	2.677	JPN $\rightarrow$ RUS	N

<sup>13</sup> Table 7.10 continues on the next page.

Table 7.10 AIC-Augmented Linear ARDL Bivariate Models Applied to Sample 2 – Short-Run Estimates, Continued.

Bivariate Model (Dependant – Explanatory)	ARDL Model Definition (AIC)	First-differenced explanatory coefficients $\theta$ (t-stats)	Wald Test: Coefficients Joint Significance - F-statistic	Short-run Causality Relationship	Short-run Causality (Y/N)
$\Delta \text{LnRUS} - \Delta \text{LnALNG}$	ARDL(2,3)	$\Delta \text{LnLNG}_{t-1} = 0.173^{***}$ (3.372) $\Delta \text{LnLNG}_{t-2} = 0.005$ (0.093) $\Delta \text{LnLNG}_{t-3} = 0.203^{***}$ (3.980)	8.989 <sup>***</sup>	LNG → RUS	Y
$\Delta \text{LnTTF} - \Delta \text{LnHH}$	ARDL(2,1)	$\Delta \text{LnHH}_{t-1} = 0.380^{***}$ (2.938)	8.632 <sup>***</sup>	HH → TTF	Y
$\Delta \text{LnTTF} - \Delta \text{LnJPN}$	ARDL(2,0)	No short-run coefficients	NA	JPN → TTF	N
$\Delta \text{LnALNG} - \Delta \text{LnHH}$	ARDL(1,0)	No short-run coefficients	NA	HH → LNG	N
$\Delta \text{LnALNG} - \Delta \text{LnTTF}$	ARDL(1,0)	No short-run coefficients	NA	TTF → LNG	N
$\Delta \text{LnALNG} - \Delta \text{LnJPN}$	ARDL(1,0)	No short-run coefficients	NA	JPN → LNG	N

Notes:  $\hat{\theta}$  are the coefficients of the ARDL models. \*\*\*, \*\* represent statistical significance at the 1% and 5% levels, respectively. The ARDL model definition ARDL(x,y) using the AIC approach represents that the lag length of the dependent variable is x and that the lag length of the explanatory variable is y.

The short-run causality relationships of Sample 2 are investigated using the results from Tables 7.9 and 7.10. The results show that 13 out of 42 ARDL bivariate models in Sample 2 have a short-run causal relationship.

Russia is the most reactive gas price in the short-run affected by shocks from all prices other than JPN. The variables TTF, JPN, and ALNG are each affected by two prices in the short-run. The TTF gas price is caused by HH and NBP, which is evidence of the short-run link between the European and the U.S. markets. These relationships can be attributed to LNG exports from the U.S. to Europe and the fact that the TTF is the commercial centre of the European gas market, resulting in a strong linkage of price signalling from other hubs (such as NBP) to the transport cost differentials to TTF. OIL and RUS cause the JPN gas market in short-run. However, the impact of OIL is much stronger, as seen from the F-statistic (Coefficients Joint Significance) shown in Table 7.9. The ALNG is affected by two European gas markets, RUS and NBP. This proves that ALNG has a strong short-run correlation with gas markets rather than the OIL market. The HH and NBP prices are only affected by the OIL market in the short-run. Finally, the OIL market is the least reactive market, as its development in the short-run is not affected by any other market in Sample 2.

Figure 7.5 illustrates the short-run causality directions between the variables in Sample 2 according to Tables 7.9 and 7.10.

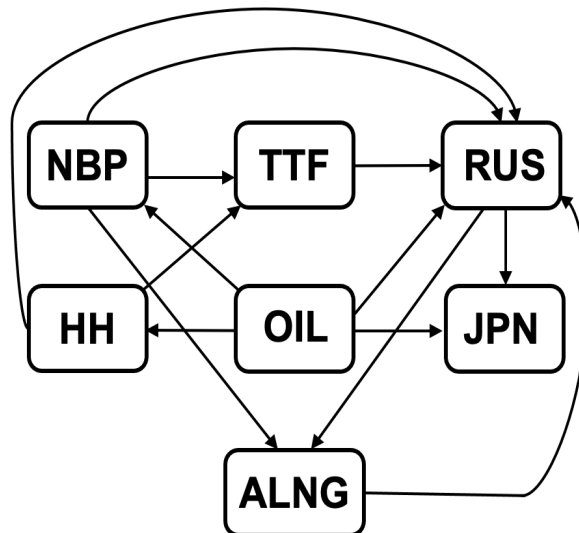


Figure 7.5 Diagram of the Linear Short-Run Causality Directions of Sample 2.

The short-run results fill the gap in explaining dynamic relationships between prices that are not cointegrated in some cases ( $\text{LnHH} - \text{LnOIL}$ ,  $\text{LnRUS} - \text{LnHH}$ , and  $\text{LnRUS} - \text{LnALNG}$ ). It is worth highlighting that there are price pairs that show no relationship at all, either in the long run or the short run. These bivariate models are in Table 7.10 (non-cointegrated models), with no short-run causality.

#### 7.2.4 Price Discovery Assessment - Leading and Lagging Prices

Building on the cointegration analysis within bivariate ARDL models presented in subsections 7.2.2 and 7.2.3, we now focus on determining between leading and lagging markets by employing a concept known in the literature as price discovery.

The bounds cointegration test in Pesaran, Shin and Smith (2001), when applied to bivariate ARDL models, can yield one of three outcomes. Firstly, it might find no evidence of cointegration, indicating that the variables operate independently in the long run without a shared equilibrium dynamic. Secondly, cointegration may be detected only when one of the variables acts as the dependent variable in the model. This suggests that this variable adjusts in response to long-term shocks from the other variable, thereby identifying it as the lagging market in this relationship. Lastly, cointegration could be established with both variables acting as dependents, indicating mutual long-term causality. In such scenarios, the leading and lagging markets within these cointegrated pairs can be distinguished using the price discovery method, which



evaluates the relative speed of adjustment coefficients to determine the share of total adjustment.

Price discovery in this context is determined by analysing the relative speed of adjustment ratios, drawing upon methodologies from Schwarz and Szakmary (1994), Foster (1996), and Theissen (2002). Within a bivariate model that incorporates prices  $p_1$  and  $p_2$ , alongside  $\alpha_1$  and  $\alpha_2$  representing the respective speed of adjustment coefficients for prices  $p_1$  and  $p_2$  towards long-term equilibrium, the comparative ratio of these adjustment speeds is articulated as follows:

$$\theta_1 = \frac{|\alpha_2|}{|\alpha_1| + |\alpha_2|}, \quad \theta_2 = \frac{|\alpha_1|}{|\alpha_1| + |\alpha_2|}, \quad \theta_1 + \theta_2 = 1 \quad (7.5)$$

In this formula,  $\theta_i$  measures the responsiveness of price  $i$  to an unexpected shock in the system. Specifically, a lower (higher)  $\theta_i$  corresponds to a higher (lower)  $\alpha_i$ , indicating that price  $i$  adjusts rapidly (gradually) back to the long-term equilibrium. Consequently, price  $i$  is classified as the lagging (leading) market within the context of the bivariate cointegration analysis.

Tables 7.11 and 7.12 offer a summary of the speed of adjustment coefficients of each cointegrated bivariate model in Samples 1 and 2, respectively. Importantly, it identifies which bivariate models have a bidirectional causality in the long-run, enabling the calculation of the relative ratio of the speed of adjustment coefficients. The tables also report the number of months that it will take for the dependent variable to return to long-run equilibrium.

Table 7.11 Summary of the Speed of Adjustment Coefficients of Cointegrated Bivariate Models in Sample 1.

Bivariate Model		Speed of Adjustment $\hat{\alpha}$	Return to Long-run Equilibrium (months)	Long run Causality Relationship
Dependant Variable	Explanatory Variable			
LnNBP	LnRUS	-0.189	5.29	Bidirectional
LnRUS	LnNBP	-0.064	15.63	
LnNBP	LnJPN	-0.124	8.06	Bidirectional
LnJPN	LnNBP	-0.030	33.33	
LnNBP	LnOIL	-0.202	4.95	Unidirectional
LnRUS	LnOIL	-0.077	12.99	Unidirectional
LnJPN	LnOIL	-0.108	9.26	Unidirectional

Table 7.12 Summary of the Speed of Adjustment Coefficients of Cointegrated Bivariate Models in Sample 2.

Bivariate Model		Speed of Adjustment $\hat{\alpha}$	Return to Long-run Equilibrium (months)	Long run Causality Relationship
Dependant Variable	Explanatory Variable			
LnHH	LnNBP	-0.206	4.85	Unidirectional
LnHH	LnALNG	-0.191	5.23	Unidirectional
LnNBP	LnRUS	-0.653	1.53	Bidirectional
LnRUS	LnNBP	-0.244	4.09	
LnNBP	LnTTF	-0.207	4.83	Bidirectional
LnTTF	LnNBP	-0.697	1.43	
LnNBP	LnALNG	-0.318	3.14	Bidirectional
LnALNG	LnNBP	-0.236	4.24	
LnNBP	LnOIL	-0.202	4.95	Bidirectional
LnOIL	LnNBP	-0.163	6.13	
LnRUS	LnOIL	-0.172	5.81	Bidirectional
LnOIL	LnRUS	-0.185	5.40	
LnTTF	LnALNG	-0.381	2.62	Unidirectional
LnTTF	LnRUS	-0.371	2.70	Unidirectional
LnTTF	LnOIL	-0.237	2.62	Bidirectional
LnOIL	LnTTF	-0.176	5.68	
LnJPN	LnALNG	-0.124	8.06	Unidirectional
LnJPN	LnRUS	-0.104	9.61	Unidirectional
LnJPN	LnNBP	-0.101	9.90	Unidirectional
LnJPN	LnTTF	-0.085	11.76	Unidirectional
LnJPN	LnHH	-0.074	13.51	Unidirectional
LnJPN	LnOIL	-0.477	2.09	Bidirectional
LnOIL	LnJPN	-0.156	6.41	
LnALNG	LnRUS	-0.360	2.78	Unidirectional
LnALNG	LnOIL	-0.221	4.52	Bidirectional
LnOIL	LnALNG	-0.169	5.92	
LnOIL	LnHH	-0.164	6.10	Unidirectional

Having identified the bidirectional causal long-run relationships, the relative speed of adjustment coefficients were calculated to determine the leading market in these bivariate models for price discovery. The summarised findings in Table 7.11 for Sample 1 reveal that the influence of OIL prices on gas prices, except HH, is predominantly one-way. Despite the bidirectional cointegration among European and Asian gas markets, OIL prices emerge as the leading market due to their significant role in shaping the long-term dynamics between NBP, RUS, and JPN.

The summary findings from Table 7.12 reveal that the RUS market leads the price discovery relative to the TTF and ALNG gas prices due to a unidirectional long-term relationship. In the Asian market, ALNG is the leading market that influences the JPN gas price. Additionally, ALNG and the NBP gas prices play significant roles in shaping the price discovery dynamics of the North American HH gas price.

When examining bidirectional long-term relationships, the analysis reveals that purely natural gas price models exhibit a higher speed of adjustment than those models integrating gas and oil prices, with the JPN gas price being an exception. In this case, it is unclear which are the leading (price discovery) and lagging gas prices. Therefore, the price discovery is assessed only for those bivariate models in Sample 2 (Table 7.12) that present a bidirectional long-run relationship. Table 7.13 presents the results.

Table 7.13 Price Discovery Assessment of Cointegrating Pairs in Sample 2.

Bivariate Model		Speed of Adjustment $\hat{\alpha}$	Relative Ratio of $\alpha_i$ $\theta_i$	Price Discovery
Dependant Variable	Explanatory Variable			
LnNBP	LnRUS	-0.653	0.272	Lagging
LnRUS	LnNBP	-0.244	0.728	Leading
LnNBP	LnTTF	-0.207	0.771	Leading
LnTTF	LnNBP	-0.697	0.229	Lagging
LnNBP	LnALNG	-0.318	0.426	Lagging
LnALNG	LnNBP	-0.236	0.574	Leading
LnNBP	LnOIL	-0.202	0.447	Lagging
LnOIL	LnNBP	-0.163	0.553	Leading
LnRUS	LnOIL	-0.172	0.518	Leading
LnOIL	LnRUS	-0.185	0.482	Lagging
LnTTF	LnOIL	-0.237	0.426	Lagging
LnOIL	LnTTF	-0.176	0.574	Leading
LnJPN	LnOIL	-0.477	0.246	Lagging
LnOIL	LnJPN	-0.156	0.754	Leading
LnALNG	LnOIL	-0.221	0.433	Lagging
LnOIL	LnALNG	-0.169	0.567	Leading

Analysing the data from Table 7.13, the RUS gas price emerges as the leader in price discovery within Sample 2, showcasing the highest  $\theta_i$  in its cointegration model with the NBP and OIL prices. Similarly, OIL assumes a leading role in price discovery across all its paired interactions, except when compared with RUS. However, in the cointegrating pairs of OIL with NBP, TTF, and ALNG, the  $\theta_i$  values for OIL do not markedly exceed 0.5, indicating a relatively equal pace of adjustment towards long-

term equilibrium between the paired prices. A significant difference is observed in the pair involving JPN and OIL prices, where JPN lags OIL. An interesting finding is that the ALNG price leads in price discovery when evaluated against its long-term equilibrium with the NBP price, underscoring its influence in the global LNG market.

### 7.3 Concluding Remarks

This chapter presented the main results of the Pesaran, Shin and Smith (2001) ARDL cointegration models that apply to the two datasets of natural gas prices and the Brent crude oil price time series, complemented with short-run dynamics analysis.

We conducted a robustness check to ensure the reliability of our long-run causalities (cointegrated bivariate models) results, which are based solely on price time series. The outcomes remained consistent when we included the GECON indicator, an exogenous variable that accounts for various global economic factors influencing commodity market dynamics. Therefore, our interpretation will focus on the original cointegration results in Tables 7.3 and 7.7. The number of causality relationships is summarized in Table 7.14.

Table 7.14 Summary of the Linear ARDL Bivariate Models Results.

	<b>Sample 1 2001 - 2020</b>	<b>Sample 2 2010 - 2020</b>
Number of Variables	5	7
Number of long-run relationships (Linear ARDL)	7 out of 20 (35%)	27 out of 42 (64.3%)
Number of short-run relationships (Linear ARDL)	9 out of 20 (45%)	13 out of 42 (31%)

Analysing Table 7.14, shows that the number of cointegrated markets increased significantly between the two sample periods. It increased from 7 out of 20 pairs (35%) in Sample 1 to 27 out of 42 (64.3%) in Sample 2 (Sample 2 has two additional time series). This comparison is interpreted as evidence of a higher degree of market integration between gas markets in the most recent sample period, even though the number of variables increased by two. Furthermore, in Sample 1, which starts in 2001, the number of short-run causality relationships far exceeded the long-run relationships, meaning that the causality dynamics between the markets were

predominantly unstable and temporary, and the markets' long-run dynamics were significantly affected by the OIL price. On the other hand, in Sample 2, the long-run relationships were predominant, indicating that the causality dynamics between the gas prices became more stable and persistent from 2010 to 2020.

An important consideration is that in Sample 1 (2001 to 2020), there is evidence of integration between the European and Asian markets. However, the leading price is OIL, which affects, in a unidirectional relationship, all other gas prices in the long run other than the HH, as the OIL price can facilitate the long-run relationship between NBP, RUS, and JPN. Also, the HH is not cointegrated with any other time series in Sample 1, which suggests that the North American market is rather independent and has its gas-on-gas price formation.

On the other hand, when analysing the cointegration relationships in bivariate models in Sample 2 (2010 to 2020) by adding two other gas prices (TTF and ALNG), the results present a different scenario than that of Sample 1. First, the speed of adjustment of the bivariate models containing only natural gas prices is more significant than the bivariate models with gas and oil prices. The result shows a greater cointegration between gas prices and that the long-term relationship between natural gas and oil prices has become bidirectional.

Second, there is a high degree of integration between the three European gas prices (RUS, NBP, TTF) and the Asian spot LNG price ALNG. Even though there is a bidirectional causality in most of these bivariate models, the RUS gas price is the leading market. The slow growth of Russian natural gas exports to Europe in the last decade can explain this. Kutcherov et al. (2020) show that more than 35% of European gas imports in 2018 came from Russia. Furthermore, Russia exports natural gas in the form of LNG to the Asia Pacific region, Europe, and other countries (the Middle East and Canada), considering the period of the sample (pre-Russia-Ukraine war).

Third, in Sample 2, there is evidence of cointegration between the U.S. (Henry Hub), Europe (NBP), and Asia (ALNG) gas markets. The results suggest that both NBP and ALNG affect the HH in the long run. This outcome could be justified by the recent increase in the U.S. domestic gas production through shale gas production, which changed the United States' status from an importer to a net exporter of natural gas. In

2018, the net natural gas exports from the U.S. to Europe and Asia reached 16.6 bcm (Kutcherov et al. 2020). Finally, regarding Sample 2, the natural gas market still highly dependent on OIL price is the JPN, which, despite being affected by all other gas prices, has a speed of adjustment with OIL price more than four times greater than any other gas price.

Another interesting finding when analysing the ARDL results for Sample 2 is that there is a bidirectional long-run causality between the natural gas prices and the OIL price in Sample 2, excluding the HH, which is not affected by OIL in the long run. Several factors could have contributed to the bidirectional causality between oil and gas prices. Technological advancements and regulatory changes in the energy industry over the past decade have increased the production and supply of natural gas, which may have led to changes in the relationship between oil and gas prices. In particular, the shale gas revolution in the US has significantly increased domestic gas production and reduced dependence on foreign oil imports, which may have influenced the dynamics of the oil-gas price relationship (Newell and Raimi 2014). Moreover, the impact of environmental policies and concerns on energy markets must be addressed. As governments and societies worldwide have become increasingly focused on reducing carbon emissions and transitioning to renewable energy sources, the demand for oil and gas has been affected (Dong, Sun and Hochman 2017). This could have spilled over the long-run relationship between oil and gas prices.

Contemplating these possibilities, the next chapter applies the Asymmetric ARDL bounds cointegration test proposed by Shin, Yu and Greenwood-Nimmo (2014) to Samples 1 and 2. Then, Chapter 9 will present a VAR assessment englobing all seven-time series in Sample 2 in a single model to assess the causality relationship of the variables as a group and also to investigate if there are some different insights from the bivariate ARDL models' results discussed in this chapter.

## CHAPTER 8

### BIVARIATE ASYMMETRIC ARDL MODELS

#### 8.1 Introduction

This chapter conducts a robustness check on the findings from Chapter 7, with an alternative Nonlinear Autoregressive Distributed Lag (NARDL) model. Unlike the symmetric ARDL models that assume uniform effects of independent variables on the dependent variable regardless of the direction of change, the NARDL model recognizes that the impacts of positive and negative changes in independent variables can differ. This advancement allows for exploring asymmetric effects, offering a more detailed and flexible analysis of causality relationships. Utilizing the NARDL model enhances our understanding of how variations in independent variables influence the dependent variable, thereby facilitating more precise and insightful decision-making based on these complex dynamics.

Therefore, the focus is to investigate if there are asymmetric influences in the long- and short-run dynamics between the gas prices by decomposing each time series into positive and negative increments. The bivariate configuration of the models will be the same as presented in Chapter 7. Hence, 20 bivariate models for Sample 1 and 42 bivariate models for Sample 2 will be assessed. As introduced in Chapter 6, the methodology chosen is the non-linear ARDL model (NARDL) developed by Shin, Yu and Greenwood-Nimmo (2014). This methodology became popular in causality assessment between time series for outperforming all other ordinary cointegration methods.

The analytical methodology of Pesaran, Shin and Smith (2001) for the bounds cointegration test used in Chapter 7 is also valid for the NARDL approach proposed by Shin, Yu and Greenwood-Nimmo (2014). Thus, the methodology explained in section 7.2 and illustrated in Figure 7.1 also applies to the asymmetric ARDL (NARDL) models. The variables' unit root and structural break analysis were already

presented in Chapter 4, suggesting dummy variables are also applied in the NARDL modelling.

## 8.2 Empirical Results of the Bivariate NARDL Models

### 8.2.1 Bivariate NARDL Models Applied to Sample 1

As explained in Chapter 6, the first step before applying the NARDL bounds cointegration Test proposed by Shin, Yu and Greenwood-Nimmo (2014) is decomposing the explanatory variables into positive and negative shocks. Recalling Equations (6.38) and (6.39), the positive and negative shocks of an explanatory variable can be formulated as follows:

$$p_{2,t}^+ = \sum_{j=1}^t \Delta p_{2,i}^+ = \sum_{j=1}^t \max(\Delta p_{2,j}, 0) \quad (8.1)$$

$$p_{2,t}^- = \sum_{k=1}^t \Delta p_{2,i}^- = \sum_{k=1}^t \min(\Delta p_{2,k}, 0) \quad (8.2)$$

As presented in Equation (6.40) of Chapter 6, the NARDL model proposed for the application of the bounds cointegration test is as follows:

$$\begin{aligned} \Delta p_{1,t} = & \beta_0 + \sum_{i=1}^p \lambda_i \Delta p_{1,t-i} + \sum_{j=0}^q \delta_{1j} \Delta p_{2,t-j}^+ + \sum_{k=0}^r \delta_{2k} \Delta p_{2,t-k}^- \\ & + \varphi_1 p_{1,t-1} + \varphi_2 p_{2,t-1}^+ + \varphi_3 p_{2,t-1}^- + q_t \end{aligned} \quad (8.3)$$

Where,  $\Delta p_1$  is the dependent variable. The coefficients  $\delta_{1i}$  and  $\delta_{2i}$  are the positive and negative asymmetric short-run terms of the explanatory variable  $p_2$ , respectively. The indices  $p$ ,  $q$ , and  $r$  are the optimal lag lengths of the explanatory variables in their first differences. The optimal lag lengths are determined by AIC. The long-run behaviour is dictated by the coefficients  $\varphi_1$ ,  $\varphi_2$ , and  $\varphi_3$ . Similar to  $\delta_{1i}$  and  $\delta_{2i}$ ,  $\varphi_2$  and  $\varphi_3$  are the coefficients representing the positive and negative long-run terms of the explanatory variable  $p_2$ , respectively. The F-statistic bounds cointegration test investigates the null hypothesis that the long-run terms  $\varphi_1$ ,  $\varphi_2$ , and  $\varphi_3$  jointly equal



to zero. If the bounds test confirms cointegration, the error correction representation of Equation (7.3) can be formulated as follows:

$$\Delta p_{1,t} = \beta_0 + \sum_{i=1}^p \lambda_i \Delta p_{1,t-i} + \sum_{j=0}^q \delta_{1j} \Delta p_{2,t-j}^+ + \sum_{k=0}^r \delta_{2k} \Delta p_{2,t-k}^- + \hat{\alpha} ECT_{t-1} + q_t \quad (8.4)$$

The error correction term ( $ECT_{t-1}$ ) substitutes the NARDL long-run terms ( $\varphi_1 p_{1,t-1} + \varphi_2 p_{2,t-1}^+ + \varphi_3 p_{2,t-1}^-$ ). The coefficient of the error correction term,  $\alpha$ , represents the speed of adjustment of the bivariate model towards long run equilibrium. There is convergence toward long-run equilibrium if the coefficient  $\alpha$  is negative and significant, otherwise ( $\alpha > 0$ ) the model is considered unstable and explosive. If  $\hat{\alpha}$  is significant, then  $p_2$  Granger-causes  $p_1$  in the long-run. The significance of  $\hat{\alpha}$  is assessed by its t-statistic value. The error correction term can be presented as follows:

$$ECT_{t-1} = p_{1,t-1} - (\beta + \beta_{expl}^+ p_{2,t}^+ + \beta_{expl}^- p_{2,t}^-) \quad (8.5)$$

where  $\beta_{expl}^+$  and  $\beta_{expl}^-$  represents the positive and negative loading factors of the explanatory variable  $p_2$  in the bivariate model.

Table 8.1 reports the results for the F-statistic bound test for the asymmetric ARDL models proposed by Shin, Yu and Greenwood-Nimmo (2014) applied to Sample 1. When applicable (null hypothesis is rejected), the speed of adjustment  $\hat{\alpha}$  and the long-run terms of  $p_2$ ,  $\varphi_2$  and  $\varphi_3$ , are also reported. The last column of the table informs by stating yes (Y) or no (N) the cases which the variables are cointegrated and the positive long-run coefficient  $\varphi_2$  is greater than the negative  $\varphi_3$ .

Table 8.1<sup>14</sup> Asymmetric ARDL (NARDL) Bivariate Models Applied to Sample 1 - Cointegration and ECM Long-Run Estimates.

Bivariate Model (Dependant – Explanatory)	NARDL Model Definition (AIC)	Bounds Test: F-statistics	I(0) 5% Critical Value	I(1) 5% Critical Value	Coint. (Y/N)	Speed of Adjustment $\hat{\alpha}$ (t-stats)	Loading Factor $\beta_{expl,t-1}^+$ (t) $\beta_{expl,t-1}^-$ (t)	$\hat{\alpha}$ $*(-\beta_{expl,t-1}^{+,-}) = \varphi_2^+$ (t-stats) $\varphi_3^-$ (t-stats)
LnHH – LnNBP	NARDL(1,0,2)	3.72	3.79	4.85	N	NA	NA	NA

<sup>14</sup> Table 8.1 continues on the next page.

Table 8.1 Asymmetric ARDL (NARDL) Bivariate Models Applied to Sample 1 - Cointegration and ECM Long-Run Estimates, Continued.

Bivariate Model (Dependant – Explanatory)	NARDL Model Definition (AIC)	Bounds Test: F-statistics	I(0) 5% Critical Value	I(1) 5% Critical Value	Coint. (Y/N)	Speed of Adjustment $\hat{\alpha}$ (t-stats)	Loading Factor $\beta_{expl_{t-1}^+}$ (t) $\beta_{expl_{t-1}^-}$ (t)	$\hat{\alpha} * (-\beta_{expl_{t-1}^+} + \beta_{expl_{t-1}^-}) = \varphi_2^+$ (t-stats) $\varphi_3^-$ (t-stats)
LnHH – LnRUS	NARDL(1,0,1)	4.44	3.79	4.85	N	NA	NA	NA
LnHH – LnJPN	NARDL(1,1,0)	4.56	3.79	4.85	N	NA	NA	NA
LnHH – LnOIL	NARDL(1,1,3)	3.27	3.79	4.85	N	NA	NA	NA
LnNBP – LnHH	NARDL(2,2,2)	4.14	3.79	4.85	N	NA	NA	NA
LnNBP – LnRUS	NARDL(2,0,1)	11.87***	3.79	4.85	Y	-0.220*** (-5.993)	1.047 <sup>+</sup> (6.745) 0.906 <sup>-</sup> (6.795)	0.230 <sup>+</sup> *** (4.308) 0.199 <sup>-</sup> *** (4.178)
LnNBP – LnJPN	NARDL(2,2,0)	6.91***	3.79	4.85	Y	-0.142*** (-4.574)	0.896 <sup>+</sup> (3.934) 0.941 <sup>-</sup> (4.102)	0.127 <sup>+</sup> *** (3.206) 0.134 <sup>-</sup> *** (3.148)
LnNBP – LnOIL	NARDL(2,0,2)	16.71***	3.79	4.85	Y	-0.208*** (-7.113)	1.196 <sup>+</sup> (7.899) 1.189 <sup>-</sup> (8.173)	0.248 <sup>+</sup> *** (6.468) 0.247 <sup>-</sup> *** (6.482)
LnRUS – LnHH	NARDL(4,0,2)	3.68	3.79	4.85	N	NA	NA	NA
LnRUS – LnNBP	NARDL(4,2,4)	5.48**	3.79	4.85	Y	-0.066*** (-4.073)	0.937 <sup>+</sup> (4.382) 0.998 <sup>-</sup> (4.714)	0.061 <sup>+</sup> *** (3.008) 0.065 <sup>-</sup> *** (3.209)
LnRUS – LnJPN	NARDL(4,4,4)	5.33**	3.79	4.85	Y	-0.112*** (-4.017)	0.679 <sup>+</sup> (5.420) 0.907 <sup>-</sup> (6.842)	0.076 <sup>+</sup> *** (2.667) 0.102 <sup>-</sup> *** (2.979)
LnRUS – LnOIL	NARDL(4,0,0)	12.65***	3.79	4.85	Y	-0.103*** (-6.189)	0.870 <sup>+</sup> (6.412) 0.970 <sup>-</sup> (7.028)	0.089 <sup>+</sup> *** (4.988) 0.099 <sup>-</sup> *** (5.365)
LnJPN – LnHH	NARDL(3,0,3)	9.12***	3.79	4.85	Y	-0.073*** (-5.254)	0.229 <sup>+</sup> (2.176) 0.037 <sup>-</sup> (0.361)	0.017 <sup>+</sup> *** (2.192) 0.003 <sup>-</sup> (0.366)
LnJPN – LnNBP	NARDL(2,0,1)	13.68***	3.79	4.85	Y	-0.100*** (-6.436)	0.328 <sup>+</sup> (5.074) 0.201 <sup>-</sup> (2.772)	0.033 <sup>+</sup> *** (4.219) 0.020 <sup>-</sup> *** (2.657)
LnJPN – LnRUS	NARDL(2,0,1)	12.15***	3.79	4.85	Y	-0.099*** (-6.064)	0.640 <sup>+</sup> (10.140) 0.313 <sup>-</sup> (2.772)	0.063 <sup>+</sup> *** (4.585) 0.031 <sup>-</sup> *** (2.660)
LnJPN – LnOIL	NARDL(3,2,2)	25.56***	3.79	4.85	Y	-0.129*** (-8.798)	0.715 <sup>+</sup> (9.949) 0.642 <sup>-</sup> (6.680)	0.092 <sup>+</sup> *** (7.774) 0.083 <sup>-</sup> *** (6.384)
LnOIL – LnHH	NARDL(2,4,1)	1.77	3.79	4.85	N	NA	NA	NA
LnOIL – LnNBP	NARDL(2,0,3)	2.31	3.79	4.85	N	NA	NA	NA
LnOIL – LnRUS	NARDL(2,0,0)	2.49	3.79	4.85	N	NA	NA	NA
LnOIL – LnJPN	NARDL(2,0,1)	1.78	3.79	4.85	N	NA	NA	NA

Notes: The I(0) and I(1) Bound Test critical values are reported from Pesaran, Shin, and Smith (2001), Critical values: Case III – constant and no trend.  $\hat{\alpha}$  is the speed of adjustment coefficient. \*\*\*, \*\* represent statistical significance at the 1% and 5% levels, respectively. The ARDL model definition ARDL(x,y,z) using the AIC approach represents that the lag length of the dependent variable is x and that the lag length explanatory variable representing positive increments is y and representing negative increments is z.

The analysis from Table 8.1 indicates that out of 20 bivariate models tested in Sample 1 using the NARDL approach, 10 exhibit long-run relationships, demonstrating cointegration. This finding includes all seven long-run relationships previously identified through the linear ARDL method, plus three additional cointegrated pairs: (LnRUS, LnJPN), (LnJPN, LnHH), and (LnJPN, LnRUS). These new pairs show significant differences between their positive ( $\varphi_2^+$ ) and negative ( $\varphi_3^-$ ) long-run coefficients. The analysis highlights the Japanese gas market (JPN) as the most responsive to other gas

prices and oil price changes. Following JPN, the UK's NBP and Russia's RUS gas markets are the next most reactive. Complementing the linear ARDL findings, the study now points to the Russian market as the leading influencer for European gas prices, overtaking oil prices in this role. Specifically, the NBP gas price tends to return to equilibrium more quickly when RUS prices are considered over oil prices. For the Russian gas market, the pair (LnRUS, LnJPN) shows the highest adjustment speed. Unlike the others, the Henry Hub (HH) and oil prices do not depend on long-run causality from the examined variables.

Following the robustness check methodology used to verify the reliability of the cointegration models in the previous chapter, we again include the GECON indicator as an exogenous variable in the asymmetric ARDL models. This ensures that the long-run relationships presented in Table 8.1 are indeed significant. Consequently, Table 8.2 presents the cointegration (long-run causalities) analysis for these bivariate ARDL models, including the GECON indicator, mirroring the format of Table 8.1.

Table 8.2<sup>15</sup> Asymmetric ARDL (NARDL) Models with GECON Indicator as Exogenous Variable Applied to Sample 1 - Cointegration and ECM Long-Run Estimates.

Bivariate Model (Dependant – Explanatory)	NARDL Model Definition (AIC)	Bounds Test: F-statistics	I(0) 5% Critical Value	I(1) 5% Critical Value	Coint. (Y/N)	Speed of Adjustment $\hat{\alpha}$ (t-stats)	Loading Factor $\beta_{expl_{t-1}}^+ (t)$ $\beta_{expl_{t-1}}^- (t)$
LnHH – LnNBP	NARDL(1,0,0)	6.55**	3.79	4.85	Y	-0.132*** (-4.453)	0.419 <sup>+</sup> (2.196) 0.553 <sup>-</sup> (2.957)
LnHH – LnRUS	NARDL(1,0,1)	4.24	3.79	4.85	N	NA	NA
LnHH – LnJPN	NARDL(1,2,0)	4.94	3.79	4.85	N	NA	NA
LnHH – LnOIL	NARDL(1,3,3)	3.65	3.79	4.85	N	NA	NA
LnNBP – LnHH	NARDL(2,2,2)	3.91	3.79	4.85	N	NA	NA
LnNBP – LnRUS	NARDL(2,0,1)	11.8***	3.79	4.85	Y	-0.222*** (-5.990)	1.059 <sup>+</sup> (6.854) 0.923 <sup>-</sup> (6.911)
LnNBP – LnJPN	NARDL(2,2,0)	6.89***	3.79	4.85	Y	-0.142*** (-4.567)	0.900 <sup>+</sup> (3.978) 0.966 <sup>-</sup> (4.216)
LnNBP – LnOIL	NARDL(2,0,2)	16.32***	3.79	4.85	Y	-0.208*** (-7.031)	1.198 <sup>+</sup> (7.846) 1.190 <sup>-</sup> (8.148)
LnRUS – LnHH	NARDL(4,0,2)	3.61	3.79	4.85	N	NA	NA
LnRUS – LnNBP	NARDL(4,2,4)	5.40**	3.79	4.85	Y	-0.065*** (-4.045)	0.943 <sup>+</sup> (4.293) 1.004 <sup>-</sup> (4.595)
LnRUS – LnJPN	NARDL(4,4,4)	4.83	3.79	4.85	N	NA	NA
LnRUS – LnOIL	NARDL(4,0,0)	13.27***	3.79	4.85	Y	-0.122*** (-6.338)	0.884 <sup>+</sup> (7.737) 0.971 <sup>-</sup> (8.409)

<sup>15</sup> Table 8.2 continues on the next page.

Table 8.2 Asymmetric ARDL (NARDL) Models with GECON Indicator as Exogenous Variable Applied to Sample 1 - Cointegration and ECM Long-Run Estimates, Continued.

Bivariate Model (Dependant – Explanatory)	NARDL Model Definition (AIC)	Bounds Test: F-statistics	I(0) 5% Critical Value	I(1) 5% Critical Value	Coint. (Y/N)	Speed of Adjustment $\hat{\alpha}$ (t-stats)	Loading Factor $\beta_{expl_{t-1}^+}$ (t) $\beta_{expl_{t-1}^-}$ (t)
LnJPN – LnHH	NARDL(3,0,3)	7.10***	3.79	4.85	Y	-0.122*** (-6.338)	0.231 <sup>+</sup> (2.135) 0.039 <sup>-</sup> (3.371)
LnJPN – LnNBP	NARDL(2,0,1)	11.27***	3.79	4.85	Y	-0.079*** (-4.851)	0.328 <sup>+</sup> (5.074) 0.201 <sup>-</sup> (2.772)
LnJPN – LnRUS	NARDL(2,0,1)	10.50***	3.79	4.85	Y	-0.097*** (-5.637)	0.650 <sup>+</sup> (9.717) 0.330 <sup>-</sup> (3.421)
LnJPN – LnOIL	NARDL(3,2,2)	24.24***	3.79	4.85	Y	-0.131*** (-8.567)	0.710 <sup>+</sup> (9.857) 0.635 <sup>-</sup> (6.582)
LnOIL – LnHH	NARDL(2,4,1)	2.16	3.79	4.85	N	NA	NA
LnOIL – LnNBP	NARDL(2,0,3)	3.15	3.79	4.85	N	NA	NA
LnOIL – LnRUS	NARDL(2,0,0)	7.16***	3.79	4.85	Y	-0.109*** (-4.657)	1.025 <sup>+</sup> (7.427) 1.109 <sup>-</sup> (5.354)
LnOIL – LnJPN	NARDL(2,2,2)	6.91***	3.79	4.85	Y	-0.065*** (-4.575)	1.517 <sup>+</sup> (4.043) 2.441 <sup>-</sup> (2.985)

Notes: The I(0) and I(1) Bound Test critical values are reported from Pesaran, Shin, and Smith (2001), Critical values: Case III – constant and no trend.  $\hat{\alpha}$  is the speed of adjustment coefficient. \*\*\*, \*\* represent statistical significance at the 1% and 5% levels, respectively. The ARDL model definition ARDL(x,y,z) using the AIC approach represents that the lag length of the dependent variable is x and that the lag length explanatory variable representing positive increments is y and representing negative increments is z.

Analysing the results from the asymmetric ARDL models with the GECON indicator included as an exogenous variable in Table 8.2, 17 out of 20 bivariate model cointegration outcomes within Sample 1 remain consistent with those in Table 8.1. Overall, the asymmetric bivariate ARDL models applied to Sample 1, with and without the GECON indicator, demonstrate an 85% consistency in cointegration results.

Figure 8.1 illustrates the long-run causality directions between the price time series in Sample 1 asymmetric ARDL models, as derived from the results in Tables 8.1 and 8.2.

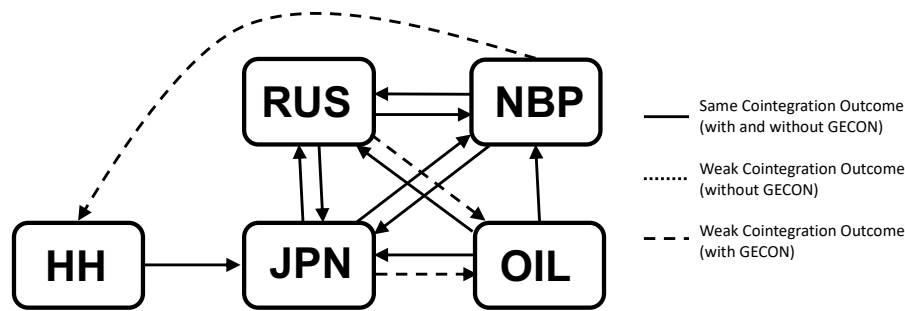


Figure 8.1 Diagram of the Asymmetric Long-Run Causality Directions of Sample 1.

As in the previous chapter, we will not extend the robustness check to asymmetric short-run causalities to avoid overly lengthening the results and discussion. This decision is also based on the high consistency observed in the cointegration analyses in Tables 8.1 and 8.2.

From Table 8.1, we can distinguish positive and negative asymmetries in the long run. When  $\varphi_2^+ > \varphi_3^-$  in Equation 8.3, positive shocks to the explanatory variable cause a larger deviation to the dependant variable from the long-run equilibrium than the negative shocks. Therefore, when shocks to the explanatory are positive, it takes more time to go back to equilibrium compared to negative shocks. Similarly, when  $\varphi_2^+ < \varphi_3^-$ , negative shocks to the explanatory variable cause a larger deviation to the dependant variable from the long-run equilibrium than the positive shocks. Figures 8.2 and 8.3 present a diagram illustrating the positive ( $\varphi_2^+ > \varphi_3^-$ ) and negative ( $\varphi_2^+ < \varphi_3^-$ ) long-run causality asymmetries in Sample 1, respectively. Using the values of the speed of adjustment  $\hat{\alpha}$  of each bivariate model, the diagram shows in each arrow the number of months necessary for the dependent variable to return to equilibrium in case of an external shock. For example, for the cointegrated pair (LnNBP, LnRUS), the speed of adjustment is  $-0.220$ ; thus, in a presence of a shock, it will take NBP  $(\frac{1}{0.220})$  4.55 months to return to equilibrium. This assessment is important to identify the leading and lagging markets in the long-run causality analysis.

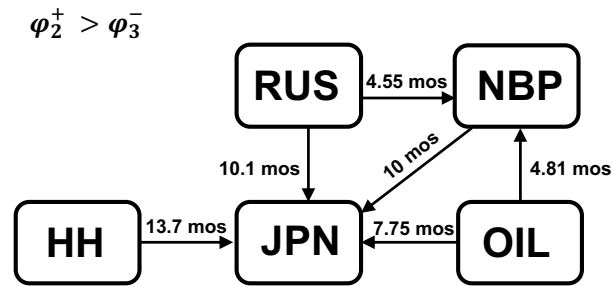


Figure 8.2 Diagram of the Positive Asymmetric Long-Run Causality Directions of Sample 1.

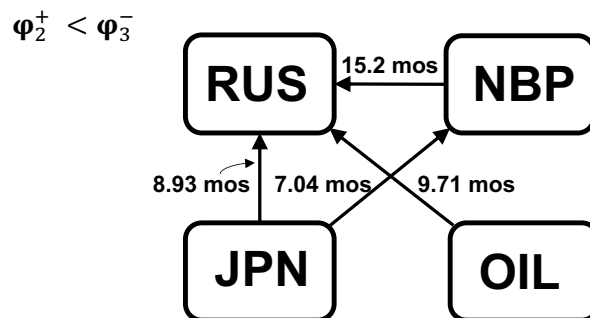


Figure 8.3 Diagram of the Negative Asymmetric Long-Run Causality Directions of Sample 1.

Analysing the 10 asymmetric long-run causality relationships derived from Table 8.1 and shown in Figure 8.1, breaking them down into positive and negative components reveals a distinct pattern. Six of these relationships exhibit a stronger reaction to positive shocks, as depicted in Figure 8.2, while the remaining four respond more to negative shocks. The analysis shows that positive shocks to the explanatory variables in the analysed cointegrated bivariate NARDL models of Sample 1 tend to push the dependent variable further away from its long-run equilibrium than negative shocks. Specifically, the Japanese (JPN) gas market is more susceptible to larger deviations from its long-run equilibrium due to positive changes in other variables within the sample. Conversely, the Russian (RUS) gas market shows a tendency to deviate more significantly from its equilibrium following negative shocks, especially from the British (NBP), and oil (OIL) markets.

The analysis of positive and negative adjustments in the explanatory variables' long-run behaviour from Table 8.1 reveals that in the two models, the positive and negative adjustments, though distinct, are considerably similar. The first instance involves the

variables LnRUS and LnNBP, where the adjustments for LnNBP show a positive change of 0.061 and a negative change of 0.065. The second instance is between LnNBP and LnOIL, where LnOIL's adjustments register a positive change of 0.248 and a nearly identical negative change of 0.247. This similarity in adjustments suggests that there is no significant asymmetry in the long-term relationships of these bivariate models.

To assess if the  $\varphi_2^+$  and  $\varphi_3^-$  are statistically equal or not, we use the Welch's t-test based on the Welch-Satterthwaite equation (Satterthwaite 1946). The Welch's t-test is used for testing whether two samples have the same population mean. Considering that  $\eta_1$  is the number of positive residuals and  $\eta_2$  is the number of negative residuals of the bivariate model;  $s_1$  and  $s_2$  are the standard errors of  $\varphi_2^+$  and  $\varphi_3^-$ , respectively. Welch's t-statistic is given by:

$$t = \frac{\widehat{\varphi}_2^+ - \widehat{\varphi}_3^-}{\sqrt{\frac{s_1^2}{\eta_1} + \frac{s_2^2}{\eta_2}}} \quad (8.6)$$

If the null hypothesis stands,  $t$  is approximately t-distributed with  $\nu$  degrees of freedom. The degrees of freedom  $\nu$  can be calculated through the Welch-Satterthwaite equation as follows:

$$\nu = \frac{\left(\frac{s_1^2}{\eta_1} + \frac{s_2^2}{\eta_2}\right)^2}{\left(\frac{s_1^4}{\eta_1^2(\eta_1 - 1)} + \frac{s_2^4}{\eta_2^2(\eta_2 - 1)}\right)} \quad (8.7)$$

Following calculating degrees of freedom, we determine the critical values for both positive and negative critical values using the inverse functions of the Student's t-distribution's right and left tails, with a significance level set at 5%. Should Welch's t-statistic, derived from Equation (8.6), fall within the critical range defined by these tails, we do not reject the null hypothesis. This outcome implies no significant difference between  $\varphi_2^+$  and  $\varphi_3^-$ . The findings, utilizing both the Welch's t-test and the Welch-Satterthwaite Equation for  $\varphi_2^+$  and  $\varphi_3^-$ , are detailed in Table 8.2.

Table 8.3 Welch's t-test and Welch-Satterthwaite results for null hypothesis,

$$H_0: \varphi_2^+ = \varphi_3^- \text{ (Sample 1).}$$

Bivariate Model	$\varphi_2^+$	$\varphi_3^-$	Welch's t-test $t$	Critical Value $p = 0.025$	Critical Value $p = 0.975$	$H_0: \varphi_2^+ = \varphi_3^-$ (Y/N)
LnNBP – LnOIL	0.248	0.247	0.1975	-1.9708	1.9708	Y
LnRUS – LnNBP	0.061	0.065	-1.4956	-1.9705	1.9705	Y

The results in Table 8.3 suggest that the pairs (LnRUS, LnNBP), and (LnNBP, LnOIL) do not have an asymmetric long-run relationship ( $\varphi_2^+ = \varphi_3^-$ ) and the causality asymmetries from these two pairs can be eliminated from Figures 8.2 and 8.3.

The assessment of asymmetric short-term causality within the cointegrated NARDL models, as shown in Table 8.1, involves conducting a joint F-test on the short-term coefficients  $\delta_{1j}$  and  $\delta_{2k}$  (Equation 8.3) for each explanatory variable. This joint F-test is applied to  $\delta_{1j}$  and  $\delta_{2k}$  separately to examine potential asymmetries in the short-term dynamics. Table 8.4 details the outcomes of these  $\delta_{1j}$  and  $\delta_{2k}$  coefficient estimations, chosen based on the optimal lag determined by the Akaike Information Criterion (AIC) for each bivariate model. Furthermore, Table 8.4 includes the results from the Wald test for joint significance, highlighting the presence of asymmetrical short-term causal relationships in the bivariate error ECMs.

Table 8.4<sup>16</sup> Asymmetric ARDL (NARDL) Bivariate Models Applied to Sample 1 - ECM Short-Run Estimates.

Bivariate Model (Dependant – Explanatory)	First-differenced explanatory coefficients $\delta_{1j}^+, \delta_{2k}^-$ (t-stats)	Wald Test: Coefficients Joint Significance - F-statistic <sup>+, -</sup>	Short-run Causality Relationship	Short-run Asymmetry (Y/N)
$\Delta \text{LnNBP} - \Delta \text{LnRUS}$	No short-run coefficients in ECM	NA	RUS $\rightarrow$ NBP	N
$\Delta \text{LnNBP} - \Delta \text{LnJPN}$	$\Delta \text{LnJPN}_{t-1}^+ = 0.95^{**}$ (2.209)	4.880 <sup>***</sup>	JPN $\rightarrow$ NBP	Y
$\Delta \text{LnNBP} - \Delta \text{LnOIL}$	$\Delta \text{LnOIL}_{t-1}^- = -0.493^{***}$ (-3.015)	9.090 <sup>***</sup>	OIL $\rightarrow$ NBP	Y
$\Delta \text{LnRUS} - \Delta \text{LnNBP}$	$\Delta \text{LnNBP}_{t-1}^+ = 0.105^{**}$ (2.038) $\Delta \text{LnNBP}_{t-1}^- = 0.021$ (0.410) $\Delta \text{LnNBP}_{t-2}^- = 0.124^{**}$ (2.566) $\Delta \text{LnNBP}_{t-3}^- = -0.105^{**}$ (-2.094)	4.153 <sup>***</sup> 3.820 <sup>***</sup>	NBP $\rightarrow$ RUS	N

<sup>16</sup> Table 8.4 continues on the next page.



Table 8.4 Asymmetric ARDL (NARDL) Bivariate Models Applied to Sample 1 - ECM Short-Run Estimates, Continued.

Bivariate Model (Dependant – Explanatory)	First-differenced explanatory coefficients $\delta_{1j}^+, \delta_{2k}^-$ (t-stats)	Wald Test: Coefficients Joint Significance - F-statistic <sup>+, -</sup>	Short-run Causality Relationship	Short-run Asymmetry (Y/N)
$\Delta \ln \text{RUS} - \Delta \ln \text{JPN}$	$\Delta \ln \text{JPN}_{t-1}^+ = 0.111$ (0.529) $\Delta \ln \text{JPN}_{t-2}^+ = 0.187$ (0.894) $\Delta \ln \text{JPN}_{t-3}^+ = 0.590^{***}$ (2.852) $\Delta \ln \text{JPN}_{t-1}^- = -0.409^{**}$ (-2.168) $\Delta \ln \text{JPN}_{t-2}^- = 0.407^{**}$ (2.129) $\Delta \ln \text{JPN}_{t-3}^- = -0.320$ (-1.849)	$3.711^{+**}$ $3.000^{-**}$	JPN → RUS	N
$\Delta \ln \text{RUS} - \Delta \ln \text{OIL}$	No short-run coefficients in ECM	NA	OIL → RUS	NA
$\Delta \ln \text{JPN} - \Delta \ln \text{HH}$	$\Delta \ln \text{HH}_{t-1} = 0.034$ (1.057) $\Delta \ln \text{HH}_{t-2} = 0.090^{***}$ (2.784)	$4.517^{-**}$	HH → JPN	Y
$\Delta \ln \text{JPN} - \Delta \ln \text{NBP}$	No short-run coefficients in ECM	NA	NBP → JPN	N
$\Delta \ln \text{JPN} - \Delta \ln \text{RUS}$	No short-run coefficients in ECM	NA	RUS → JPN	N
$\Delta \ln \text{JPN} - \Delta \ln \text{OIL}$	$\Delta \ln \text{OIL}_{t-1}^+ = -0.037$ (-0.652) $\Delta \ln \text{OIL}_{t-1}^- = -0.080$ (-1.674)	$0.425^+$ $2.802^-$	OIL → JPN	N

Notes:  $\hat{\delta}$  are coefficients of the short-run terms of the bivariate ECM. \*\*\*, \*\* represent statistical significance at the 1% and 5% levels, respectively.

When cointegration is not confirmed, the following model is proposed using the dependent and explanatory variables in their first differences to investigate short-run causality:

$$\Delta p_{1,t} = \gamma_0 + \sum_{i=1}^r \eta_i \Delta p_{1,t-i} + \sum_{j=0}^s \theta_{1j} \Delta p_{2,t-j}^+ + \sum_{k=0}^t \theta_{2k} \Delta p_{2,t-k}^- + v_t \quad (8.8)$$

In this equation,  $\eta_i$  represents the short-term coefficients of the dependent variable, while  $\theta_{1j}$  and  $\theta_{2k}$  correspond to the explanatory variable's positive and negative asymmetric short-term coefficients, respectively. The optimal lag lengths  $r$ ,  $s$  and  $t$  are selected based on the AIC. Assessing asymmetric short-term causality involves conducting separate joint F-tests on the short-term coefficients  $\theta_{1j}$  and  $\theta_{2k}$ . The causality in the short term is evaluated using the t-statistic of the short-term coefficients of the explanatory variable. The findings are presented in Table 8.5.

Table 8.5 AIC-Augmented Asymmetric ARDL (NARDL) Bivariate Models  
Applied to Sample 1 – Short-Run Estimates.

Bivariate Model (Dependant – Explanatory)	NARDL Model Definition (AIC)	First-differenced explanatory coefficients $\theta_{1j}^+, \theta_{2k}^-$ (t-stats)	Wald Test: Coeff. Joint Significance - F-statistic <sup>+, -</sup>	Short-run Causality Relationship	Short-run Asymm. (Y/N)
$\Delta \ln HH - \Delta \ln NBP$	ARDL(1,3,1)	$\Delta \ln NBP_{t-1}^+ = 0.053$ (0.451) $\Delta \ln NBP_{t-2}^+ = 0.144$ (1.313) $\Delta \ln NBP_{t-3}^+ = 0.150$ (1.647) $\Delta \ln NBP_{t-1}^- = 0.183$ (1.830)	4.811 <sup>+</sup> *** 3.349 <sup>-</sup>	NBP → HH	Y
$\Delta \ln HH - \Delta \ln RUS$	ARDL(1,3,0)	$\Delta \ln RUS_{t-1}^+ = -0.047$ (-0.195) $\Delta \ln RUS_{t-2}^+ = 0.765$ *** (3.269) $\Delta \ln RUS_{t-3}^+ = -0.482$ *** (-2.632)	4.058 <sup>+</sup> ***	RUS → HH	Y
$\Delta \ln HH - \Delta \ln JPN$	ARDL(1,0,0)	No short-run coefficients	NA	JPN → HH	N
$\Delta \ln HH - \Delta \ln OIL$	ARDL(1,2,2)	$\Delta \ln OIL_{t-1}^+ = -0.208$ (-1.043) $\Delta \ln OIL_{t-2}^+ = 0.366$ *** (2.602) $\Delta \ln OIL_{t-1}^- = -0.109$ -0.485) $\Delta \ln OIL_{t-2}^- = 0.375$ ** (2.138)	3.604 <sup>+</sup> ** 2.647 <sup>-</sup>	OIL → HH	Y
$\Delta \ln NBP - \Delta \ln HH$	ARDL(1,3,0)	$\Delta \ln HH_{t-1}^+ = -0.053$ (-0.412) $\Delta \ln HH_{t-2}^+ = -0.121$ (-1.075) $\Delta \ln HH_{t-3}^+ = 0.319$ *** (3.498)	5.041 <sup>+</sup> ***	HH → NBP	Y
$\Delta \ln RUS - \Delta \ln HH$	ARDL(4,1,1)	$\Delta \ln HH_{t-1}^+ = 0.113$ ** (2.232) $\Delta \ln HH_{t-1}^- = 0.093$ ** (2.117)	4.982 <sup>+</sup> ** 4.482 <sup>-</sup> **	HH → RUS	N
$\Delta \ln OIL - \Delta \ln HH$	ARDL(1,0,0)	No short-run coefficients	NA	HH → OIL	N
$\Delta \ln OIL - \Delta \ln NBP$	ARDL(1,0,0)	No short-run coefficients	NA	NBP → OIL	N
$\Delta \ln OIL - \Delta \ln RUS$	ARDL(1,0,0)	No short-run coefficients	NA	RUS → OIL	N
$\Delta \ln OIL - \Delta \ln JPN$	ARDL(1,0,0)	No short-run coefficients	NA	JPN → OIL	N

Notes:  $\hat{\theta}$  are the coefficients of the ARDL models. \*\*\*, \*\* represent statistical significance at the 1% and 5% levels, respectively. The ARDL model definition ARDL(x,y,z) using the AIC approach represents that the lag length of the dependent variable is x and that the lag length explanatory variable representing positive increments is y and representing negative increments is z.

From Tables 8.4 and 8.5, we examine the short-term dynamics through both the short-run estimates of the ECM and the estimates from the bivariate NARDL models (which are not cointegrated) when applied to first-differenced variables. The findings reveal that, within Sample 1, 10 out of 20 bivariate NARDL models exhibit a short-run causal relationship. Among these, 7 models demonstrate an asymmetric causal relationship. By comparing these outcomes with the results from the linear ARDL models applied to Sample 1, as discussed in Chapter 7, it is evident that the directions of causality are quite similar between the short-run relationships of the NARDL models and those observed in the linear ARDL models.

The primary distinction is that, within the NARDL framework, the short-term impact on NBP gas price is influenced not only by OIL but also by HH and JPN variables. In

these instances, the variables exert an asymmetric effect on NBP in the short term, which accounts for the emergence of new relationships identified through the NARDL model analysis. Figure 8.4 depicts the directions of short-run causality among the variables in Sample 1, as detailed in the findings presented in Tables 8.3 and 8.4.

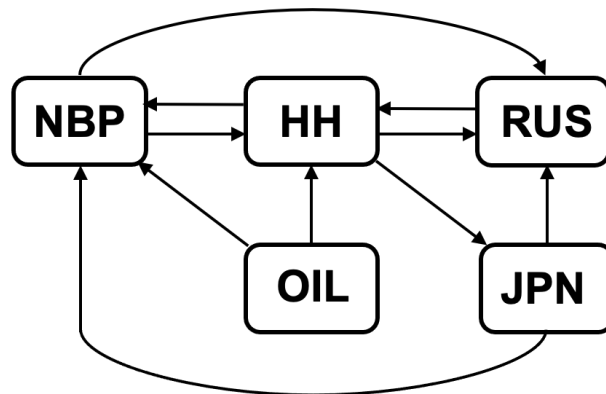


Figure 8.4 Diagram of the NARDL Short-Run Causality Directions of Sample 1.

In the next section, the results of the NARDL bounds cointegration test applied to Sample 2 will be presented.

### 8.2.2 Bivariate NARDL Models Applied to Sample 2

Using the methodology outlined in section 8.2.1, Table 8.6 displays the NARDL cointegration bounds test outcomes for Sample 2. It presents the F-statistic value for each bivariate model. In cases where cointegration is confirmed, it includes the speed of adjustment  $\alpha$  and the long-run asymmetric terms of the explanatory variable,  $\varphi_2$  and  $\varphi_3$ .

Table 8.6<sup>17</sup> Asymmetric ARDL (NARDL) Bivariate Models Applied to Sample 2 - Cointegration and ECM Long-Run Estimates.

Bivariate Model (Dependant – Explanatory)	NARDL Model Definition (AIC)	Bounds Test: F-statistics	I(0) 5% I(1) 5% Critical Value	Coint. (Y/N)	Speed of Adjustment $\hat{\alpha}$ (t-stats)	Loading Factor $\beta_{\text{expl}_{t-1}^+}$ (t-stats) $\beta_{\text{expl}_{t-1}^-}$ (t-stats)	$\hat{\alpha}$ $*(-\beta_{\text{expl}_{t-1}^{+,-}})$ = $\varphi_2^+$ (t-stats) $\varphi_3^-$ (t-stats)
LnHH – LnNBP	NARDL(1,0,0)	6.149**	3.79 4.85	Y	-0.201*** (-4.334)	0.722+*** (2.772) 0.693-*** (3.297)	0.146+*** (2.894) 0.140-*** (3.247)
LnHH – LnRUS	NARDL(1,3,0)	3.71	3.79 4.85	N	NA	NA	NA

<sup>17</sup> Table 8.6 continues on the next pages.

Table 8.6 Asymmetric ARDL (NARDL) Bivariate Models Applied to Sample 2 - Cointegration and ECM Long-Run Estimates, Continued.

Bivariate Model (Dependant – Explanatory)	NARDL Model Definition (AIC)	Bounds Test: F-statistics	I(0) 5% I(1) 5% Critical Value	Coint. (Y/N)	Speed of Adjustment $\hat{\alpha}$ (t-stats)	Loading Factor $\beta_{\text{expl}_{t-1}}^+$ (t-stats) $\beta_{\text{expl}_{t-1}}^-$ (t-stats)	$\hat{\alpha}$ $*(-\beta_{\text{expl}_{t-1}}^{+,-})$ = $\phi_2^+$ (t-stats) $\phi_3^-$ (t-stats)
LnHH – LnTTF	NARDL(1,1,0)	3.87	3.79 4.85	N	NA	NA	NA
LnHH – LnJPN	NARDL(1,0,0)	2.50	3.79 4.85	N	NA	NA	NA
LnHH – LnALNG	NARDL(1,1,0)	4.35	3.79 4.85	N	NA	NA	NA
LnHH – LnOIL	NARDL(1,1,3)	2.41	3.79 4.85	N	NA	NA	NA
LnNBP – LnHH	NARDL(2,0,1)	3.53	3.79 4.85	N	NA	NA	NA
LnNBP – LnRUS	NARDL(2,0,1)	13.35***	3.79 4.85	Y	-0.658*** (-6.390)	0.708+*** (10.149) 0.717-*** (15.369)	0.466+*** (4.473) 0.472-*** (4.891)
LnNBP – LnTTF	NARDL(4,3,2)	12.07***	3.79 4.85	Y	-0.591*** (-6.078)	0.765+*** (13.953) 0.804-*** (17.340)	0.452+*** (5.263) 0.476-*** (5.452)
LnNBP – LnJPN	NARDL(2,0,0)	3.34	3.79 4.85	N	NA	NA	NA
LnNBP – LnALNG	NARDL(2,1,0)	8.16***	3.79 4.85	Y	-0.358*** (-4.994)	0.567+*** (6.552) 0.601-*** (8.811)	0.203+*** (3.600) 0.215-*** (4.017)
LnNBP – LnOIL	NARDL(2,2,0)	5.61**	3.79 4.85	Y	-0.194*** (-4.142)	0.724+*** (3.000) 0.715-*** (4.162)	0.141+*** (2.759) 0.139-*** (3.324)
LnRUS – LnHH	NARDL(2,0,2)	1.61	3.79 4.85	N	NA	NA	NA
LnRUS – LnNBP	NARDL(4,2,2)	4.54	3.79 4.85	N	NA	NA	NA
LnRUS – LnTTF	NARDL(4,2,2)	8.92***	3.79 4.85	Y	-0.206*** (-5.221)	0.876+*** (7.064) 0.951-*** (8.827)	0.180+*** (4.606) 0.196-*** (4.882)
LnRUS – LnJPN	NARDL(2,4,0)	1.62	3.79 4.85	N	NA	NA	NA
LnRUS – LnALNG	NARDL(4,0,4)	7.808***	3.79 4.85	Y	-0.260*** (-4.889)	0.806+*** (8.332) 0.821-*** (11.038)	0.209+*** (4.429) 0.214-*** (4.341)
LnRUS – LnOIL	NARDL(2,3,0)	5.96**	3.79 4.85	Y	-0.150*** (-4.269)	0.925+*** (3.746) 0.923-*** (5.308)	0.139+*** (3.456) 0.138-*** (3.946)
LnTTF – LnHH	NARDL(3,3,2)	1.06	3.79 4.85	N	NA	NA	NA
LnTTF – LnNBP	NARDL(3,3,2)	27.28***	3.79 4.85	Y	-0.695*** (-9.133)	1.167+*** (11.609) 1.111-*** (14.524)	0.811+*** (8.114) 0.772-*** (8.248)
LnTTF – LnRUS	NARDL(4,0,1)	10.60***	3.79 4.85	Y	-0.467*** (-5.693)	1.116+*** (5.860) 0.910-*** (8.404)	0.521+*** (4.678) 0.425-*** (4.430)
LnTTF – LnJPN	NARDL(3,0,1)	1.44	3.79 4.85	N	NA	NA	NA
LnTTF – LnALNG	NARDL(3,0,0)	7.43***	3.79 4.85	Y	-0.332*** (-4.765)	0.857+*** (4.386) 0.796-*** (5.540)	0.284+*** (4.234) 0.264-*** (4.318)
LnTTF – LnOIL	NARDL(3,1,0)	4.34	3.79 4.85	N	NA	NA	NA

Table 8.6 Asymmetric ARDL (NARDL) Bivariate Models Applied to Sample 2 - Cointegration and ECM Long-Run Estimates, Continued.

Bivariate Model (Dependant – Explanatory)	NARDL Model Definition (AIC)	Bounds Test: F-statistics	I(0) 5% I(1) 5% Critical Value	Coint. (Y/N)	Speed of Adjustment $\hat{\alpha}$ (t-stats)	Loading Factor $\beta_{\text{expl}_{t-1}^+}$ (t-stats) $\beta_{\text{expl}_{t-1}^-}$ (t-stats)	$\hat{\alpha}$ $*(-\beta_{\text{expl}_{t-1}^+}, -)$ = $\phi_2^+$ (t-stats) $\phi_3^-$ (t-stats)
LnJPN – LnHH	NARDL(2,0,0)	6.06**	3.79 4.85	Y	-0.086*** (-4.304)	0.482+** (1.886) 0.414- (1.569)	0.042+** (2.290) 0.036- (1.903)
LnJPN – LnNBP	NARDL(2,1,0)	9.92***	3.79 4.85	Y	-0.144*** (-5.507)	0.489+*** (4.067) 0.354-*** (3.162)	0.070+*** (3.683) 0.051-*** (3.088)
LnJPN – LnRUS	NARDL(2,3,4)	10.90***	3.79 4.85	Y	-0.196*** (-5.776)	0.501+*** (4.905) 0.303-** (2.862)	0.098+*** (3.211) 0.059-** (2.213)
LnJPN – LnTTF	NARDL(2,0,0)	7.85***	3.79 4.85	Y	-0.127*** (-4.897)	0.369+*** (3.255) 0.275-** (2.556)	0.047+*** (3.081) 0.035-** (2.577)
LnJPN – LnALNG	NARDL(2,0,4)	7.57***	3.79 4.85	Y	-0.140*** (-4.813)	0.407+*** (3.717) 0.315-** (2.795)	0.057+*** (2.895) 0.044-** (2.306)
LnJPN – LnOIL	NARDL(2,4,4)	62.86***	3.79 4.85	Y	-0.465*** (-13.874)	0.811+*** (25.002) 0.823-*** (22.461)	0.378+*** (12.332) 0.383-*** (12.235)
LnALNG – LnHH	NARDL(2,0,2)	2.14	3.79 4.85	N	NA	NA	NA
LnALNG – LnNBP	NARDL(1,0,2)	8.41***	3.79 4.85	Y	-0.289*** (-5.070)	1.224+*** (6.590) 1.251-*** (8.461)	0.353+*** (3.924) 0.361-*** (4.206)
LnALNG – LnRUS	NARDL(3,4,1)	8.56***	3.79 4.85	Y	-0.447*** (-5.117)	1.034+*** (7.948) 1.020-*** (12.282)	0.463+*** (3.781) 0.456-*** (4.155)
LnALNG – LnTTF	NARDL(2,1,1)	3.848	3.79 4.85	N	NA	NA	NA
LnALNG – LnJPN	NARDL(2,0,0)	2.143	3.79 4.85	N	NA	NA	NA
LnALNG – LnOIL	NARDL(2,2,0)	8.52***	3.79 4.85	Y	-0.243*** (-5.102)	1.269+*** (5.029) 1.188-*** (6.677)	0.308+*** (4.364) 0.288-*** (4.801)
LnOIL – LnHH	NARDL(2,0,0)	6.38***	3.79 4.85	Y	-0.197*** (-4.414)	0.207+ (1.158) 0.114- (0.626)	0.041+ (1.103) 0.022- (0.614)
LnOIL – LnNBP	NARDL(2,2,3)	9.70***	3.79 4.85	Y	-0.298*** (-5.449)	0.634+*** (4.966) 0.485-*** (3.980)	0.189+*** (3.638) 0.145-*** (3.174)
LnOIL – LnRUS	NARDL(2,0,0)	10.99***	3.79 4.85	Y	-0.338*** (-5.795)	0.551+*** (5.422) 0.345-*** (3.784)	0.186+*** (3.604) 0.117-*** (2.891)
LnOIL – LnTTF	NARDL(2,0,0)	9.03***	3.79 4.85	Y	-0.271*** (-5.254)	0.360+*** (3.574) 0.260-*** (2.762)	0.098+*** (2.911) 0.070-** (2.432)
LnOIL – LnJPN	NARDL(2,1,1)	6.68***	3.79 4.85	Y	-0.202*** (-4.518)	0.324+ (1.130) 0.021- (0.059)	0.066+ (0.870) 0.004- (0.060)
LnOIL – LnALNG	NARDL(2,0,2)	7.63***	3.79 4.85	Y	-0.254*** (-4.829)	0.327+*** (2.859) 0.228- (2.005)	0.083+** (2.213) 0.058- (1.688)

Notes: The I(0) and I(1) Bound Test critical values are reported from Pesaran, Shin, and Smith (2001), Critical values: Case III – constant and no trend.  $\hat{\alpha}$  is the speed of adjustment coefficient. \*\*\*, \*\* represent statistical significance at the 1% and 5% levels, respectively. ARDL(x,y,z) using the AIC approach represents that the lag length of the dependent variable is x and that the lag length explanatory variable representing positive increments is y and representing negative increments is z.

Table 8.6 reveals that among 42 NARDL bivariate models in Sample 2, 26 exhibit long-run relationships (cointegrated models). Comparing the linear ARDL bounds

cointegration test results in Chapter 7 (Table 7.4) with the NARDL approach in Table 8.6, 24 cointegrating pairs are consistent across both methods. However, three pairs identified in the linear ARDL test were not confirmed in Table 8.6: (LnHH, LnALNG), (LnRUS, LnNBP), and (LnTTF, LnOIL). These may require alternative methods for long-run equilibrium confirmation. Conversely, Table 8.6 uncovers two new cointegrating pairs not found in the linear assessment: (LnRUS, LnTTF), and (LnRUS, LnALNG). Cointegration here relies on a stronger long-run negative loading factor ( $\varphi_3^-$ ), which was detected through the breakdown of the explanatory variable into positive and negative increments.

The JPN gas price is the most reactive variable in the long run, affecting all other six variables in Sample 2. The coefficients  $\hat{\alpha}$  (speed of adjustment) in the bivariate Error Correction Model (ECM) involving the Japanese (JPN) gas price as the dependent variable and the other five gas prices range from 0.086 to 0.196. However, in the case of the bivariate ECM pairing JPN with OIL, the speed of adjustment is 0.465, indicating a robust long-run relationship where OIL appears to be the leading market influencing JPN gas prices.

The NBP gas price is the second most reactive variable and is influenced in the long run by the other four variables. Also, NBP is strongly affected in the long run by the two European gas prices, RUS and TTF, with a speed of adjustment in the ECM models equal to 0.658 and 0.591, respectively. Thus, based on the results, RUS is considered the leading market over NBP.

The third most reactive gas prices are the TTF, RUS, and ALNG, and are affected by three variables in the long run. While NBP is the leading market causing TTF in the long run with a speed of adjustment of 0.695, the RUS gas price leads the long-run relationship over ALNG with a speed of adjustment of 0.477. HH is the least reactive gas price in Sample 2. HH is influenced only by NBP in the long run, with a speed of adjustment of 0.201. Therefore, despite the HH being considered an isolated market in Sample 1, there is evidence from Table 8.6 that the HH has a long-run relation with at least one European gas price.

Similar to the cointegration analysis robustness check conducted in the previous subsection for Sample 1, Table 8.7 presents the results of the bounds cointegration

assessment for the NARDL models, including the GECON indicator as an exogenous variable.

Table 8.7<sup>18</sup> Asymmetric ARDL (NARDL) Models with GECON Indicator as Exogenous Variable Applied to Sample 2 - Cointegration and ECM Long-Run Estimates.

Bivariate Model (Dependant – Explanatory)	NARDL Model Definition (AIC)	Bounds Test: F-statistics	I(0) 5% I(1) 5% Critical Value	Coint. (Y/N)	Speed of Adjustment $\hat{\alpha}$ (t-stats)	Loading Factor $\beta_{expl_{t-1}}^+$ (t-stats) $\beta_{expl_{t-1}}^-$ (t-stats)
LnHH – LnJPN	NARDL(1,0,0)	4.57	3.79 4.85	N	NA	NA
LnHH – LnALNG	NARDL(1,1,0)	4.35	3.79 4.85	N	NA	NA
LnHH – LnOIL	NARDL(1,1,3)	2.39	3.79 4.85	N	NA	NA
LnHH – LnNBP	NARDL(1,0,0)	7.27***	3.79 4.85	Y	-0.245*** (-4.714)	0.507*** (2.304) 0.526**** (2.974)
LnHH – LnRUS	NARDL(1,1,3)	4.82	3.79 4.85	N	NA	NA
LnHH – LnTTF	NARDL(1,1,3)	4.70	3.79 4.85	N	NA	NA
LnNBP – LnHH	NARDL(2,0,1)	3.46	3.79 4.85	N	NA	NA
LnNBP – LnRUS	NARDL(2,0,1)	13.15***	3.79 4.85	Y	-0.658*** (-6.390)	0.726**** (9.258) 0.728**** (14.169)
LnNBP – LnTTF	NARDL(4,3,2)	11.97***	3.79 4.85	Y	-0.593*** (-6.052)	0.777**** (12.713) 0.809**** (15.885)
LnNBP – LnJPN	NARDL(2,0,0)	3.38	3.79 4.85	N	NA	NA
LnNBP – LnALNG	NARDL(2,1,0)	8.02***	3.79 4.85	Y	-0.298*** (-3.983)	0.537**** (6.552) 0.651**** (8.811)
LnNBP – LnOIL	NARDL(2,2,0)	5.81**	3.79 4.85	Y	-0.194*** (-4.221)	0.857**** (2.847) 0.809**** (3.791)
LnRUS – LnHH	NARDL(2,0,2)	1.71	3.79 4.85	N	NA	NA
LnRUS – LnNBP	NARDL(4,2,2)	6.66**	3.79 4.85	Y	-0.194*** (-4.221)	1.072**** (9.650) 1.124**** (12.907)
LnRUS – LnTTF	NARDL(4,2,2)	8.83***	3.79 4.85	Y	-0.205*** (-5.197)	0.876**** (7.064) 0.951**** (8.827)
LnRUS – LnJPN	NARDL(2,4,0)	2.02	3.79 4.85	N	NA	NA
LnRUS – LnALNG	NARDL(4,0,4)	7.73***	3.79 4.85	Y	-0.263*** (-4.865)	0.789**** (7.552) 0.809**** (10.064)

<sup>18</sup> Table 8.7 continues on the next pages.

Table 8.7 Asymmetric ARDL (NARDL) Models with GECON Indicator as Exogenous Variable Applied to Sample 2 - Cointegration and ECM Long-Run Estimates, Continued.

Bivariate Model (Dependant – Explanatory)	NARDL Model Definition (AIC)	Bounds Test: F-statistics	I(0) 5% I(1) 5% Critical Value	Coint. (Y/N)	Speed of Adjustment $\hat{\alpha}$ (t-stats)	Loading Factor $\beta_{expl,t-1}^+$ (t-stats) $\beta_{expl,t-1}^-$ (t-stats)
LnRUS – LnOIL	NARDL(2,3,0)	6.21**	3.79 4.85	Y	-0.148*** (-4.359)	1.071+*** (3.409) 1.024-*** (4.653)
LnTTF – LnHH	NARDL(3,3,4)	2.94	3.79 4.85	N	NA	NA
LnTTF – LnNBP	NARDL(3,3,2)	29.95***	3.79 4.85	Y	-0.761*** (-9.571)	1.060+*** (10.860) 1.033-*** (13.915)
LnTTF – LnRUS	NARDL(4,0,1)	10.60***	3.79 4.85	Y	-0.484*** (-5.693)	1.034+*** (5.090) 0.866-*** (7.488)
LnTTF – LnJPN	NARDL(3,0,0)	1.77	3.79 4.85	N	NA	NA
LnTTF – LnALNG	NARDL(3,0,0)	7.79***	3.79 4.85	Y	-0.369*** (-4.878)	0.718+*** (3.955) 0.696-*** (5.186)
LnTTF – LnOIL	NARDL(3,1,0)	3.83	3.79 4.85	N	NA	NA
LnJPN – LnHH	NARDL(2,0,0)	5.75**	3.79 4.85	Y	-0.090*** (-4.191)	0.522+*** (2.027) 0.444- (1.703)
LnJPN – LnNBP	NARDL(2,2,3)	10.60***	3.79 4.85	Y	-0.183*** (-5.595)	0.537+*** (4.767) 0.397-*** (3.640)
LnJPN – LnRUS	NARDL(2,3,4)	14.30***	3.79 4.85	Y	-0.248*** (-6.618)	0.553+*** (7.108) 0.341-*** (4.292)
LnJPN – LnTTF	NARDL(2,0,0)	7.98***	3.79 4.85	Y	-0.143** (-4.939)	0.405+*** (3.839) 0.302-*** (3.061)
LnJPN – LnALNG	NARDL(2,1,4)	8.65***	3.79 4.85	Y	-0.180*** (-5.145)	0.450+*** (5.158) 0.349-*** (3.936)
LnJPN – LnOIL	NARDL(2,4,4)	65.15***	3.79 4.85	Y	-0.475*** (-14.126)	0.822+*** (25.671) 0.839-*** (22.917)
LnALNG – LnHH	NARDL(2,0,2)	2.35	3.79 4.85	N	NA	NA
LnALNG – LnNBP	NARDL(1,0,2)	9.95***	3.79 4.85	Y	-0.292*** (-6.399)	1.184+*** (5.925) 1.221-*** (7.718)
LnALNG – LnRUS	NARDL(3,4,1)	8.46***	3.79 4.85	Y	-0.447*** (-5.088)	1.015+*** (6.730) 1.009-*** (10.707)
LnALNG – LnTTF	NARDL(2,1,1)	3.84	3.79 4.85	N	NA	NA
LnALNG – LnJPN	NARDL(2,0,0)	2.39	3.79 4.85	N	NA	NA
LnALNG – LnOIL	NARDL(2,2,0)	9.24***	3.79 4.85	Y	-0.245*** (-5.316)	1.495+*** (4.762) 1.344-*** (6.106)
LnOIL – LnHH	NARDL(2,0,0)	6.42***	3.79 4.85	Y	-0.195*** (-4.431)	0.122+ (0.600) 0.039- (0.195)



Table 8.7 Asymmetric ARDL (NARDL) Models with GECON Indicator as Exogenous Variable Applied to Sample 2 - Cointegration and ECM Long-Run Estimates, Continued.

LnOIL – LnNBP	NARDL(2,1,3)	9.20***	3.79 4.85	Y	-0.290*** (-5.306)	0.594+*** (4.415) 0.458*** (3.593)
LnOIL – LnRUS	NARDL(2,0,0)	10.85***	3.79 4.85	Y	-0.318*** (-5.225)	0.529+*** (4.922) 0.329*** (3.455)
LnOIL – LnTTF	NARDL(2,0,0)	8.82***	3.79 4.85	Y	-0.268*** (-5.191)	0.334+*** (3.095) 0.243*** (2.403)
LnOIL – LnJPN	NARDL(4,0,1)	7.72***	3.79 4.85	Y	0.171*** (4.858)	2.230+*** (1.985) 2.476*** (1.757)
LnOIL – LnALNG	NARDL(2,0,2)	7.56***	3.79 4.85	Y	-0.250*** (-4.807)	0.298+*** (2.426) 0.202*** (1.683)

Notes: The  $I(0)$  and  $I(1)$  Bound Test critical values are reported from Pesaran, Shin, and Smith (2001), Critical values: Case III – constant and no trend.  $\hat{\alpha}$  is the speed of adjustment coefficient. \*\*\*, \*\* represent statistical significance at the 1% and 5% levels, respectively. The ARDL model definition ARDL(x,y,z) using the AIC approach represents that the lag length of the dependent variable is x and that the lag length explanatory variable representing positive increments is y and representing negative increments is z.

Comparing the results from Tables 8.6 and 8.7, we notice that there are no discrepancies in the cointegration results. When the GECON indicator is added as an exogenous variable to the NARDL bivariate models, 100% of the models consistently identify cointegration.

Figure 8.5 illustrates the long-run causality directions between the time series according to the results reported in Table 8.6 and 8.7.

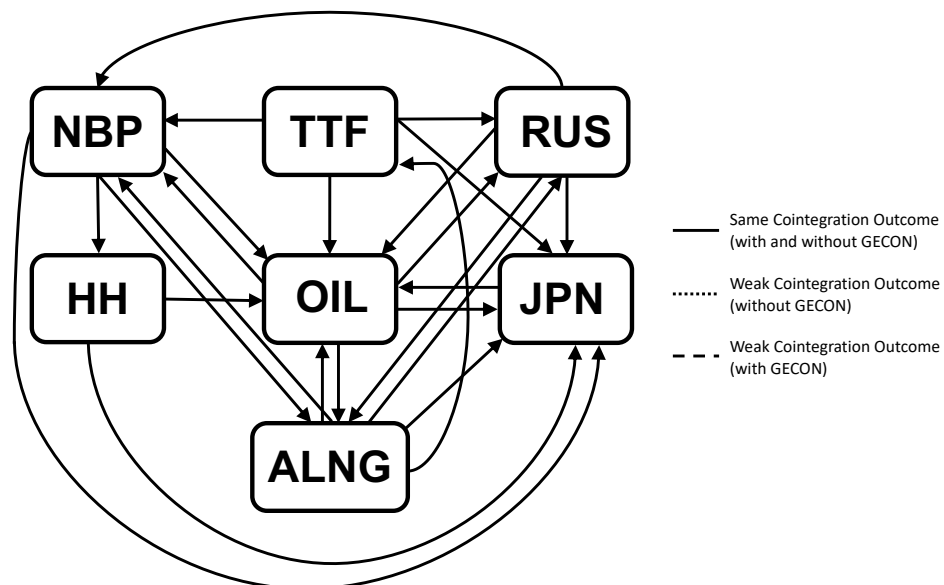


Figure 8.5 Diagram of the Asymmetric Long-Run Causality Directions of Sample 2.

Table 8.6 allows us to distinguish cointegrated models exhibiting positive and negative asymmetries in the long run by examining their long-run loading factors,  $\varphi_2^+$  and  $\varphi_3^-$ . Figures 8.6 and 8.7 visually depict these positive ( $\varphi_2^+ > \varphi_3^-$ ) and negative ( $\varphi_2^+ < \varphi_3^-$ ) long-run causality asymmetries in Sample 2, respectively. Similar to the diagrams in Sample 1, the arrows indicating the direction of causality illustrate the number of months required for the dependent variable to adjust following an external shock.

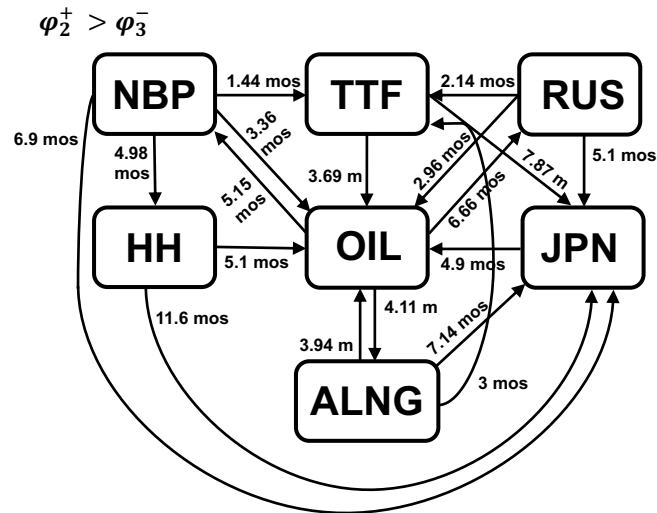


Figure 8.6 Diagram of the Positively Asymmetric Long-Run Causality Directions of Sample 2.

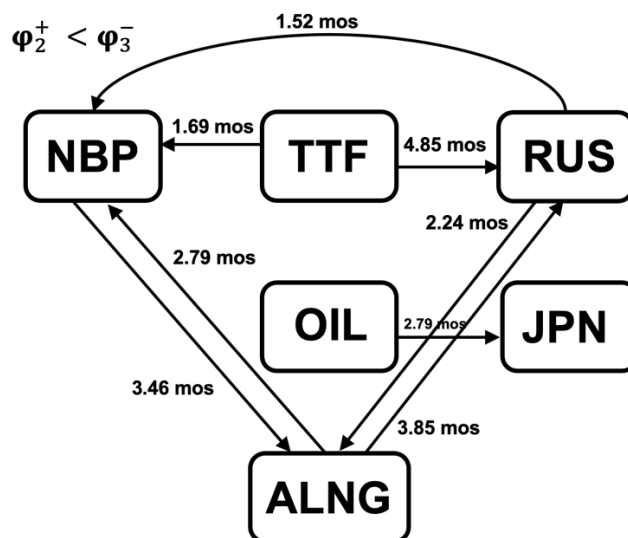


Figure 8.7 Diagram of the Negatively Asymmetric Long-Run Causality Directions of Sample 2.

By splitting the long-run causality relationships depicted in Figure 8.5 by positive and negative long-run asymmetries, there are 18 bivariate cointegrated models with greater positive loading factors and 8 with greater negative loading factors. For example, in Sample 1, there is a greater proportion of positive asymmetries in the long run, which makes a positive increment of the explanatory variable more effective than the dependent variable. Comparing the results from Sample 1, in this new sample period, the JPN gas price will deviate more through positive increments of all other five gas prices. In contrast, negative increments of OIL will have a larger effect on its long-run deviation. TTF and OIL suffer a larger deviation from positive increments of all its cointegrating pairs.

On the other hand, the RUS gas will have a larger deviation in the long run from negative increments of TTF and ALNG, as well as from positive increments of OIL. Similarly, the NBP gas price is caused by greater intensity in the long run by negative increments of TTF, ALNG, and RUS. Finally, HH, which is affected in the long-run only by NBP, has a greater impact from positive increments of the explanatory variable.

In Sample 2, similar to Sample 1, we observe three instances where the positive and negative asymmetric loading factors exhibit very close values. Firstly, the cointegrating pair (LnNBP, LnOIL) demonstrates positive ( $\varphi_2^+$ ) and negative ( $\varphi_3^-$ ) loading factors of 0.141 and 0.139, respectively. Secondly, (LnRUS, LnOIL) exhibits positive ( $\varphi_2^+$ ) and negative ( $\varphi_3^-$ ) loading factors of 0.139 and 0.138, respectively. Thirdly, (LnNBP, LnOIL) shows positive ( $\varphi_2^+$ ) and negative ( $\varphi_3^-$ ) loading factors of 0.378 and 0.383, respectively. To confirm the asymmetric long-run causality in these cases, we apply Welch's t-test and the Welch-Satterthwaite equation, as detailed in the previous subsection. Table 8.8 presents the results of these tests applied to  $\varphi_2^+$  and  $\varphi_3^-$  for the described cases.

Table 8.8 Welch's t-test and Welch-Satterthwaite results for null hypothesis,  $H_0: \varphi_2^+ = \varphi_3^-$  (Sample 2).

Bivariate Model	$\varphi_2^+$	$\varphi_3^-$	Welch's t-test <i>t</i>	Critical Value <i>p</i> = 0.025	Critical Value <i>p</i> = 0.975	$H_0: \varphi_2^+ = \varphi_3^-$ (Y/N)
LnNBP – LnOIL	0.141	0.139	0.2315	-1.9814	1.9814	Y
LnRUS – LnOIL	0.139	0.138	0.1425	-1.9820	1.9820	Y
LnJPN – LnOIL	0.378	0.383	-0.8691	-1.9812	1.9812	Y

The findings in Table 8.8 suggest that the pairs (LnNBP, LnOIL), (LnRUS, LnOIL), and (LnJPN, LnOIL) lack an asymmetric long-run relationship, indicating that  $\varphi_2^+ = \varphi_3^-$ . Subsequently, we evaluate the short-run terms of the depicted cointegrating pairs in Figure 8.5 by examining a joint F-statistic of their coefficients ( $\delta_{1j}$  and  $\delta_{2k}$ ). Table 8.9 shows the results of the Wald test for joint significance, which investigates the asymmetric short-run causality of the cointegrated bivariate ECMs.

Table 8.9<sup>19</sup> Asymmetric ARDL (NARDL) Bivariate Models Applied to Sample 2 - Cointegration and ECM Short-Run Estimates.

Bivariate Model (Dependant – Explanatory)	First-differenced explanatory coefficients $\delta_{1j}^+, \delta_{2k}^-$ (t-stats)	Wald Test: Coefficients Joint Significance - F-statistic <sup>+, -</sup>	Short-run Causality Relationship	Short-run Asymmetry (Y/N)
$\Delta \text{LnHH} - \Delta \text{LnNBP}$	No short-run coefficients in ECM	NA	NBP → HH	N
$\Delta \text{LnNBP} - \Delta \text{LnRUS}$	No short-run coefficients in ECM	NA	RUS → NBP	N
$\Delta \text{LnNBP} - \Delta \text{LnTTF}$	$\Delta \text{LnTTF}_{t-1}^+ = -0.015 (-0.145)$ $\Delta \text{LnTTF}_{t-2}^+ = -0.298^{***} (-2.981)$ $\Delta \text{LnTTF}_{t-1}^- = -0.467^{***} (-3.983)$	4.471 <sup>+</sup> *** 15.864 <sup>-</sup> ***	TTF → NBP	N
$\Delta \text{LnNBP} - \Delta \text{LnALNG}$	No short-run coefficients in ECM	NA	LNG → NBP	N
$\Delta \text{LnNBP} - \Delta \text{LnOIL}$	$\Delta \text{LnOIL}_{t-1}^+ = -0.429 (-1.889)$	3.568 <sup>+</sup>	OIL → NBP	N
$\Delta \text{LnRUS} - \Delta \text{LnTTF}$	No short-run coefficients in ECM	NA	TTF → RUS	N
$\Delta \text{LnRUS} - \Delta \text{LnALNG}$	$\Delta \text{LnLNG}_{t-1}^- = -0.012 (-0.135)$ $\Delta \text{LnLNG}_{t-2}^- = -0.141 (-1.687)$ $\Delta \text{LnLNG}_{t-3}^- = 0.149 (1.745)$	2.123 <sup>-</sup>	LNG → RUS	N
$\Delta \text{LnRUS} - \Delta \text{LnOIL}$	$\Delta \text{LnOIL}_{t-1}^+ = 0.269 (1.620)$ $\Delta \text{LnOIL}_{t-2}^+ = -0.552^{***} (-3.310)$	6.332 <sup>+</sup> ***	OIL → RUS	Y
$\Delta \text{LnTTF} - \Delta \text{LnNBP}$	$\Delta \text{LnNBP}_{t-1}^+ = -0.631^{***} (-2.915)$ $\Delta \text{LnNBP}_{t-1}^- = -0.403 (-1.853)$	8.497 <sup>+</sup> *** 3.434 <sup>-</sup>	NBP → TTF	Y
$\Delta \text{LnTTF} - \Delta \text{LnNBP}$	No short-run coefficients in ECM	NA	RUS → TTF	N
$\Delta \text{LnTTF} - \Delta \text{LnALNG}$	No short-run coefficients in ECM	NA	LNG → TTF	N
$\Delta \text{LnJPN} - \Delta \text{LnHH}$	No short-run coefficients in ECM	NA	HH → JPN	N
$\Delta \text{LnJPN} - \Delta \text{LnNBP}$	No short-run coefficients in ECM	NA	NBP → JPN	N
$\Delta \text{LnJPN} - \Delta \text{LnRUS}$	$\Delta \text{LnRUS}_{t-1}^+ = -0.042 (-0.488)$ $\Delta \text{LnRUS}_{t-2}^+ = -0.229^{**} (-2.599)$ $\Delta \text{LnRUS}_{t-1}^- = -0.073 (-0.884)$ $\Delta \text{LnRUS}_{t-2}^- = -0.019 (-2.632)$ $\Delta \text{LnRUS}_{t-3}^- = 0.215^{***} (2.704)$	3.462 <sup>+</sup> *** 3.302 <sup>-</sup> ***	RUS → JPN	N
$\Delta \text{LnJPN} - \Delta \text{LnTTF}$	No short-run coefficients in ECM	NA	TTF → JPN	N
$\Delta \text{LnJPN} - \Delta \text{LnALNG}$	$\Delta \text{LnLNG}_{t-1}^- = 0.015 (0.313)$ $\Delta \text{LnLNG}_{t-2}^- = 0.001 (0.028)$ $\Delta \text{LnLNG}_{t-3}^- = 0.116^{**} (2.446)$	1.995 <sup>-</sup>	LNG → JPN	N

<sup>19</sup> Table 8.9 continues on the next page.

Table 8.9 Asymmetric ARDL (NARDL) Bivariate Models Applied to Sample 2 - Cointegration and ECM Short-Run Estimates, Continued.

Bivariate Model (Dependant – Explanatory)	First-differenced explanatory coefficients $\delta_{1j}^+, \delta_{2k}^-$ (t-stats)	Wald Test: Coefficients Joint Significance - F-statistic <sup>+, -</sup>	Short-run Causality Relationship	Short-run Asymmetry (Y/N)
$\Delta \text{LnJPN} - \Delta \text{LnOIL}$	$\Delta \text{LnOIL}_{t-1}^+ = -0.366^{***} (-5.271)$ $\Delta \text{LnOIL}_{t-2}^+ = -0.354^{***} (-5.261)$ $\Delta \text{LnOIL}_{t-3}^+ = -0.337^{***} (-5.275)$ $\Delta \text{LnOIL}_{t-1}^- = -0.393^{***} (-6.985)$ $\Delta \text{LnOIL}_{t-2}^- = -0.381^{***} (-6.859)$ $\Delta \text{LnOIL}_{t-3}^- = -0.232^{***} (-3.755)$	24.582 <sup>***</sup> 30.787 <sup>***</sup>	OIL → JPN	N
$\Delta \text{LnALNG} - \Delta \text{LnNBP}$	$\Delta \text{LnNBP}_{t-1} = 0.360^{**} (1.995)$	3.980 <sup>**</sup>	NBP → LNG	Y
$\Delta \text{LnALNG} - \Delta \text{LnRUS}$	$\Delta \text{LnRUS}_{t-1}^+ = -0.477 (-1.789)$ $\Delta \text{LnRUS}_{t-2}^+ = 0.539^{**} (2.169)$ $\Delta \text{LnRUS}_{t-3}^+ = 1.206^{***} (5.407)$	3.544 <sup>***</sup>	RUS → LNG	Y
$\Delta \text{LnALNG} - \Delta \text{LnOIL}$	$\Delta \text{LnOIL}_{t-1}^+ = -0.555 (-1.938)$	3.756 <sup>+</sup>	OIL → LNG	N
$\Delta \text{LnOIL} - \Delta \text{LnHH}$	No short-run coefficients in ECM	NA	HH → OIL	N
$\Delta \text{LnOIL} - \Delta \text{LnNBP}$	$\Delta \text{LnNBP}_{t-1}^+ = -0.170 (-1.358)$ $\Delta \text{LnNBP}_{t-1} = 0.098 (0.749)$ $\Delta \text{LnNBP}_{t-2}^- = -0.347^{***} (-2.834)$	1.844 <sup>+</sup> 4.137 <sup>**</sup>	NBP → OIL	Y
$\Delta \text{LnOIL} - \Delta \text{LnRUS}$	No short-run coefficients in ECM	NA	RUS → OIL	N
$\Delta \text{LnOIL} - \Delta \text{LnTTF}$	No short-run coefficients in ECM	NA	TTF → OIL	N
$\Delta \text{LnOIL} - \Delta \text{LnJPN}$	No short-run coefficients in ECM	NA	JPN → OIL	N
$\Delta \text{LnOIL} - \Delta \text{LnALNG}$	$\Delta \text{LnLNG}_{t-1}^- = -0.181 (-1.974)$	3.897 <sup>-</sup>	LNG → OIL	N

Notes:  $\delta$  are coefficients of the short-run terms of the bivariate ECM. \*\*\*, \*\* represent statistical significance at the 1% and 5% levels, respectively.

Additionally, we employ the NARDL model using the variables in their first differences to examine the short-run causality for models lacking a long-run relationship, as depicted in Table 8.6. The outcomes are presented in Table 8.10.

Table 8.10<sup>20</sup> AIC-Augmented Asymmetric ARDL (NARDL) Bivariate Models Applied to Sample 2 - Short-Run Estimates.

Bivariate Model (Dependant – Explanatory)	NARDL Model Definition (AIC)	First-differenced explanatory coefficients $\theta_{1j}^+, \theta_{2k}^-$ (t-stats)	Wald Test: Coefficients Joint Signif. - F-stat. <sup>+, -</sup>	Short-run Causality	Short-run Asym. (Y/N)
$\Delta \text{LnHH} - \Delta \text{LnRUS}$	ARDL(1,3,0)	$\Delta \text{LnRUS}_{t-1}^+ = -0.103 (-0.359)$ $\Delta \text{LnRUS}_{t-2}^+ = 0.861^{***} (3.041)$ $\Delta \text{LnRUS}_{t-3}^+ = -0.475^{**} (-2.148)$	3.378 <sup>***</sup>	RUS → HH	Y

<sup>20</sup> Table 8.10 continues on the next pages.

Table 8.10 AIC-Augmented Asymmetric ARDL (NARDL) Bivariate Models  
Applied to Sample 2 - Short-Run Estimates, Continued.

Bivariate Model (Dependant – Explanatory)	NARDL Model Definition (AIC)	First-differenced explanatory coefficients $\theta_{1j}^+, \theta_{2k}^-$ (t-stats)	Wald Test: Coefficients Joint Signif. - F-stat. <sup>+, -</sup>	Short-run Causality	Short-run Asym. (Y/N)
$\Delta \ln HH - \Delta \ln TTF$	ARDL(1,1,2)	$\Delta \ln TTF_{t-1}^+ = 0.411^{***} (2.691)$ $\Delta \ln TTF_{t-1}^- = -0.150 (-1.849)$ $\Delta \ln TTF_{t-2}^- = 0.199^{**} (2.215)$	$7.241^{+***}$ $2.924^-$	RUS $\rightarrow$ HH	Y
$\Delta \ln HH - \Delta \ln JPN$	ARDL(1,0,0)	No short-run coefficients	NA	JPN $\rightarrow$ HH	N
$\Delta \ln HH - \Delta \ln ALNG$	ARDL(1,4,0)	$\Delta \ln LNG_{t-1}^+ = 0.056 (0.382)$ $\Delta \ln LNG_{t-2}^+ = 0.064 (0.452)$ $\Delta \ln LNG_{t-3}^+ = 0.340^{**} (2.413)$ $\Delta \ln LNG_{t-4}^+ = -0.261^{**} (-2.272)$	$2.485^+$	LNG $\rightarrow$ HH	N
$\Delta \ln HH - \Delta \ln OIL$	ARDL(1,3,2)	$\Delta \ln OIL_{t-1}^+ = -0.171 (-0.687)$ $\Delta \ln OIL_{t-2}^+ = 0.117 (0.512)$ $\Delta \ln OIL_{t-3}^+ = 0.420^{**} (2.206)$ $\Delta \ln OIL_{t-1}^- = 0.021 (0.066)$ $\Delta \ln OIL_{t-2}^- = 0.817^{***} (2.876)$	$3.625^{+**}$ $5.157^{-***}$	OIL $\rightarrow$ HH	N
$\Delta \ln NBP - \Delta \ln HH$	ARDL(4,3,2)	$\Delta \ln HH_{t-1}^+ = 0.261 (1.307)$ $\Delta \ln HH_{t-2}^+ = -0.396^{**} (-2.228)$ $\Delta \ln HH_{t-3}^+ = 0.487^{***} (3.295)$ $\Delta \ln HH_{t-1}^- = -0.249 (-1.412)$ $\Delta \ln HH_{t-2}^- = 0.356^{**} (2.232)$	$3.756^{+***}$ $2.523^-$	HH $\rightarrow$ NBP	Y
$\Delta \ln NBP - \Delta \ln JPN$	ARDL(4,0,0)	No short-run coefficients	NA	JPN $\rightarrow$ NBP	N
$\Delta \ln RUS - \Delta \ln HH$	ARDL(1,4,0)	$\Delta \ln HH_{t-1}^+ = 0.349^{***} (3.891)$ $\Delta \ln HH_{t-2}^+ = -0.293^{***} (-3.069)$ $\Delta \ln HH_{t-3}^+ = -0.072 (-0.799)$ $\Delta \ln HH_{t-4}^+ = 0.210^{***} (2.818)$	$5.480^{+***}$	HH $\rightarrow$ RUS	Y
$\Delta \ln RUS - \Delta \ln NBP$	ARDL(1,2,4)	$\Delta \ln NBP_{t-1}^+ = 0.360^{***} (3.218)$ $\Delta \ln NBP_{t-2}^+ = 0.241^{**} (2.391)$ $\Delta \ln NBP_{t-1}^- = 0.561^{***} (5.429)$ $\Delta \ln NBP_{t-2}^- = -0.043 (-0.427)$ $\Delta \ln NBP_{t-3}^- = -0.181^{**} (-2.036)$ $\Delta \ln NBP_{t-4}^- = 0.270^{***} (3.579)$	$12.882^{+***}$ $11.704^{-***}$	NBP $\rightarrow$ RUS	N
$\Delta \ln RUS - \Delta \ln JPN$	ARDL(1,0,2)	$\Delta \ln JPN_{t-1}^- = -0.647 (-1.893)$ $\Delta \ln JPN_{t-2}^- = 0.775^{**} (2.587)$	$3.646^{-**}$	JPN $\rightarrow$ RUS	Y
$\Delta \ln TTF - \Delta \ln HH$	ARDL(2,0,2)	$\Delta \ln HH_{t-1}^- = 0.707^{***} (3.775)$ $\Delta \ln HH_{t-2}^- = -0.263 (-1.736)$	$7.141^{-***}$	HH $\rightarrow$ NBP	Y
$\Delta \ln TTF - \Delta \ln JPN$	ARDL(2,0,2)	No short-run coefficients	NA	JPN $\rightarrow$ TTF	N
$\Delta \ln TTF - \Delta \ln OIL$	ARDL(2,0,1)	$\Delta \ln OIL_{t-1}^- = 0.788^{**} (2.222)$	$4.937^{-**}$	OIL $\rightarrow$ TTF	Y
$\Delta \ln ALNG - \Delta \ln HH$	ARDL(1,3,0)	$\Delta \ln HH_{t-1}^+ = 0.164 (0.823)$ $\Delta \ln HH_{t-2}^+ = -0.051 (-0.323)$ $\Delta \ln HH_{t-3}^+ = 0.293^{**} (2.228)$	$3.220^{+**}$	HH $\rightarrow$ LNG	Y
$\Delta \ln ALNG - \Delta \ln TTF$	ARDL(1,0,0)	No short-run coefficients	NA	TTF $\rightarrow$ LNG	N

Table 8.10 AIC-Augmented Asymmetric ARDL (NARDL) Bivariate Models  
Applied to Sample 2 - Short-Run Estimates, Continued.

Bivariate Model (Dependant – Explanatory)	NARDL Model Definition (AIC)	First-differenced explanatory coefficients $\theta_{1j}^+, \theta_{2k}^-$ (t-stats)	Wald Test: Coefficients Joint Signif. - F-stat. <sup>+, -</sup>	Short-run Causality	Short-run Asym. (Y/N)
$\Delta \ln \text{ALNG} - \Delta \ln \text{JPN}$	ARDL(1,0,0)	No short-run coefficients	NA	JPN → LNG	N

Notes:  $\hat{\theta}$  are the coefficients of the ARDL models. \*\*\*, \*\* represent statistical significance at the 1% and 5% levels, respectively. The ARDL model definition ARDL(x,y,z) using the AIC approach represents that the lag length of the dependent variable is x and that the lag length explanatory variable representing positive increments is y and representing negative increments is z.

Analysing Tables 8.9 and 8.10, we find that among the 42 NARDL bivariate models, 18 demonstrate short-run causality of the explanatory variable over the dependent variable. Within these, 13 exhibit asymmetrical relationships, while 5 show symmetrical ones. Additionally, among the models lacking a long-run relationship (as seen in Table 8.8), short-run causality is confirmed in 10 out of 16 cases. Considering both the long- and short-run causality assessments, 6 models indicate an absence of any causal connection.

Comparing these findings with the ARDL linear short-run assessment in Chapter 6 reveals significant disparities in short-run causality directions. Notably, in the NARDL analysis (Tables 8.9 and 8.10), four other variables influence the RUS gas price as the most reactive variable in the short run. Interestingly, although previously identified in the ARDL assessment, the short-run causality of TTF and ALNG over RUS is absent in the refined NARDL bivariate models.

The second most reactive gas prices in the short run are HH, TTF, and ALNG, each affected by three prices. Notably, two additional short-run causalities over HH are observed in the asymmetrical assessment. Comparing linear and asymmetrical ARDL assessments, TTF and ALNG are affected by an additional variable in the short run. While JPN is influenced by RUS and OIL, similar to the linear ARDL assessment, NBP is affected by OIL and HH. The asymmetrical NARDL model reveals the short-run causality of HH over NBP. Lastly, OIL is the least reactive market but experiences asymmetrical short-run effects from NBP.

The short-run causality directions in Tables 8.9 and 8.10 are illustrated in Figure 8.8 below.

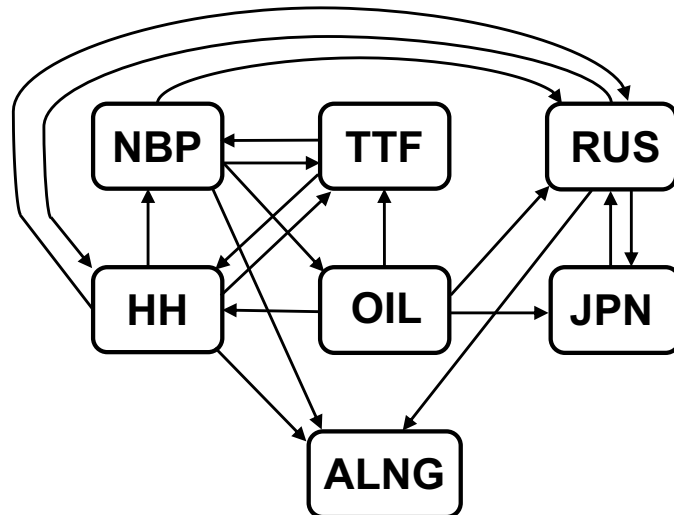


Figure 8.8 Diagram of the NARDL Short-Run Causality Directions of Sample 2.

### 8.3 Concluding Remarks

This chapter explored the connection between natural gas and oil prices using the Non-linear ARDL (NARDL) bounds cointegration test, a refined method proposed by Shin, Yu and Greenwood-Nimmo (2014). This approach examines how changes in one variable affect the other, accommodating positive and negative shocks. We applied this test to the same datasets as Chapter 7, allowing for a comparison between Linear and Non-linear ARDL methods in assessing the relationship between gas and oil prices.

A similar robustness check to the one used in the previous chapter was conducted in this chapter's NARDL bounds cointegration test. We included the GECON indicator as an exogenous variable in the asymmetric bivariate ARDL to assess the reliability of the cointegrated models based solely on price time series. The results showed considerable consistency between the two assessments (with and without the GECON variable), particularly in the cointegration outcomes related to Sample 2, which showed 100% consistency. Therefore, the discussion of asymmetric long-run causality (cointegration) will focus on Tables 8.1 and 8.6.

In summary, Table 8.11 outlines the main findings of this chapter. It presents the number of causal relationships discovered through the NARDL bivariate models applied to Samples 1 and 2. Additionally, it illustrates the proportion of these causal relationships that display asymmetry and categorizes them accordingly.



In certain instances, asymmetrical causality in a bivariate NARDL model implies that the dependent variable responds differently to positive or negative changes in the explanatory variable. In other cases, it suggests that only a positive or negative alteration of the explanatory variable holds statistical significance over the dependent variable.

Table 8.11 Summary of the Asymmetric ARDL (NARDL) Bivariate Models Results.

	<b>Sample 1 2001 - 2020</b>	<b>Sample 2 2010 - 2020</b>
Number of Variables	5	7
Number of long-run relationships (NARDL)	10 out of 20	26 out of 42
Positively Asymmetric Long-run Causality	5	16
Negatively Asymmetric Long-run Causality	3	7
Symmetric Long-run Causality	2	3
Number of short-run relationships (NARDL)	10 out of 20	18 out of 42
Positively Asymmetric Short-run Causality	5	9
Negatively Asymmetric Short-run Causality	2	4
Symmetric Short-run Causality	3	5

This chapter's analysis reveals an enhancement in detecting long-run and short-run relationships in Samples 1 and 2, respectively, by comparing the findings in Table 8.11 with those from Chapter 7 (Table 7.14) concerning Linear ARDL bivariate models. Specifically, adopting the asymmetrical cointegration method uncovered three more long-run cointegrating pairs in Sample 1, increasing the total to 10 beyond the 7 identified using linear ARDL models. In Sample 2, the application of Nonlinear ARDL (NARDL) bivariate models led to identifying 18 short-run causality relationships, marking a 38% increase over those found with linear ARDL models. This improvement is attributed to the NARDL models' capacity to account for

asymmetrical effects between the variables, an aspect the linear models could not capture. Nevertheless, the comparison shows no significant difference in the number of long-run relationships in Sample 2 and short-run relationships in Sample 1 when contrasting symmetric and asymmetric ARDL models. For short-run causality, the NARDL method highlighted that approximately 70% of causalities in each sample were asymmetric, with a tendency towards more positive asymmetries.

The chapter brings to light several key insights. Firstly, it demonstrates that traditional linear ARDL models may overlook complex causal dynamics between variables, such as gas and oil prices, that can be captured through the disaggregation of variables into positive and negative changes. Secondly, it establishes that a significant portion of both short- and long-run relationships exhibit asymmetry, with positive changes in variables exerting a stronger causal influence than negative ones in both samples.

# CHAPTER 9

## VAR ASSESSMENT OF CAUSALITY RELATIONSHIPS BETWEEN NATURAL GAS AND OIL PRICES

### 9.1 Introduction

The two previous chapters explored the integration within natural gas markets using linear and asymmetrical ARDL models across two distinct datasets. These investigations revealed an increase in the causality relationships among natural gas prices when comparing an earlier dataset from 2001 with a more recent one from 2010. Additionally, the dynamics between natural gas and oil prices were examined, leading to the observation that oil prices have less influence on natural gas prices than before in the latter and more contemporary sample period. Instead, the relationship among natural gas prices has emerged as more significant.

A main drawback of these bivariate ARDL and NARDL models is that they tend to overestimate the actual number of causal links. Addressing this issue, this chapter uses a VAR that includes all prices to discard incorrect causality links.

A three-variable example can demonstrate the emergence of the drawback mentioned above. Assume that time series A, B and C belong to a data-generating process in which A causes B, A causes C, and no other causal link exists. Furthermore, assume that the lag structure defining the propagation delay is shorter for  $A \rightarrow B$  than for  $A \rightarrow C$ . These causal relationships should be accurately detectable in a three-variable VAR model. The problem occurs when bivariate models (VAR or ARDL) are used. The bivariate specification will tend to detect an apparent  $B \rightarrow C$  that does not exist in the data-generating process. This is illustrated in Figure 9.1. In reality, A causes both B and C with different delays. Hence, causal relationships such as  $B \rightarrow C$  that do not appear in the full VAR of this chapter should be rejected.

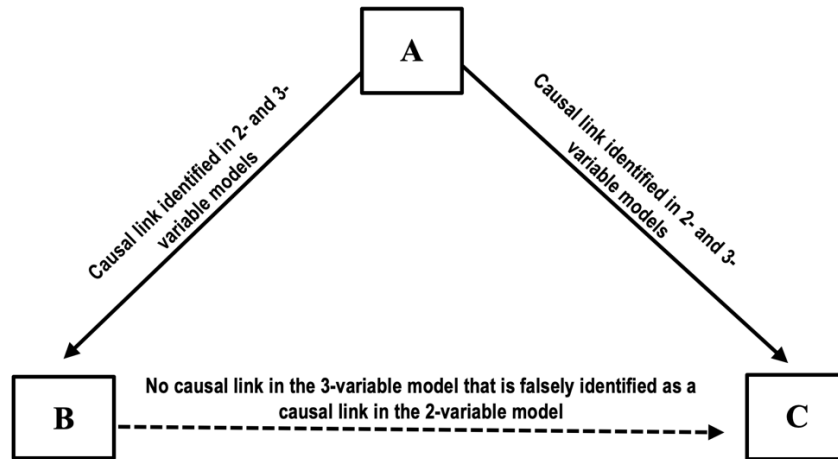


Figure 9.1 Causality Detection Trap in Bivariate Models.

To address the above and obtain refined information on dynamics interrelationships, a unified VAR (Vector Autoregression) model encompassing all relevant time series is estimated. The estimation is implemented on Sample 2 from Chapters 7 and 8 only, which includes seven time series (6 representing natural gas prices and one for Brent crude oil) from July 2010 to February 2020, because it is the most relevant subsample.

The VAR model's causality will be examined using the Toda and Yamamoto (1995) test, complemented by in-depth insights gained from impulse response functions and forecast error variance decomposition. While the ARDL approach previously helped identify long-term equilibrium relationships and causality directions, the forthcoming VAR analysis aims to quantify the effects of dynamic shocks. This methodological shift is also adequate by the nature of the variables within Sample 2, which, unlike Sample 1's mix of  $I(0)$  and  $I(1)$  variables, exclusively contains  $I(1)$  variables, aligning with the findings from the Unit Root tests discussed in Chapter 4.

## 9.2 Empirical Results of the VAR Model Applied to Sample 2

To explore the interrelationships between all variables in Sample 2, we have developed a VAR model that consists of 7 equations. Within this framework, every time series from the dataset is treated as an endogenous component within the system's 7-equation structure. An equation for each variable incorporates  $n$  lagged values within the system's variables. Thus, the structure of the VAR model examined in this chapter is outlined as follows:

$$\begin{aligned}
LnHH_t = & \beta_0 + \sum_{i=1}^n \beta_{1i} LnHH_{t-i} + \sum_{i=1}^n \beta_{2i} LnNBP_{t-i} + \sum_{i=1}^n \beta_{3i} LnRUS_{t-i} \\
& + \sum_{i=1}^n \beta_{4i} LnTTF_{t-i} + \sum_{i=1}^n \beta_{5i} LnJPN_{t-i} + \sum_{i=1}^n \beta_{6i} LnALNG_{t-i} \\
& + \sum_{i=1}^n \beta_{7i} LnOIL_{t-i} + u_{1t}
\end{aligned} \tag{9.1}$$

$$\begin{aligned}
LnNBP_t = & \gamma_0 + \sum_{i=1}^n \gamma_{1i} LnHH_{t-i} + \sum_{i=1}^n \gamma_{2i} LnNBP_{t-i} + \sum_{i=1}^n \gamma_{3i} LnRUS_{t-i} \\
& + \sum_{i=1}^n \gamma_{4i} LnTTF_{t-i} + \sum_{i=1}^n \gamma_{5i} LnJPN_{t-i} + \sum_{i=1}^n \gamma_{6i} LnALNG_{t-i} \\
& + \sum_{i=1}^n \gamma_{7i} LnOIL_{t-i} + u_{2t}
\end{aligned} \tag{9.2}$$

$$\begin{aligned}
LnRUS_t = & \delta_0 + \sum_{i=1}^n \delta_{1i} LnHH_{t-i} + \sum_{i=1}^n \delta_{2i} LnNBP_{t-i} + \sum_{i=1}^n \delta_{3i} LnRUS_{t-i} \\
& + \sum_{i=1}^n \delta_{4i} LnTTF_{t-i} + \sum_{i=1}^n \delta_{5i} LnJPN_{t-i} + \sum_{i=1}^n \delta_{6i} LnALNG_{t-i} \\
& + \sum_{i=1}^n \delta_{7i} LnOIL_{t-i} + u_{3t}
\end{aligned} \tag{9.3}$$

$$\begin{aligned}
LnTTF_t = & \theta_0 + \sum_{i=1}^n \theta_{1i} LnHH_{t-i} + \sum_{i=1}^n \theta_{2i} LnNBP_{t-i} + \sum_{i=1}^n \theta_{3i} LnRUS_{t-i} \\
& + \sum_{i=1}^n \theta_{4i} LnTTF_{t-i} + \sum_{i=1}^n \theta_{5i} LnJPN_{t-i} + \sum_{i=1}^n \theta_{6i} LnALNG_{t-i} \\
& + \sum_{i=1}^n \theta_{7i} LnOIL_{t-i} + u_{4t}
\end{aligned} \tag{9.4}$$

$$\begin{aligned}
LnJPN_t = & \vartheta_0 + \sum_{i=1}^n \vartheta_{1i} LnHH_{t-i} + \sum_{i=1}^n \vartheta_{2i} LnNBP_{t-i} + \sum_{i=1}^n \vartheta_{3i} LnRUS_{t-i} \\
& + \sum_{i=1}^n \vartheta_{4i} LnTTF_{t-i} + \sum_{i=1}^n \vartheta_{5i} LnJPN_{t-i} + \sum_{i=1}^n \vartheta_{6i} LnALNG_{t-i} \\
& + \sum_{i=1}^n \vartheta_{7i} LnOIL_{t-i} + u_{5t}
\end{aligned} \tag{9.5}$$

$$\begin{aligned}
LnALNG_t = & \lambda_0 + \sum_{i=1}^n \lambda_{1i} LnHH_{t-i} + \sum_{i=1}^n \lambda_{2i} LnNBP_{t-i} + \sum_{i=1}^n \lambda_{3i} LnRUS_{t-i} \\
& + \sum_{i=1}^n \lambda_{4i} LnTTF_{t-i} + \sum_{i=1}^n \lambda_{5i} LnJPN_{t-i} + \sum_{i=1}^n \lambda_{6i} LnALNG_{t-i} \\
& + \sum_{i=1}^n \lambda_{7i} LnOIL_{t-i} + u_{6t}
\end{aligned} \tag{9.6}$$

$$\begin{aligned}
LnOIL_t = & \mu_0 + \sum_{i=1}^n \mu_{1i} LnHH_{t-i} + \sum_{i=1}^n \mu_{2i} LnNBP_{t-i} + \sum_{i=1}^n \mu_{3i} LnRUS_{t-i} \\
& + \sum_{i=1}^n \mu_{4i} LnTTF_{t-i} + \sum_{i=1}^n \mu_{5i} LnJPN_{t-i} + \sum_{i=1}^n \mu_{6i} LnALNG_{t-i} \\
& + \sum_{i=1}^n \mu_{7i} LnOIL_{t-i} + u_{7t}
\end{aligned} \tag{9.7}$$

Here,  $n$  denotes the number of lags in the VAR model, while  $u_{1t}$  to  $u_{7t}$  represent the normally and independently distributed (NID) residuals for each equation. Initially, we aim to determine the optimal number of lags for the VAR model, ensuring the absence of autocorrelation and the confirmation of normal distribution criteria within the residuals of each equation.

### 9.2.1 Optimal VAR Lag Length

The optimal lag length of the specified VAR model is analysed by utilising the information criterion outlined in sub-section 6.3.1 of Chapter 6. The value of each information criterion (IC) statistic is calculated for each number of lags, and the minimum value will dictate the appropriate number of lags. Table 9.1 presents the results of the IC minimum values according to the calculations of the three main ICs, Akaike's IC, Schwarz's IC, and Hannan-Quinn's IC, in trials allowing for a maximum of 13 lags.

Table 9.1 Optimal Lag Length Selection Using Information Criterion.

Information Criterion	Minimum Value	Number of Lags
Akaike's IC	-25.605	13
Schwarz's IC	-13.780	1
Hannan-Quinn's IC	-18.933	13

Based on Akaike's and Hannan-Quinn's IC values, the optimal lag length for the VAR model is 13 lags. However, this may lead to overfitting due to the large number of coefficients and constraints. Schwarz's IC suggests one lag, but it results in residuals that are not normally distributed. Further analyses will be conducted to determine the optimal lag length. First, the Lagrange Multiplier Test and autocorrelograms will assess autocorrelation in the residuals (Breusch 1978, Godfrey 1978). Second, the Jarque-Bera test will verify normal distribution. Table 9.2 shows the probability results of the Lagrange Multiplier Test for up to 5 lags.

Table 9.2 LM Test Results for Serial Correlation in the Residuals.

Number of Lags	P-value
1	0.0001
2	0.0005
3	0.0383
4	0.4434
5	0.0247

Table 9.2 suggests that a lag length of four is suitable, as it does not reject the null hypothesis of autocorrelation in the residuals. However, when applying the Jarque-Bera normality test (which accounts for the kurtosis and skewness of the distribution) to the VAR with four lags, one residual in the equation with the European gas price as the dependent variable is not normally distributed. This anomaly was attributed to outliers in the Dutch gas price time series, notably during periods of high volatility in 2011 and a spike in 2018 due to freezing weather (Kotek, Tóth and Mezösi 2018). Dummy variables were introduced into the VAR model to address these outlier periods without altering the model's causal relationships. Thus, the appropriate lag length for the proposed VAR model is concluded to be four. Though typical VAR analyses include Granger-causality tests, impulse responses, and forecast error variance decompositions, this chapter omits reporting the estimated VAR regression coefficients.

## 9.2.2 Toda and Yamamoto (1995) Causality Test

After identifying the optimal lag length for the VAR model, the directional causality between the seven variables in Sample 2 is evaluated using Granger causality test on an unrestricted VAR with variables in levels, following the approach laid out by Toda and Yamamoto (1995). This test is preferred over standard causality tests due to its robustness and alignment with the stationarity and cointegration of the data, thus avoiding pre-test bias (Zapata and Rambaldi 1997). The results of the Toda-Yamamoto (T-Y) causality test applied to the VAR model are presented in Table 9.3, where the null hypothesis of no Granger causality is examined.

Table 9.3<sup>21</sup> Toda and Yamamoto Granger Causality Test in VAR model.

Dependent Variable	Granger-causality	$\chi^2$	Probability	Granger-cause (Y/N)
HH	NBP $\nrightarrow$ HH	4.490	0.344	N
	TTF $\nrightarrow$ HH	2.585	0.630	N
	RUS $\nrightarrow$ HH	5.490	0.241	N
	JPN $\nrightarrow$ HH	1.070	0.899	N
	ALNG $\nrightarrow$ HH	3.770	0.4380	N
	OIL $\nrightarrow$ HH	6.277	0.1794	N
NBP	HH $\nrightarrow$ NBP	4.122	0.3897	N
	TTF $\nrightarrow$ NBP	8.806	0.063	N
	RUS $\nrightarrow$ NBP	8.923	0.063	N
	JPN $\nrightarrow$ NBP	4.576	0.334	N
	ALNG $\nrightarrow$ NBP	11.253**	0.024	Y
	OIL $\nrightarrow$ NBP	1.835	0.766	N
TTF	HH $\nrightarrow$ TTF	3.089	0.543	N
	NBP $\nrightarrow$ TTF	11.725**	0.019	Y
	RUS $\nrightarrow$ TTF	19.729***	0.0006	Y
	JPN $\nrightarrow$ TTF	8.092	0.088	N
	ALNG $\nrightarrow$ TTF	7.897	0.095	N
	OIL $\nrightarrow$ TTF	1.713	0.788	N

<sup>21</sup> Table 9.3 continues on the next page.



Table 9.3 Toda and Yamamoto Granger Causality Test in VAR model, Continued.

Dependent Variable	Granger-causality	$\chi^2$	Probability	Granger-cause (Y/N)
RUS	HH $\nrightarrow$ RUS	3.865	0.425	N
	NBP $\nrightarrow$ RUS	49.310***	0.000	Y
	TTF $\nrightarrow$ RUS	4.766	0.312	N
	JPN $\nrightarrow$ RUS	8.746	0.068	N
	ALNG $\nrightarrow$ RUS	14.138***	0.007	Y
	OIL $\nrightarrow$ RUS	1.072	0.899	N
JPN	HH $\nrightarrow$ JPN	0.776	0.942	N
	NBP $\nrightarrow$ JPN	0.944	0.918	N
	TTF $\nrightarrow$ JPN	1.683	0.794	N
	RUS $\nrightarrow$ JPN	6.041	0.196	N
	ALNG $\nrightarrow$ JPN	13.604***	0.009	Y
	OIL $\nrightarrow$ JPN	179.46***	0.000	Y
ALNG	HH $\nrightarrow$ ALNG	0.497	0.974	N
	NBP $\nrightarrow$ ALNG	22.238***	0.000	Y
	TTF $\nrightarrow$ ALNG	5.293	0.259	N
	RUS $\nrightarrow$ ALNG	3.255	0.516	N
	JPN $\nrightarrow$ ALNG	1.478	0.831	N
	OIL $\nrightarrow$ ALNG	3.917	0.417	N
OIL	HH $\nrightarrow$ OIL	2.023	0.732	N
	NBP $\nrightarrow$ OIL	7.079	0.132	N
	TTF $\nrightarrow$ OIL	1.199	0.878	N
	RUS $\nrightarrow$ OIL	6.863	0.143	N
	JPN $\nrightarrow$ OIL	5.770	0.217	N
	ALNG $\nrightarrow$ OIL	1.465	0.833	N

Notes: \*\*\*, \*\* represent statistical significance at the 1% and 5% levels, respectively.

If the modified Wald chi-square ( $\chi^2$ ) shows significance at a 5% level, we can reject the null hypothesis of no Granger causality. Analysing the outcomes in Table 9.3, there's bidirectional causality between NBP and ALNG gas prices, while unidirectional causality runs from NBP to RUS and TTF. Additionally, there's unidirectional causality from RUS to TTF, OIL to JPN, and ALNG to RUS and JPN. According to the T-Y test, no variable demonstrates any causal relationship over the HH or OIL prices.

As a robustness check to validate the results in Table 9.3, we included the GECON indicator as an exogenous variable in the VAR model. The T-Y test results for this adjusted VAR model are presented in Table 9.4.

Table 9.4<sup>22</sup> Toda and Yamamoto Granger Causality Test in VAR model with GECON Indicator as Exogenous Variable.

Dependent Variable	Granger-causality	$\chi^2$	Probability	Granger-cause (Y/N)
HH	NBP $\Rightarrow$ HH	4.057	0.398	N
	TTF $\Rightarrow$ HH	2.448	0.654	N
	RUS $\Rightarrow$ HH	5.368	0.252	N
	JPN $\Rightarrow$ HH	1.093	0.895	N
	ALNG $\Rightarrow$ HH	3.118	0.538	N
	OIL $\Rightarrow$ HH	5.144	0.273	N
NBP	HH $\Rightarrow$ NBP	3.985	0.408	N
	TTF $\Rightarrow$ NBP	8.627	0.071	N
	RUS $\Rightarrow$ NBP	8.765	0.067	N
	JPN $\Rightarrow$ NBP	4.524	0.340	N
	ALNG $\Rightarrow$ NBP	11.148**	0.025	Y
	OIL $\Rightarrow$ NBP	1.803	0.772	N
TTF	HH $\Rightarrow$ TTF	4.235	0.375	N
	NBP $\Rightarrow$ TTF	10.644**	0.031	Y
	RUS $\Rightarrow$ TTF	21.060***	0.0003	Y
	JPN $\Rightarrow$ TTF	9.399	0.0539	N
	ALNG $\Rightarrow$ TTF	9.150	0.058	N
	OIL $\Rightarrow$ TTF	1.874	0.760	N
RUS	HH $\Rightarrow$ RUS	4.339	0.362	N
	NBP $\Rightarrow$ RUS	49.538***	0.000	Y
	TTF $\Rightarrow$ RUS	5.175	0.270	N
	JPN $\Rightarrow$ RUS	8.833	0.066	N
	ALNG $\Rightarrow$ RUS	14.542***	0.006	Y
	OIL $\Rightarrow$ RUS	1.026	0.906	N

<sup>22</sup> Table 9.4 continues on the next page.

Table 9.4 Toda and Yamamoto Granger Causality Test in VAR model with GECON Indicator as Exogenous Variable, Continued.

Dependent Variable	Granger-causality	$\chi^2$	Probability	Granger-cause (Y/N)
JPN	HH $\Rightarrow$ JPN	0.777	0.942	N
	NBP $\Rightarrow$ JPN	0.790	0.940	N
	TTF $\Rightarrow$ JPN	1.687	0.793	N
	RUS $\Rightarrow$ JPN	5.697	0.223	N
	ALNG $\Rightarrow$ JPN	11.893***	0.018	Y
	OIL $\Rightarrow$ JPN	175.29***	0.000	Y
ALNG	HH $\Rightarrow$ ALNG	0.502	0.973	N
	NBP $\Rightarrow$ ALNG	21.268***	0.000	Y
	TTF $\Rightarrow$ ALNG	5.352	0.253	N
	RUS $\Rightarrow$ ALNG	3.165	0.531	N
	JPN $\Rightarrow$ ALNG	1.511	0.825	N
	OIL $\Rightarrow$ ALNG	3.166	0.531	N
OIL	HH $\Rightarrow$ OIL	2.048	0.727	N
	NBP $\Rightarrow$ OIL	5.013	0.286	N
	TTF $\Rightarrow$ OIL	0.816	0.937	N
	RUS $\Rightarrow$ OIL	6.948	0.139	N
	JPN $\Rightarrow$ OIL	6.108	0.191	N
	ALNG $\Rightarrow$ OIL	1.446	0.836	N

Notes: \*\*\*, \*\* represent statistical significance at the 1% and 5% levels, respectively.

As expected from the consistency in the cointegration results in Chapters 7 and 8, the robustness check confirmed all causal directions derived from the VAR T-Y Granger causality test. This outcome ensures the reliability of our results and subsequent interpretations.

Following the diagrams illustrating causality directions in Chapters 7 and 8, Figure 9.2 represents the causality relationships in Sample 2 based on the long-run T-Y causality test results in Tables 9.3 and 9.4.

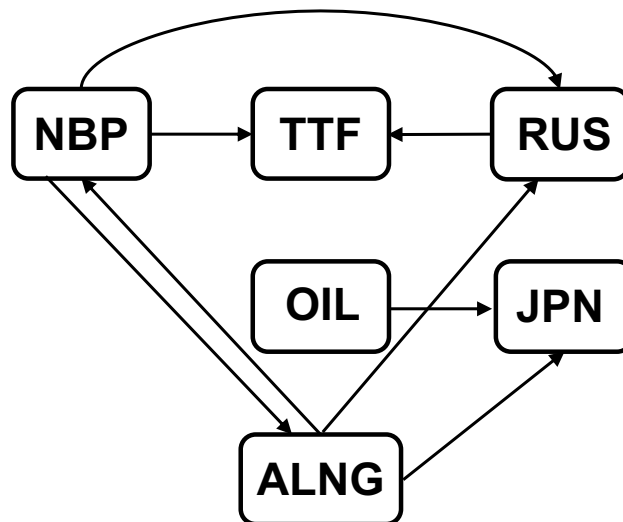


Figure 9.2 Diagram of the Granger Causality Directions Based on the T-Y Causality Test Applies to Sample 2.

There are 8 Granger causality pairs identified through T-Y causality test applied to the VAR model. A significant reduction from the findings of Chapter 7 which identified via linear ARDL bivariate models a total of 27 long-run causality relationships in Sample 2. A detailed discussion of the results will be conducted in the concluding section of this chapter. The next sub-sections will present other two methods that will help on the interpretation of the VAR model. They are the impulse responses and decomposition of forecast error variance.

### 9.2.3 VAR Analysis of Impulse Responses

The impulse responses provide the future observations of each time series in the VAR model in response to a one-unit increase in the current value of one of the VAR residuals. This analysis is feasible because the residuals exhibit no correlation across equations in the VAR. The Cholesky decomposition generates impulse responses in a Vector Autoregression (VAR) model. The technique converts the system of equations into a corresponding set of orthogonal errors by establishing an order among the variables in the VAR model and subsequently decomposing the variance-covariance matrix of the errors via the Cholesky decomposition. A specific ordering of the variables in the VAR model must be selected. Lütkepohl (2005) states that this ordering should reflect the causal relationships among the variables, such that the contemporaneous shocks affect the variables in the order specified by the ordering.

Upon reviewing Table 9.3 and Figure 9.2, we can identify the time series that exerts the most influence on others within the dataset. The Cholesky ordering, designed to prioritize time series based on their responsiveness to causal relationships within the VAR framework, starts with the least responsive and progresses to the most responsive variables. Following Figure 9.2, the proposed order for Cholesky decomposition is  $LnOIL - LnHH - LnALNG - LnNBP - LnRUS - LnTTF - LnJPN$ . Figures 9.3 to 9.9 illustrate the impact of an unexpected one percentage point increase in each residual on all other variables in the VAR system. Dashed lines on the graphs represent  $\pm 2$  standard error bands, establishing 95% confidence intervals.

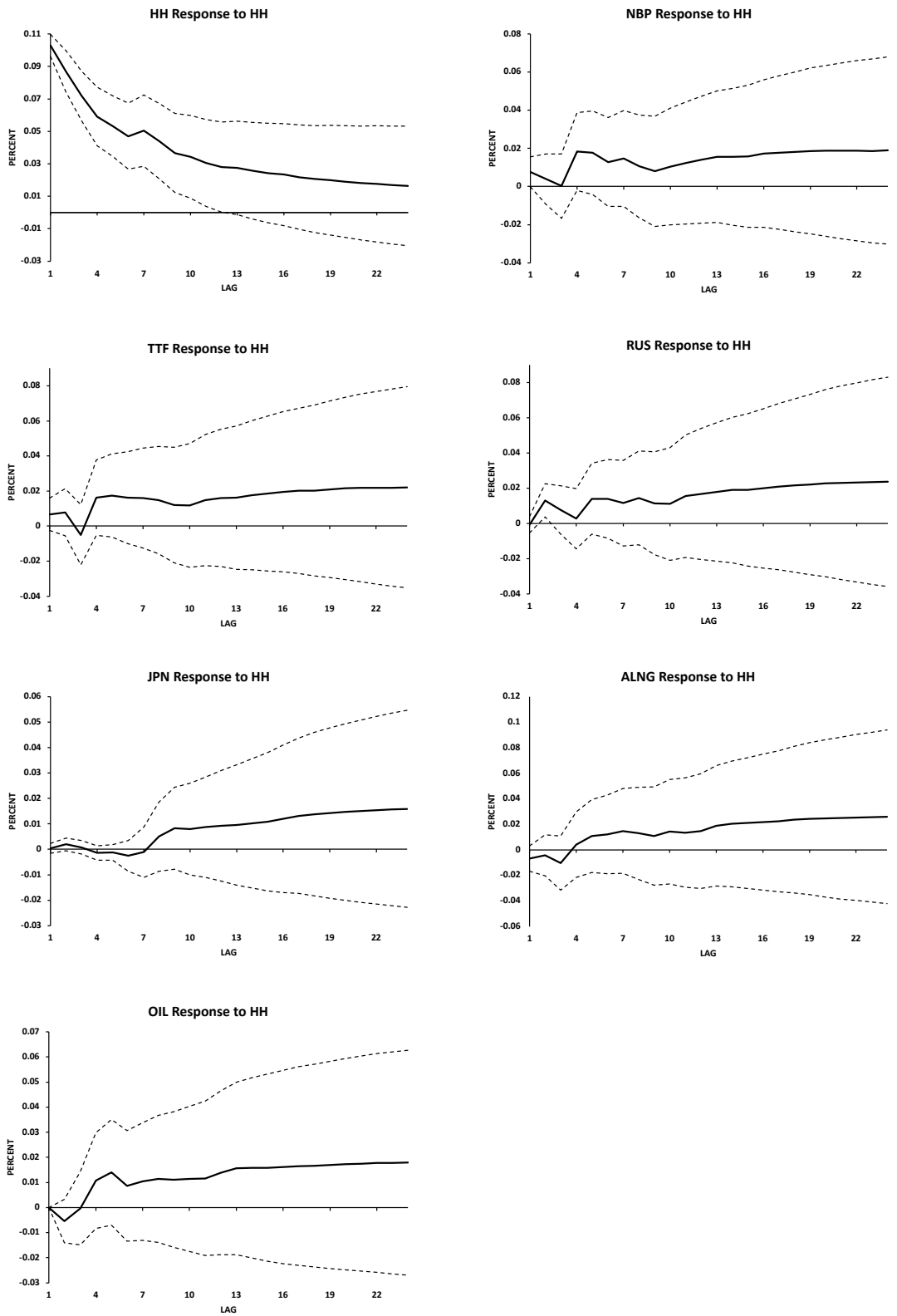


Figure 9.3 Responses to a Henry Hub (HH) gas price shock.

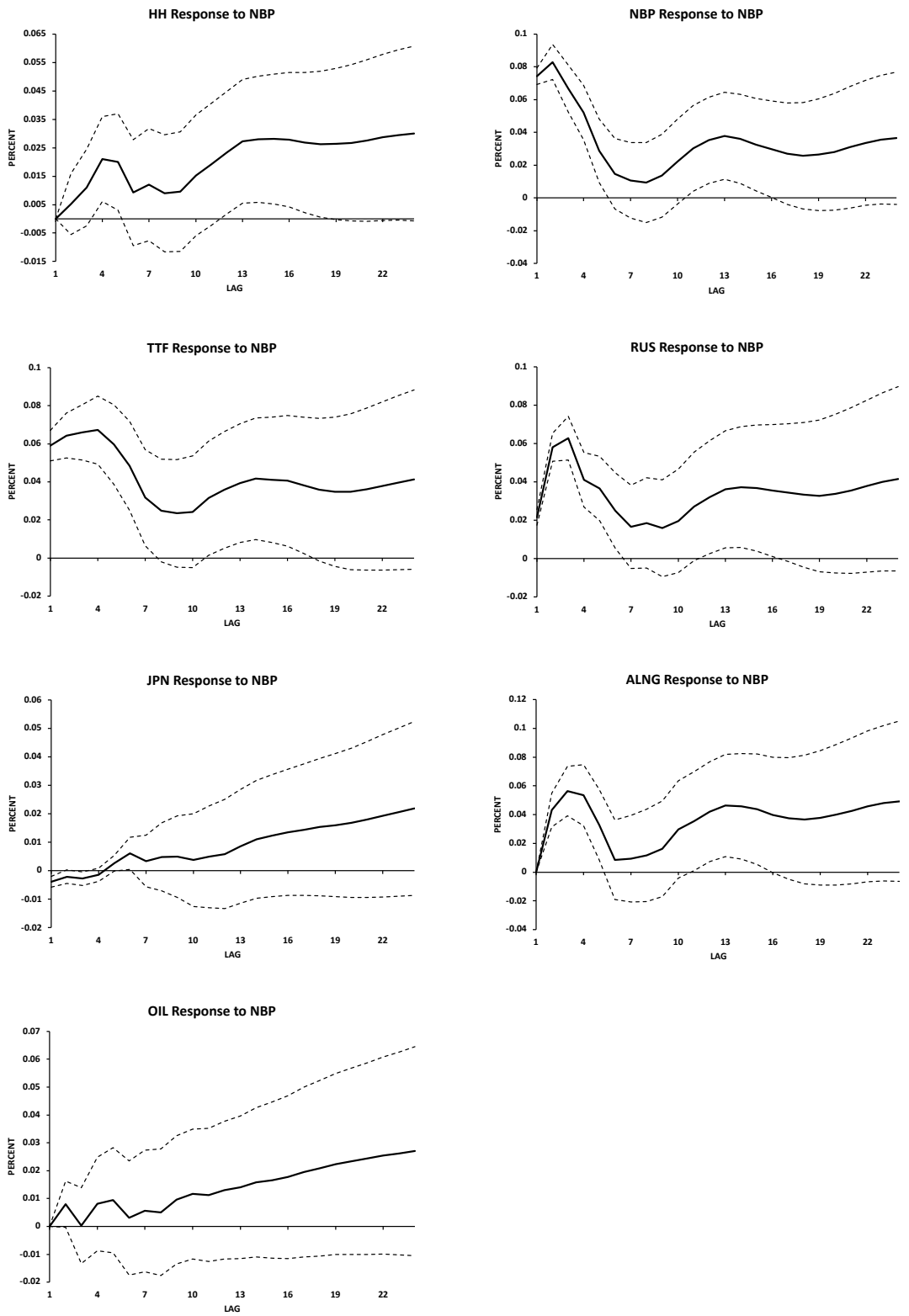


Figure 9.4 Responses to a National Balancing Point (NBP) gas price shock.

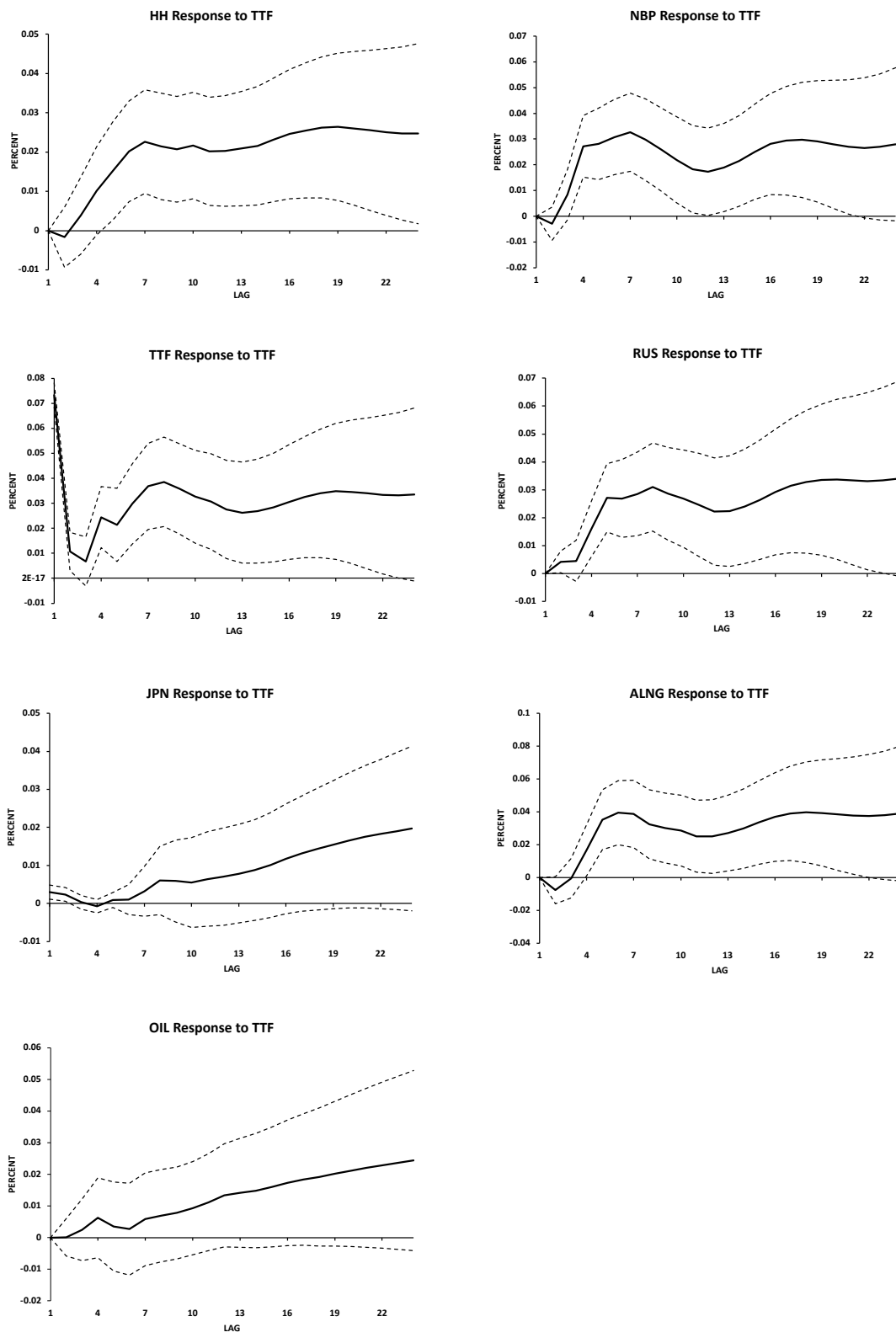


Figure 9.5 Responses to a Title Transfer Facility (TTF) gas price shock.



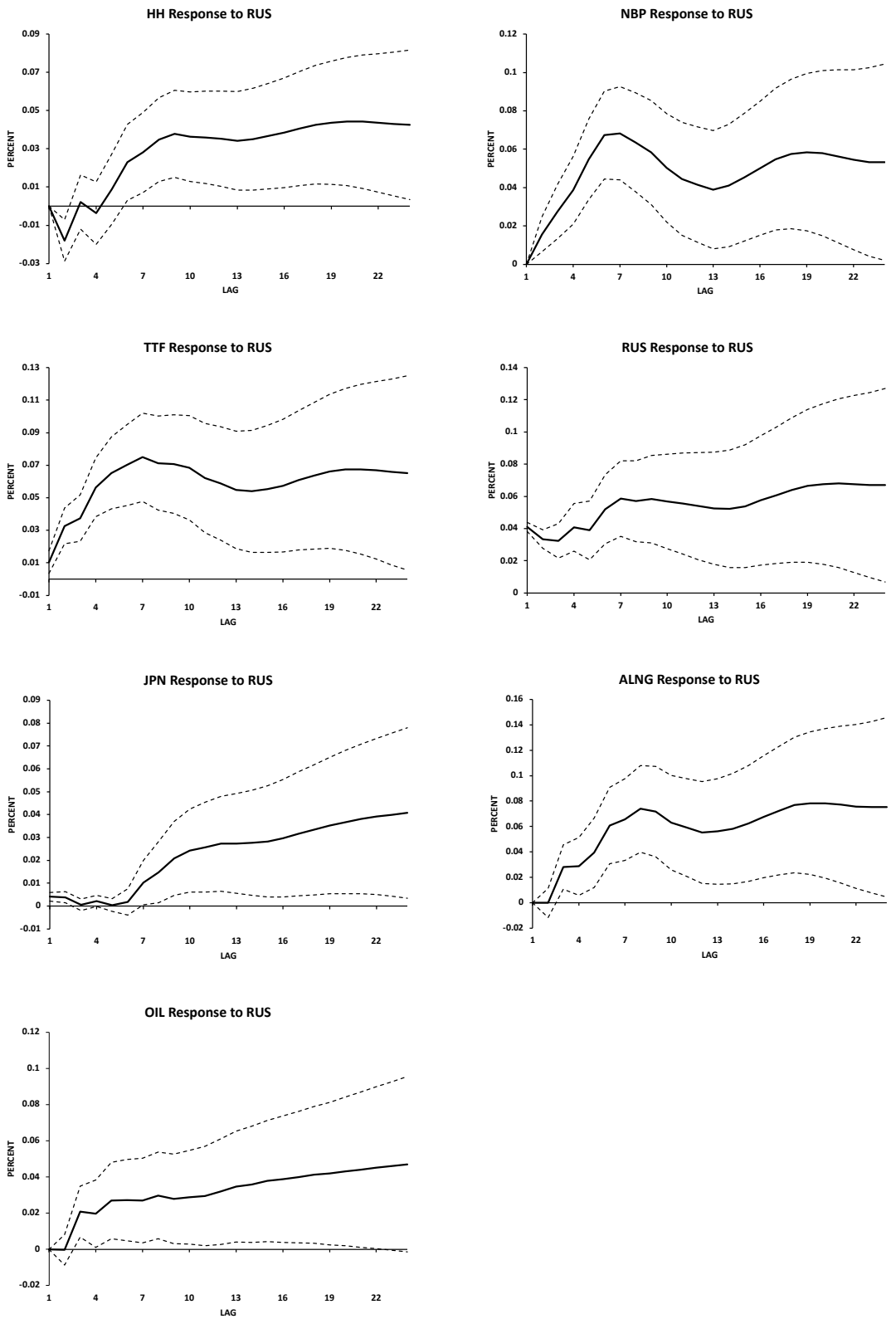


Figure 9.6 Responses to a Russian (RUS) gas price shock.

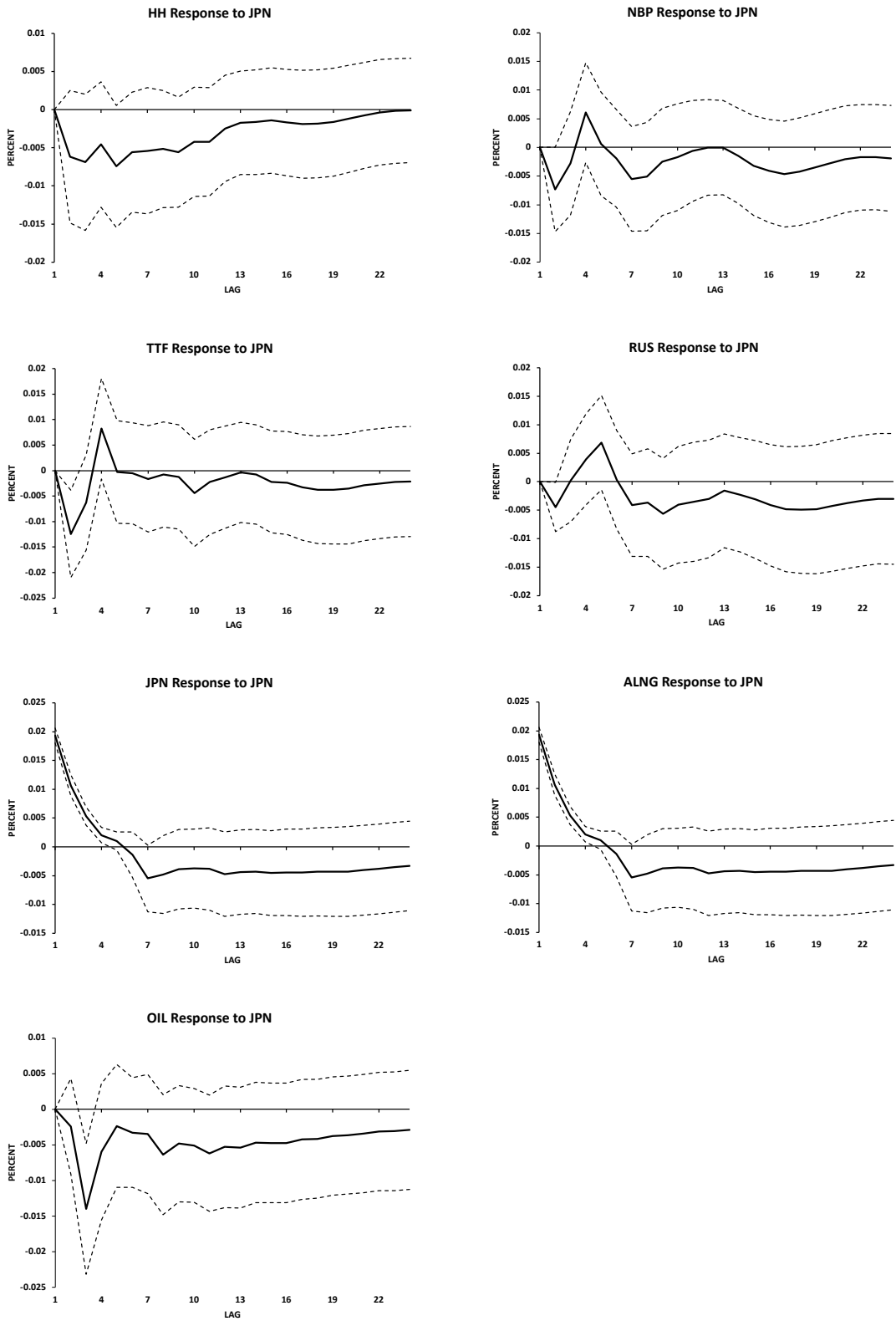


Figure 9.7 Responses to a Japan's Monthly Average LNG Import (JPN) gas price shock.

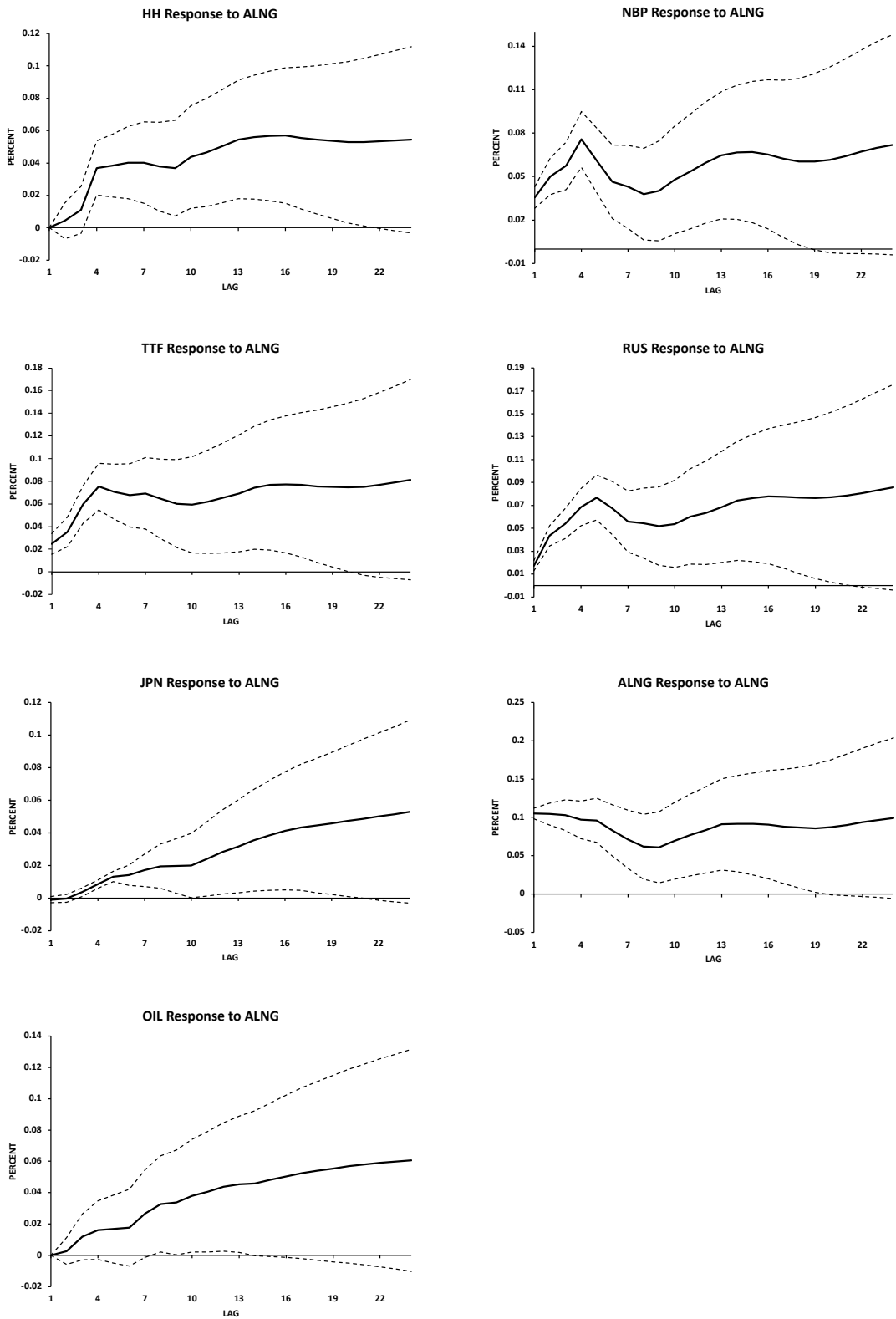


Figure 9.8 Responses to a shock in LNG prices imported into Northeast Asia (ALNG).

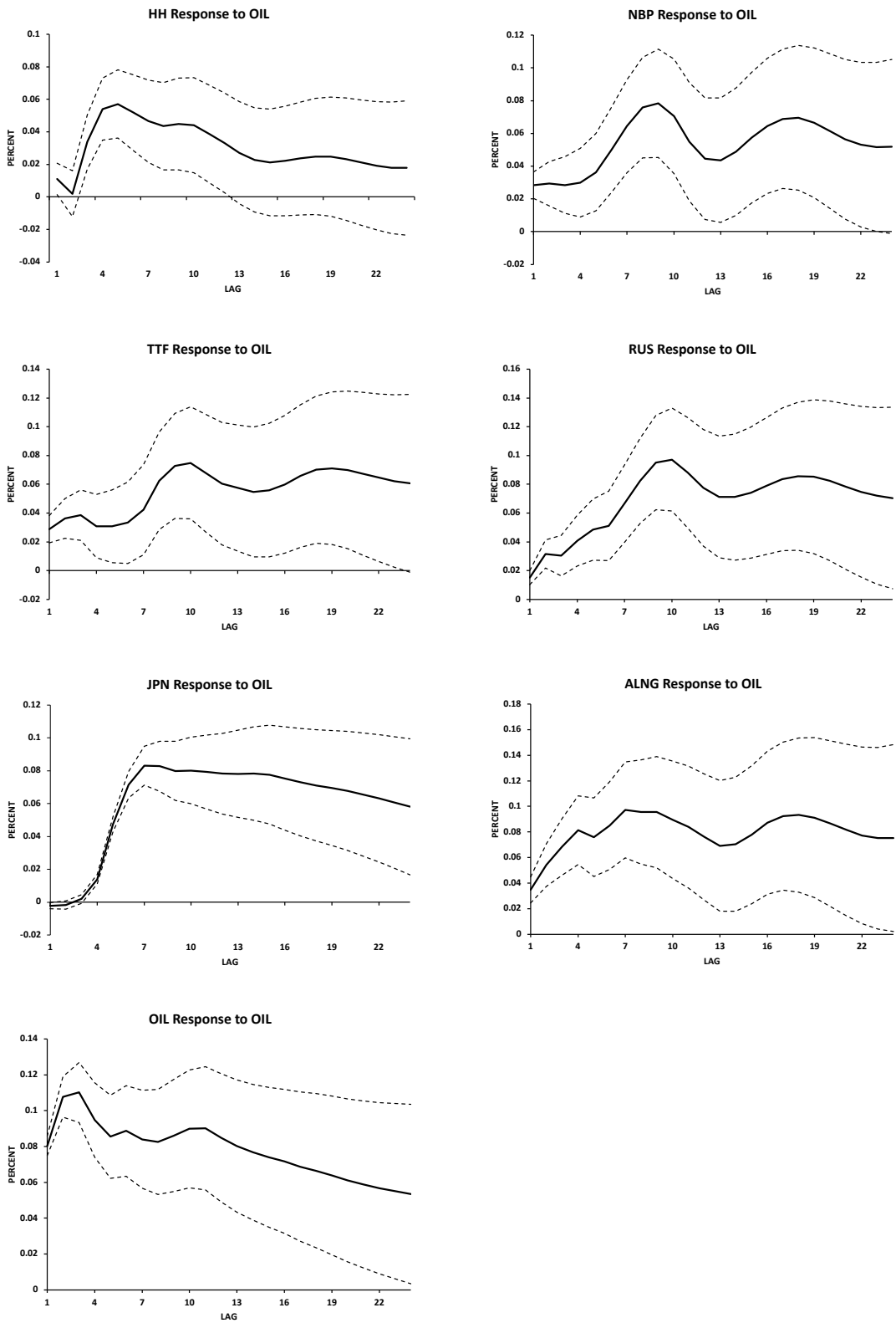


Figure 9.9 Responses to a Brent crude oil (OIL) price shock.

The impulse responses in Figures 9.3 to 9.9 suggest a considerable variation in the reactions to a 1% shock to the variables OIL, ALNG, and RUS. The variable's responses from a shock to NBP and TTF are moderate, while a 1% shock to HH appears to have a small effect of no more than 0.02% over the other variables in 24 months. Furthermore, a one-percent shock to JPN has a negligible effect on the other variables as their variations quickly fade to 0% over 24 months.

Before drawing conclusions based on the above results, the next sub-section will report the results of the forecast error decompositions of each variable in the VAR model. These findings will then be interpreted collectively to ensure a comprehensive understanding.

#### **9.2.4 VAR Analysis of Forecast Error Variance Decomposition**

The forecast error decomposition, or just variance decomposition, shows the percentage of forecast error due to each variable in the VAR model over time in response to a specific shock. Essentially, the variance decomposition shows the extent of variation in the dependent variable, which is explained by its shocks and the shocks of the other variables in the VAR system.

Just as in the assessment of Impulse Responses, Cholesky decomposition is a crucial tool for analysing variance decomposition forecasts within a VAR model. By doing so, Cholesky's decomposition enables the estimation of the variance of shocks to each variable independently, ensuring that each variable is only affected by its shocks and the shocks of variables preceding it in the ordering. This accurately determines each variable's contribution to the forecast error variance (H. Lütkepohl 2005). Therefore, once again, the order proposed for the Cholesky decomposition is as follows: LnOIL - LnHH - LnALNG - LnNBP - LnRUS - LnTTF - LnJPN.

Table 9.5 presents the variance decomposition of each of the seven variables in the VAR system over 24 months.

Table 9.5<sup>23</sup> Forecast Error Variance Decomposition from the VAR Model.

<b>A – Forecast Error Variance Decomposition of LnHH</b>								
Forecast Months	Forecast Standard Error	Variance Decomposition (Percentage)						
		LnHH	LnNBP	LnTTF	LnRUS	LnJPN	LnALNG	LnOIL
1	0.10	98.87	0.00	0.00	0.00	0.00	0.00	1.13
4	0.18	79.63	1.75	0.35	1.01	0.32	4.47	12.46
8	0.25	56.39	2.03	2.69	4.57	0.39	11.87	22.07
12	0.30	44.34	2.78	3.77	8.99	0.35	17.07	22.70
16	0.35	36.18	4.73	4.61	11.28	0.28	23.66	19.27
20	0.38	30.30	5.73	5.55	14.06	0.23	27.01	17.12
24	0.42	26.23	6.74	6.11	16.13	0.20	29.33	15.26
<b>B – Forecast Error Variance Decomposition of LnNBP</b>								
Forecast Months	Forecast Standard Error	Variance Decomposition (Percentage)						
		LnHH	LnNBP	LnTTF	LnRUS	LnJPN	LnALNG	LnOIL
1	0.09	0.77	72.13	0.00	0.00	0.00	16.64	10.46
4	0.20	1.04	49.43	2.06	6.29	0.25	32.40	8.52
8	0.29	1.44	24.64	5.34	22.17	0.19	25.99	20.23
12	0.35	1.39	18.81	5.01	22.57	0.13	25.67	26.41
16	0.41	1.63	16.60	5.00	21.18	0.12	29.11	26.35
20	0.47	1.84	13.93	5.31	21.97	0.11	28.83	28.00
24	0.52	2.01	13.05	5.40	22.20	0.10	30.28	26.96
<b>C - Forecast Error Variance Decomposition of LnTTF</b>								
Forecast Months	Forecast Standard Error	Variance Decomposition (Percentage)						
		LnHH	LnNBP	LnTTF	LnRUS	LnJPN	LnALNG	LnOIL
1	0.10	0.43	33.37	51.48	1.04	0.00	5.79	7.90
4	0.21	0.89	36.92	13.73	12.89	0.58	24.75	10.24
8	0.32	1.39	23.20	9.97	24.73	0.26	28.57	11.89
12	0.40	1.35	16.82	8.81	26.12	0.18	27.48	19.25
16	0.47	1.58	15.36	7.90	24.76	0.14	30.23	20.03
20	0.54	1.79	13.45	7.59	24.51	0.12	30.76	21.79
24	0.60	1.97	12.48	7.35	24.59	0.10	31.49	22.03

<sup>23</sup> Table 9.5 continues on the next page.

Table 9.5 Forecast Error Variance Decomposition from the VAR Model,  
Continued.

<b>D - Forecast Error Variance Decomposition of LnRUS</b>								
Forecast Months	Forecast Standard Error	Variance Decomposition (Percentage)						
		LnHH	LnNBP	LnTTF	LnRUS	LnJPN	LnALNG	LnOIL
1	0.05	0.02	17.35	0.00	62.72	0.00	11.39	8.53
4	0.17	0.81	32.38	1.02	18.89	0.12	33.70	13.08
8	0.28	1.21	15.14	4.45	20.63	0.14	33.23	25.19
12	0.38	1.21	10.06	4.31	20.22	0.13	27.66	36.42
16	0.46	1.52	9.47	4.23	19.52	0.10	29.64	35.52
20	0.54	1.76	8.41	4.56	19.95	0.11	29.66	35.54
24	0.61	1.97	8.23	4.79	20.56	0.10	30.57	33.78
<b>E - Forecast Error Variance Decomposition of LnJPN</b>								
Forecast Months	Forecast Standard Error	Variance Decomposition (Percentage)						
		LnHH	LnNBP	LnTTF	LnRUS	LnJPN	LnALNG	LnOIL
1	0.02	0.02	3.57	2.10	3.96	89.01	0.24	1.09
4	0.03	0.73	3.20	1.67	4.06	57.67	9.73	22.94
8	0.15	0.17	0.46	0.27	1.55	2.45	4.78	90.33
12	0.23	0.62	0.38	0.41	5.17	1.19	6.10	86.13
16	0.30	0.90	0.83	0.68	6.80	0.82	9.93	80.04
20	0.35	1.28	1.39	1.21	8.66	0.65	13.71	73.11
24	0.40	1.59	2.08	1.81	10.63	0.54	17.10	66.26
<b>F - Forecast Error Variance Decomposition of LnALNG</b>								
Forecast Months	Forecast Standard Error	Variance Decomposition (Percentage)						
		LnHH	LnNBP	LnTTF	LnRUS	LnJPN	LnALNG	LnOIL
1	0.11	0.39	0.00	0.00	0.00	0.00	89.77	9.84
4	0.26	0.28	11.74	0.50	2.37	0.14	62.08	22.88
8	0.38	0.57	6.35	3.87	11.33	0.10	45.65	32.11
12	0.47	0.70	6.08	3.90	14.57	0.10	39.88	34.76
16	0.55	1.06	6.92	4.15	15.41	0.08	39.64	32.73
20	0.64	1.36	6.61	4.63	17.35	0.09	37.24	32.73
24	0.71	1.60	7.01	4.86	18.46	0.08	36.99	31.00
<b>G - Forecast Error Variance Decomposition of LnOIL</b>								
Forecast Months	Forecast Standard Error	Variance Decomposition (Percentage)						
		LnHH	LnNBP	LnTTF	LnRUS	LnJPN	LnALNG	LnOIL
1	0.08	0.00	0.00	0.00	0.00	0.00	0.00	100.0
4	0.20	0.35	0.32	0.11	1.99	0.58	0.98	95.68
8	0.28	0.86	0.37	0.19	5.10	0.40	3.63	89.45
12	0.34	1.04	0.68	0.50	6.23	0.36	7.53	83.66
16	0.40	1.41	1.16	0.98	8.04	0.32	11.29	76.78
20	0.45	1.70	1.86	1.56	9.87	0.29	15.00	69.72
24	0.49	1.93	2.64	2.19	11.63	0.26	18.30	63.05

The results reported in Table 9.5 suggest a significant interaction between the variables. More specifically, the variables ALNG and OIL have the most significant spillover effect on the changes in all the other variables at the end of the time horizon. Also, the RUS price is shown to have a considerable impact on the changes in most variables and is the third most impactful variable in the variance decomposition table. The NBP price has a small impact on the forecasting changes in the RUS, TTF, and ALNG prices. The prices that have the least influence over the variability of the other variables are the JPN, HH, and TTF.

### **9.3 Concluding Remarks**

The literature commonly employs bivariate models like the linear and asymmetric ARDL to explore short- and long-term causality and assess cointegration relationships between natural gas markets and oil prices. While useful for examining long-run equilibrium interactions between two variables, bivariate models may not fully capture the complex dynamics of the entire system. To address this limitation, this chapter introduces the VAR model, allowing for the simultaneous modelling of all variables within the system, thereby offering deeper insights into their mutual influences and complementing the findings of bivariate ARDL models.

Employing the T-Y causality test within the VAR framework enhances the robustness of Granger causality detection, a feature absent in traditional linear ARDL bivariate models. Furthermore, following the robustness checks conducted in the bounds cointegration test models in Chapters 7 and 8, we included the GECON indicator as an exogenous variable in the VAR model. We compared the results to the original model based solely on price time series. The robustness check confirmed all results of the T-Y Granger causality test.

In our analysis, the T-Y causality test applied to the VAR model revealed 8 Granger causality pairs, a significant decrease from the 27 long-run causality relationships identified in Sample 2 through linear ARDL bivariate models, as discussed in Chapter 7. It is essential to emphasize that this comparison is not intended to discredit ARDL results. Rather, the VAR model complements the findings by shedding light on spillover effects inherent in bivariate relationships, offering a more comprehensive perspective on the interplay among variables.



The results of the T-Y causality test presented in Table 9.3 reveal that NBP, RUS, and ALNG have a strong causal influence on each other. While NBP and RUS do not exhibit bidirectional causality at a 5% significance level, the p-value of RUS causing NBP (0.063) almost satisfies the maximum significance level. This finding corroborates Chapters 7 and 8, which indicate a correlation between European and Asian markets. TTF is heavily impacted by NBP and RUS gas prices. At the same time, ALNG significantly influences JPN gas prices, indicating strong regional integration between European gas prices and the two Asian gas benchmarks.

However, the causality relationships between the HH gas price and the Brent crude oil price to the other gas prices identified in the ARDL assessment are not detected in the VAR framework. This finding once again present evidence that the relationship between OIL and gas prices has weakened over the last decade. The only exception is the strong causality from OIL to JPN, which results from the high proportion of oil indexation in the imported gas to Japan. As a result, JPN is the only price with a causality dependency on OIL price. The results showed that the HH price representing the U.S. gas market is independent, as there is no causal relationship in any direction between HH and all the other price time series.

After examining the VAR Impulse Responses and Variance Decomposition Forecasts, we found that the OIL time series has a certain impact on gas prices other than HH. Additionally, the analysis indicates that ALNG and RUS significantly influence the other gas prices present in Sample 2. Overall, the findings from Tables 9.3 and 9.4, as well as Figures 9.3 to 9.9, lead us to conclude the following:

- The T-Y causality test within the VAR framework enhances Granger causality detection compared to traditional linear ARDL bivariate models, revealing 8 Granger causality pairs. This not only complements the findings but also highlights the possibility of spillover effects, offering a more comprehensive perspective on the integration of natural gas prices.
- The findings in this chapter echo those in Chapter 7, particularly in the assessment of leading and lagging markets outlined in Section 7.3.5. Russia emerges as the leading market in Europe, while ALNG leads in Asia. Additionally, a correlation between European and Asian gas markets is observed, notably with bidirectional causality between NBP and ALNG, where

ALNG is the leading market. However, HH does not exhibit any strong relationship with any of the time series in the sample.

- The reduction from 27 causality relationships among cointegrated pairs discussed in Chapter 7 to just 8 VAR T-Y causality pairs aligns with explanations in the Introduction and is illustrated in Figure 9.1. The process successfully addressed the causality detection trap by removing previously identified bidirectional causalities between OIL and other gas prices. Only a few causal links were found within JPN, and all involving the HH were discarded. The T-Y findings clarify that OIL impacts only JPN, which ALNG influences, but no other gas prices. Furthermore, the research confirms a causal connection between the three European gas prices and ALNG.
- Analysing the overall results, OIL has notably reduced its influence on gas markets. However, it is important to note that OIL still retains some degree of impact on certain gas prices, as indicated by the Impulse Response and Variance Decomposition Forecasts. This could offer insight into the causality relationships uncovered in the ARDL assessments, which were not corroborated by the VAR T-Y Causality test. Chapter 7 has already examined potential explanations for these findings.

Moving forward, Chapter 10 will explore the factors influencing natural gas demand, specifically focusing on key LNG importer markets such as Japan and Korea. This last methodology will expand the thesis' investigation into the intricate relationship between natural gas demand and its underlying determinants. By utilizing autoregressive distributed lag models, Chapter 10 aims to uncover valuable insights into the elasticity of natural gas demand across industrial and residential sectors. The study will enrich our understanding of price formation dynamics in the global natural gas market.

# **CHAPTER 10**

## **PRICE AND INCOME ELASTICITY OF NATURAL GAS DEMAND: A MULTI-SECTOR ANALYSIS IN JAPAN AND KOREA**

### **10.1 Introduction**

In an era marked by increasing global concerns over energy security, environmental sustainability, and economic growth, the role of natural gas as a pivotal energy source has gained significant attention. As nations strive to diversify their energy mix and transition towards cleaner alternatives, the demand for natural gas has surged, contributing to meeting energy needs and reducing greenhouse gas emissions. Within this landscape, fast-growing gas importers in Asia emerge as vital players. Although the literature has extensively analysed the gas demand factors in China, demand in Japan and South Korea has been overlooked. This chapter fills this gap by estimating long-run elasticities for these countries.

Natural gas, characterised by its relatively lower carbon intensity than other fossil fuels, has become a transitional energy source in the journey towards a more sustainable energy future. As Japan and Korea forge their paths towards economic development and energy security, their reliance on natural gas has become increasingly pronounced. Both nations have strategically positioned themselves to tap into global natural gas supplies, underlining the importance of understanding their demand behaviour in changing market dynamics.

Japan, a nation historically grappling with limited regional energy resources, has turned to natural gas as a vital component of its energy strategy. Following the Fukushima nuclear disaster in 2011, Japan accelerated its transition away from nuclear power, elevating the role of natural gas in its energy portfolio. Similarly, while grappling with energy resource constraints, Korea has placed considerable emphasis on natural gas to support its expanding economy and meet the growing energy demands of its industries.

In this chapter, the elasticity of natural gas demand considering the major importer markets is investigated, and it is found that the price and income elasticities of natural gas demand in the industrial and residential sectors are distinctive. This chapter's methodology constructs autoregressive distribution lag models to study the elasticity of natural gas demand in these two sectors of Japan and Korea. South Korea and Japan were chosen for this thesis's focus for four specific reasons.

First, previous literature has thoroughly investigated the natural gas demand elasticities of significant importer markets like Europe and China (Andersen, Nilsen and Tveteras 2011; Zhang, Ji and Fan 2018; Lin and Li 2020; Erias and Iglesias 2022; Erias and Iglesias 2022; Wang, et al. 2022). However, there are significant knowledge gaps regarding the natural gas demand patterns in Japan and Korea. These gaps emphasise the need for studies on these regions, considering the unique energy market dynamics that influence how natural gas is consumed in these nations.

Second, the significance of natural gas as an energy source holds substantial weight in the neighbouring East Asian countries of South Korea and Japan. Japan and South Korea are the world's second and third-largest liquefied natural gas (LNG) consumers, respectively. The global LNG spot market relies on two key price indicators: the Title Transfer Facility (TTF) and the Japan Korea Maker (JKM) indices. Essentially, these nations wield considerable influence in the LNG market to the extent that an international LNG price index is specifically directed towards South Korea and Japan. While the traditional trend of the JKM index surpassing the TTF index has recently been reversed, maintaining a stable supply of LNG imports remains a longstanding energy concern for these countries (Lee, Kim and Yoo 2023).

Third, South Korea and Japan are chosen for comparison due to their similar natural gas supply conditions. Despite being geographically distinct, both nations heavily rely on imported natural gas, given their limited domestic production. This shared dependence on imported natural gas, primarily sourced from Middle Eastern countries like Qatar and Oman, highlights their common ground. Despite South Korea's peninsula status and geopolitical tensions with North Korea, its natural gas supply dynamics mirror those of island nations due to the absence of piped natural gas. This convergence underscores the importance of examining LNG imports in both countries for a comprehensive understanding of their energy landscapes.

Fourth, South Korea and Japan use natural gas in similar ways. Both countries use about the same amount of natural gas to make electricity and supply gas to cities. Since these countries are situated around the same distance from the equator, they also have similar patterns of using natural gas for heating and cooling (International Energy Agency 2021, International Energy Agency 2020). This means that South Korea and Japan compete to bring LNG from the global LNG market. In short, because South Korea and Japan are alike in how they use natural gas, how they bring it in, and how it is supplied, it makes sense to compare them. Of course, there could be some differences, but these can be seen as reasons for any variations in the comparison results.

By focusing on Japan and Korea, we aim to bridge this gap in the literature and contribute valuable insights into the determinants of natural gas demand. By exploring the price and income elasticities of natural gas demand in the industrial and residential sectors, we seek to uncover the underlying relationships between these key variables and the consumption patterns in these nations. By shedding light on the factors driving natural gas demand in Japan and Korea, we not only enhance our understanding of their energy landscapes but also pave the way for more informed energy policy decisions, sustainable growth strategies, and effective energy transition planning.

The chapter is organised as follows. Section 10.2 presents a brief outline of the natural gas industry in Japan and South Korea. Section 10.3 presents a brief literature review on using the demand function to assess natural gas demand elasticities. Section 10.4 introduces the methodology and data used for this chapter's assessment. Section 10.5 presents specific models and their respective results for the industrial and residential sectors, as well as a comparison discussion of natural gas demand elasticity in both nations and sectors. Finally, Section 10.6 presents the concluding comments and policy implications.

## **10.2 Overview of the Natural Gas Markets in Japan and South Korea**

### **10.2.1 Japan**

The Fukushima nuclear disaster in 2011 prompted Japan to overhaul its energy strategy, leading to the phasing out of nuclear power and a significant shift in its energy mix. Natural gas emerged as a key player, providing a reliable and cleaner alternative to bridge the energy gap left by nuclear shutdowns. Its versatility in electricity generation and compatibility with renewable sources helped ensure a stable power supply and mitigate the risk of energy shortages. This transformation extended beyond energy sources, as Japan underwent comprehensive liberalisation in both the electricity and gas sectors, fostering market convergence and facilitating cross-sector ventures. Notably, Japan's gas market transitioned from a regional monopoly framework to an open competitive landscape, driven by the surge in natural gas demand following the Fukushima disaster, resulting in a staggering 21% increase in total natural gas consumption in 2012 compared to 2010 (International Energy Agency 2021).

According to the IEA Japan Energy Policy Review (2021) report, Japan is taking significant steps to reshape its energy landscape for a more environmentally responsible future. The country's commitment to curbing climate change is evident as it transitions from coal-based power generation to more sustainable alternatives like natural gas cogeneration facilities and fuel cells. Japan is dedicated to a greener path with a bold target of reducing greenhouse gas emissions by 80% by 2050 compared to 2013. As a nation isolated without energy pipelines connecting to other countries, Japan heavily relies on imported liquefied natural gas (LNG) to meet its energy demands. Most (over 95%) of the natural gas used is imported as LNG, given that only about 2% is produced domestically. The imported LNG serves a dual purpose, with around 66% used for power generation and the remainder as city gas. Industries such as steel and chemicals are the primary natural gas consumers for their operations, underlining their importance in the industrial sector. Japan's strategy involves sourcing around one-third of its imported LNG from Australia, while the rest is divided among Malaysia, Qatar, and Russia, showcasing a diverse supply chain approach.

### 10.2.2 South Korea

The natural gas sector in South Korea has experienced significant changes, primarily due to the establishment of KOGAS (Korea Gas Corporation). Formed in 1983, KOGAS was founded to ensure a steady supply of natural gas to meet the nation's energy demands. KOGAS assumed a crucial role as a transmission system operator overseeing natural gas import and wholesale distribution. Given the absence of pipelines connecting South Korea to other nations, KOGAS played a vital part in managing the importation of natural gas, particularly liquefied natural gas (LNG). However, a notable shift occurred in 2005 when the government allowed the direct import of LNG for private consumption. This decision marked a pivotal moment, bringing about changes in the industry's framework and fostering competition and variety in the market. As of 2020, roughly 22% of the nation's total domestic LNG imports were directly procured without involving KOGAS, highlighting the industry's adaptability and readiness to embrace evolving energy consumption trends. While KOGAS remains central to South Korea's energy infrastructure, the coexistence of new LNG import avenues underscores the country's dedication to a resilient and flexible energy strategy (International Energy Agency 2020, U.S. Energy Information Administration 2023).

LNG is significant in South Korea's energy landscape, making up the third-largest portion of its primary energy sources. As of 2021, the country imported around 46 million tons of natural gas, underscoring its reliance on imports, as its self-sufficiency in NG is just 0.6%. Key exporters of LNG to South Korea include the Middle East (especially Qatar), Southeast Asia, Russia, Australia, and the United States (U.S. Energy Information Administration 2023).

The use of natural gas in South Korea is diverse, spanning the industrial, building, and transportation sectors. Industries like petrochemicals, metal assembly, and steel rely on natural gas. In urban areas, natural gas primarily serves as city gas for cooking and heating. Natural gas is also a key fuel for compressed natural gas buses. The demand for natural gas varies with the seasons, peaking in winter due to heating needs. For instance, during the coldest months from November to February, which is winter in South Korea, a significant 45% of the total annual natural gas consumption occurs. KOGAS operates an LNG receiving terminal where the imported LNG is stored.

KOGAS then converts this LNG back into its gaseous form, transported through an extensive network of natural gas pipelines nationwide (International Energy Agency 2020). Notably, 45% of the imported natural gas is used for generating power, reflecting South Korea's efforts to transition away from coal-fired power plants to reduce pollution and greenhouse gas emissions. This shift is expected to increase natural gas consumption for power generation (Ministry of Trade, Industry and Energy 2020).

### **10.3 Literature Review on the Energy Demand Equation**

Much research has been carried out to understand how responsive the demand for natural gas is, mainly looking at it from the perspective of the demand equation. Price elasticity of demand is a concept that deals with how much the demand for a product changes in response to variations in its price. When looking at the short term, the focus is on how quickly demand reacts to price changes within that specific time frame. In contrast, the long-term price elasticity of demand looks at the bigger picture, considering how the overall demand adjusts to a price change over time until a new balance is achieved (Donnelly 1987). Thus, it is important to consider both immediate and gradual effects when studying the impact of price changes on natural gas demand.

Hunt and Manning (1989) and Bentzen and Engsted (1993) investigated how the demand for energy changes in response to various factors. They chose to focus on the UK and Denmark, respectively. They used the error correction model to understand the short-term changes. This method helped them determine how quickly the energy demand adjusts when variables like prices and income change. For a broader and longer view, they used cointegration analysis, which allowed them to see the bigger patterns of how energy demand changes over longer periods. Using these different techniques, they gained a more detailed understanding of how energy demand reacts immediately and gradually when changing variables like prices or other factors.

Contemporary studies exploring how the demand for natural gas behaves often adopt the demand function approach. Asche, Nilsen, and Tveteras (2008), Erdogdu (2010), Dagher (2012), and Yu, Zheng, and Han (2014) have used this approach. These studies consider many factors, like natural gas prices, other energy sources that could be used as substitutes, the climate, and income, among other demand determinants.



There are two main methods when assessing how these factors affect demand. The first is called the partial adjustment model. This model considers gradual demand changes that follow from changes in endogenous demand determinants. The long-term elasticity derived from this model equals short-term elasticity divided by the adjustment rate. Nonetheless, this model's postulation of a uniform responsiveness rate has sparked concerns regarding its applicability across diverse contexts.

On the other hand, the autoregressive distributed lag model (ARDL) introduces an alternative perspective. This model incorporates a multistep view by considering instantaneous and delayed responses. This approach offers an understanding of the intricate dynamics of demand. Notably, it tackles the constraints of the partial adjustment model by directly affording insights into both short-term and long-term elasticities, thus sidestepping the presumptive uniformity in responsiveness.

Both models can be written in their ECM representations, but level-representation equations are used instead, as ECM representations are not always relevant. In addition, these models can be implemented in an instrumental-variable SEQ demand-and-supply system when adequate data, including cost data, is available with enough observations for the identification of such a system.

The ARDL approach has been extensively applied in energy demand estimations when limited data is available. Bentzen and Engsted (2001) assess how the ARDL model and the ECM approach yield comparable outcomes regarding short-term and long-term elasticities and the dynamic adjustment pattern of energy demand. According to Cuddington and Dagher (2015), the ARDL method is a comprehensive version of the ECM and the partial adjustment model. Furthermore, the analysis of income elasticity is a pertinent aspect in this domain. Dagher (2012) demonstrates that attaining long-term equilibrium occurs within a relatively rapid timeframe, approximately 18 months, following changes in income or price. The ARDL model also finds relevant use in examining income and price elasticity within the context of natural gas demand, facilitating the assessment of the enduring relationship between demand and various influencing factors (Farhani, et al. 2014, Furuoka 2016, Zhang, Ji and Fan 2018).

## 10.4 Methodology and Data

### 10.4.1 Model Specification

The ARDL method is used to model natural gas demand in Japan and Korea, offering advantages over the ECM and partial adjustment models, which restrict elasticity analyses. Pesaran and Shin (1999) showed that ARDL reliably estimates coefficients, even when variables differ in their levels of integration  $I(0)$  or  $I(1)$ , making it a robust choice for analysing the responsiveness of natural gas demand.

Examining the relationship between natural gas demand and factors like natural gas prices, alternative energy prices, and income levels across different sectors is the primary objective of the ARDL demand function. We ensure the suitable lag length for these variables through a series of tests, including the F-test, t-test, and the Akaike Information Criterion (AIC), following a similar method described by Bentzen and Engsted (2001). The general model formula can be represented as follows:

$$\begin{aligned} \ln D_t = & \vartheta + \sum_{i=1}^m \vartheta_{0,i} \ln D_{t-i} \\ & + \sum_{j=0}^n \vartheta_{1,j} \ln P_{t-j} \\ & + \sum_{r=1}^p \sum_{s=0}^q \vartheta_{2,r,s} \ln P_{s,t-s} + \sum_{v=0}^u \vartheta_{3,v} \ln I_{t-v} + \varepsilon \end{aligned} \quad (10.1)$$

Where  $\sum_{i=1}^m \vartheta_{0,i} > 0$ ,  $D_t$  refers to the quantity of natural gas demand in year  $t$ ;  $P_{t-j}$  and  $P_{s,t-s}$  refers to the real price of natural gas in year  $t-j$  and the real price of the energy substitute  $r$  in year  $t-s$ , respectively.  $I_{t-v}$  represents the real income level in year  $t-v$ ; and  $\varepsilon$  represents the random error term.

The short-run and long-run self-price elasticity of natural gas demand are represented as Equations (10.2) and (10.3), respectively.

$$\phi_{self}^S = \vartheta_{1,0} \quad (10.2)$$

$$\phi_{self}^L = \frac{\sum_{j=0}^n \vartheta_{1,j}}{1 - \sum_{i=1}^m \vartheta_{0,i}} \quad (10.3)$$

The short-run and long-run cross-price elasticity of natural gas demand for energy substitute (e.g. coal, electricity, crude oil, and LPG) are represented as (10.4) and (10.5), respectively.

$$\phi_{cross}^S = \vartheta_{2,r,0} \quad (10.4)$$

$$\phi_{self}^L = \frac{\sum_{s=0}^q \vartheta_{2,r,s}}{1 - \sum_{i=1}^m \vartheta_{0,i}} \quad (10.5)$$

The short-run and long-run income elasticity of natural gas demand are represented as (10.6) and (10.7), respectively.

$$\phi_{cross}^S = \vartheta_{3,0} \quad (10.6)$$

$$\phi_{self}^L = \frac{\sum_{v=0}^u \vartheta_{3,v}}{1 - \sum_{i=1}^m \vartheta_{0,i}} \quad (10.7)$$

#### 10.4.2 Variables and Data Sources

To determine how changes in income and prices affect the demand for natural gas in specific sectors in Japan and Korea, we collect pertinent data to construct and apply the ARDL model. It is crucial to thoroughly explain each variable in our dataset to ensure a clear understanding of their roles and importance in our analysis. In this section, we provide a detailed description of each variable.

Datasets on natural gas prices, substitute prices, and natural gas demand were sourced from the OECD iLibrary databases. We relied on relevant variables from the World Bank database to assess industrial and residential income levels in Japan and Korea. The time series were obtained at a yearly frequency with a sample period from 2001 to 2021(21 observations). The decision to focus on a sample period from 2001 to 2021, with yearly data points, was mainly influenced by constraints related to data availability. However, it is important to note that the ARDL model can still yield meaningful results even with limited observations. We can find examples in the existing literature, such as Zhang et al. 2018, where researchers have successfully

employed this methodology with a 20-year sample period of yearly observations. Nominal price or value data are adjusted to real prices or values using 2001 as the base year. Furthermore, we applied a logarithmic transformation to all the variables to mitigate heteroscedasticity. The list of variables is as follows:

- **Price of natural gas:** We gather yearly average natural gas prices in Japan ( $P^{Jpn}$ ) and Korea ( $P^{Kor}$ ), expressed in local currency per megawatt-hour (MWh). This data is disaggregated to distinguish between prices applicable to the residential ( $P_{Res}^{Jpn}, P_{Res}^{Kor}$ ) and industrial sectors ( $P_{Ind}^{Jpn}, P_{Ind}^{Kor}$ ).
- **Prices of substitutes:** Prices of alternative energy sources vary across different sectors, so we selected the relevant variables tailored to the specific characteristics of sectoral natural gas demand. According to Liu et al. 2018, it is evident that electricity, liquefied petroleum gas (LPG), and natural gas play pivotal roles in heating and cooking within the residential sector. Therefore, in our analysis of the residential sector, we incorporate the prices of residential electricity ( $E_{Res}^{Jpn}, E_{Res}^{Kor}$ ) and LPG ( $L^{Jpn}, L^{Kor}$ ) as representative alternative energy prices. In contrast, when focusing on the industrial sector in Japan and Korea, natural gas is primarily used for industrial fuels and gas-fired power generation. Therefore, for our industrial sector analysis, we consider coal and electricity, as proposed by Zhang et al. 2018, as alternative energy sources relative to natural gas. To capture this, we include the prices of industrial electricity ( $E_{Ind}^{Jpn}, E_{Ind}^{Kor}$ ) and the coal price ( $C^{Jpn}, C^{Kor}$ ) as indicators of alternative energy prices. All alternative prices time series were obtained in local currency per megawatt-hour (MWh).
- **Income:** Given the geographical concentration of residential natural gas consumption within urban and town areas, we opt to utilise the annual total Gross Domestic Product ( $GDP^{Jpn}, GDP^{Kor}$ ) in the local currency as the income indicator for the residential sector. Furthermore, we designate the Industrial Value Added ( $IVA^{Jpn}, IVA^{Kor}$ ) in the local currency as the income variable representing the industrial sector.

A summary of the variables' description and abbreviation is presented in Table 10.1 as follows:

Table 10.1 Variables' Description and Abbreviation.

Variable description	Abbrev.
Demand for natural gas (MCM logarithm form) in Japan's industrial sector	$LnD_{Ind}^{Jpn}$
Demand for natural gas (MCM logarithm form) in Korea's industrial sector	$LnD_{Ind}^{Kor}$
Demand for natural gas (MCM logarithm form) in Japan's residential sector	$LnD_{Res}^{Jpn}$
Demand for natural gas (MCM logarithm form) in Korea's residential sector	$LnD_{Res}^{Kor}$
Real price (logarithm form) of natural gas in Japan's industrial sector	$LnP_{Ind}^{Jpn}$
Real price (logarithm form) of natural gas in Korea's industrial sector	$LnP_{Ind}^{Kor}$
Real price (logarithm form) of natural gas in Japan's residential sector	$LnP_{Res}^{Jpn}$
Real price (logarithm form) of natural gas in Korea's residential sector	$LnP_{Res}^{Kor}$
Real price (logarithm form) of electricity in Japan's industrial sector	$LnE_{Ind}^{Jpn}$
Real price (logarithm form) of electricity in Korea's industrial sector	$LnE_{Ind}^{Kor}$
Real price (logarithm form) of electricity in Japan's residential sector	$LnE_{Res}^{Jpn}$
Real price (logarithm form) of electricity in Korea's residential sector	$LnE_{Res}^{Kor}$
Real price (logarithm form) of LPG in Japan's residential sector	$LnL^{Jpn}$
Real price (logarithm form) of LPG in Korea's residential sector	$LnL^{Kor}$
Real price (logarithm form) of coal in Japan's industrial sector	$LnC^{Jpn}$
Real price (logarithm form) of coal in Korea's industrial sector	$LnC^{Kor}$
Total yearly Gross Domestic Product (logarithm form) in Japan	$LnGDP^{Jpn}$
Total yearly Gross Domestic Product (logarithm form) in Korea	$LnGDP^{Kor}$
Total yearly Industrial Value Added (logarithm form) in Japan	$LnIVA^{Jpn}$
Total yearly Industrial Value Added (logarithm form) in Korea	$LnIVA^{Kor}$

## 10.5 Empirical Results and Discussion

### 10.5.1 Unit Root Test

Before constructing the ARDL model, it is crucial to evaluate the stationarity of our variables and determine their integration orders. It is important to note that while the ARDL model does not demand uniform integration orders among variables, our calculation of the F-statistic in the boundary test relies on the assumption that each variable adheres to either  $I(0)$  or  $I(1)$ , as suggested by the research of Pesaran, Shin, and Smith (2001). This leads us to a fundamental requirement for applying the ARDL

model: the integration order of our variables must not exceed 1. If it does, the validity of the co-integrated F-test comes into question.

To assess the stationarity of our variables, we turn to the Augmented Dickey-Fuller (ADF) and Philips-Perron (PP) tests. The results of the unit root test for all the variables are displayed in Table 10.2 below. These results confirm that all our variables exhibit stationarity when differenced once, assuring that they conform to either the I(0) or I(1) process. With this verification, we can confidently proceed with applying the ARDL model.

Table 10.2 Results of Unit Root Tests.

LEVELS					FIRST DIFFERENCE				
Variables	ADF (t-Stat.)		PP (Adj. t-Stat.)		Variables	ADF (t-Stat.)		PP (Adj. t-Stat.)	
	Const.	Const. & Trend	Const.	Const. & Trend		Const.	Const. & Trend	Const.	Const. & Trend
$LnD_{Ind}^{Jpn}$	-4.385***	-3.095	-7.092***	-4.240**	$LnD_{Ind}^{Jpn}$	-3.101**	-3.127	-3.716**	-2.756
$LnD_{Ind}^{Kor}$	-1.593	-1.534	-1.873	-1.084	$LnD_{Ind}^{Kor}$	-2.845*	-2.955	-3.035**	-2.788
$LnD_{Res}^{Jpn}$	-2.756*	-3.667**	-2.683*	-3.696**	$LnD_{Res}^{Jpn}$	-6.206***	-6.425***	-9.730***	-11.84***
$LnD_{Res}^{Kor}$	-3.921***	-3.847**	-3.797**	-5.517***	$LnD_{Res}^{Kor}$	-5.912***	-5.608***	-9.377***	-8.722***
$LnP_{Ind}^{Jpn}$	-2.207	-2.577	-1.711	-2.887*	$LnP_{Ind}^{Jpn}$	-3.889***	-3.730**	-2.887*	-2.727
$LnP_{Ind}^{Kor}$	-1.535	-1.190	-1.562	-1.190	$LnP_{Ind}^{Kor}$	-4.105***	-3.940**	-4.105***	-4.425**
$LnP_{Res}^{Jpn}$	-2.348	-2.336	-1.969	-1.896	$LnP_{Res}^{Jpn}$	-4.116***	-3.956**	-3.180**	-3.140
$LnP_{Res}^{Kor}$	-1.042	-0.862	-1.191	-0.862	$LnP_{Res}^{Kor}$	-3.641**	-4.132**	-3.641**	-4.140**
$LnE_{Ind}^{Jpn}$	-1.149	-1.498	-1.248	-1.484	$LnE_{Ind}^{Jpn}$	-3.039**	-2.838	-3.049**	-2.893
$LnE_{Ind}^{Kor}$	-0.789	-2.184	-0.839	-1.624	$LnE_{Ind}^{Kor}$	-3.318**	-3.145	-3.318**	-3.145
$LnE_{Res}^{Jpn}$	-1.030	-2.553	-1.104	-1.624	$LnE_{Res}^{Jpn}$	-3.981***	-3.807**	-4.005**	-3.777**
$LnE_{Res}^{Kor}$	-1.271	-1.759	-1.202	-1.557	$LnE_{Res}^{Kor}$	-3.053**	-3.025	-3.047**	-3.025*
$LnL^{Jpn}$	-1.750	-2.341	-1.553	-2.341	$LnL^{Jpn}$	-5.076***	-4.983***	-5.265***	-5.682***
$LnL^{Kor}$	-2.219	-2.135	-2.219	-2.092	$LnL^{Kor}$	-4.079***	-4.387**	-4.079***	-4.385**
$LnC^{Jpn}$	-3.108**	-2.086	-1.824	-1.950	$LnC^{Jpn}$	-4.451***	-4.856***	-4.453***	-6.628***
$LnC^{Kor}$	-1.134	-1.858	-1.135	-1.858	$LnC^{Kor}$	-4.295***	-4.192***	-4.297***	-4.189**
$LnGDP^{Jpn}$	-1.723	-2.390	-1.698	-2.390	$LnGDP^{Jpn}$	-4.626***	-4.667***	-4.627***	-4.669***
$LnGDP^{Kor}$	-4.119***	-1.228	-10.00***	-2.178	$LnGDP^{Kor}$	-4.645***	-3.775**	-4.635***	-14.43***
$LnIVA^{Jpn}$	-2.577	-2.547	-1.865	-1.746	$LnIVA^{Jpn}$	-4.069***	-3.923**	-4.098***	-3.968**
$LnIVA^{Kor}$	-3.133**	0.207	-4.994***	-1.071	$LnIVA^{Kor}$	-3.308**	-4.219**	-3.212**	-5.369***

Note: \* p < 0.1, \*\* p < 0.5, \*\*\* p < 0.01.

## 10.5.2 Analysis of the Natural Gas Demand in the Industrial Sector: Japan and Korea

Natural gas serves as a high-quality industrial fuel, finding extensive application in sectors like oil refinement, chemical production, metal processing, electrical and

thermal power generation, and oil and gas extraction. Consequently, natural gas can substitute or complement coal, fuel oil, and electricity. In developing an ARDL model to study the demand for natural gas within the industrial sector, our primary emphasis lies in examining the real prices of industrial natural gas, coal, and industrial electrical power. Furthermore, as economic progress advances, it is expected that there will be a heightened demand for energy consumption. Hence, we also consider the influence of the real Industrial Value Added on the natural gas demand.

We verify whether the ARDL model includes the autoregressive term and the lag order through a sequence of evaluations, similar to the approach introduced by Bentzen and Engsted (2001). The process involves performing the F-test, t-test and evaluating the Akaike Information Criterion (AIC). Representing the natural gas demand function for the industrial sectors in Japan and Korea, Equations (10.8) and (10.9) serve as the foundation for this analysis.

$$\begin{aligned}
 LnD_{Ind\ t}^{Jpn} = & \phi_{Ind}^{Jpn} + \alpha_0 LnD_{Ind\ t-1}^{Jpn} + \alpha_1 LnP_{Ind\ t}^{Jpn} + \alpha_2 LnC_{Ind\ t}^{Jpn} \\
 & + \alpha_3 LnC_{Ind\ t-1}^{Jpn} + \alpha_4 LnC_{Ind\ t-2}^{Jpn} + \alpha_5 LnE_{Ind\ t}^{Jpn} \\
 & + \alpha_6 LnIVA_{Ind\ t}^{Jpn} + \alpha_7 LnIVA_{Ind\ t-1}^{Jpn} + \alpha_8 LnIVA_{Ind\ t-2}^{Jpn} \\
 & + \alpha_9 TREND + \varepsilon_{Ind}^{Jpn}
 \end{aligned} \tag{10.8}$$

Based on the significance of the parameters and collinearity issues, the coal price is not included in the demand model for Korea's industrial sector. Consequently, the resulting equation is as follows:

$$\begin{aligned}
 LnD_{Ind\ t}^{Kor} = & \phi_{Ind}^{Kor} + \beta_0 LnD_{Ind\ t-1}^{Kor} + \beta_1 LnP_{Ind\ t}^{Kor} + \beta_2 LnP_{Ind\ t-1}^{Kor} \\
 & + \beta_3 LnE_{Ind\ t}^{Kor} \\
 & + \beta_4 LnE_{Ind\ t-1}^{Kor} + \beta_5 LnE_{Ind\ t-2}^{Kor} + \beta_6 LnIVA_{Ind\ t}^{Kor} \\
 & + \beta_7 LnIVA_{Ind\ t-1}^{Kor} + \beta_8 LnIVA_{Ind\ t-2}^{Kor} + \varepsilon_{Ind}^{Kor}
 \end{aligned} \tag{10.9}$$

The coefficient results for the two ARDL models demonstrated above are reported in Table 10.3. We derive the short-term and long-term elasticities for natural gas demand by substituting these coefficient estimations alongside Equations (10.2) through (10.7).

Table 10.3 Coefficient Estimation of Natural Gas Demand in Japan and Korea:  
Industrial Sector.

<b>Japan – Industrial Sector</b>				
Coefficient	ARDL estimated value	Standard Error	t-Statistic	P-value
$\phi_{Ind}^{Jpn}$	25.805**	8.623	2.993	0.017
$\alpha_0$	0.224	0.186	1.202	0.264
$\alpha_1$	0.079	0.084	0.941	0.374
$\alpha_2$	0.049	0.058	0.844	0.423
$\alpha_3$	0.102	0.071	1.433	0.190
$\alpha_4$	-0.215**	0.070	-3.093	0.015
$\alpha_5$	-0.803***	0.166	-4.849	0.001
$\alpha_6$	0.433**	0.188	2.310	0.049
$\alpha_7$	-0.325	0.252	-1.291	0.233
$\alpha_8$	-0.448**	0.180	-2.488	0.038
$\alpha_9$	0.020***	0.005	3.777	0.005

Adjusted R<sup>2</sup> = 0.86 , F-statistic = 11.99 , AIC = -4.12

<b>Korea – Industrial Sector</b>				
Coefficient	ARDL estimated value	Standard Error	t-Statistic	P-value
$\phi_{Ind}^{Kor}$	-8.850	6.472	-1.368	0.205
$\beta_0$	0.817***	0.228	3.580	0.001
$\beta_1$	-0.245	0.250	-0.979	0.353
$\beta_2$	-0.357**	0.155	-2.305	0.047
$\beta_3$	-0.072	0.328	-0.220	0.831
$\beta_4$	0.214	0.302	0.709	0.496
$\beta_5$	-1.258***	0.336	-3.748	0.005
$\beta_6$	0.751	0.400	1.878	0.093
$\beta_7$	1.104	0.724	1.525	0.162
$\beta_8$	-0.989	0.505	-1.959	0.082

Adjusted R<sup>2</sup> = 0.98 , F-statistic = 91.75 , AIC = -3.08

Notes: \* p < 0.1, \*\* p < 0.5, \*\*\* p < 0.01.

Referring to Equations (10.2) through (10.7) and the data presented in Table 10.3 for the industrial sector, we can determine natural gas demand's short-run and long-run elasticities. We assess the significance of the long-term elasticities using the Wald test for the joint significance of each variable's contemporary and lagged coefficients. This test is a statistical tool used to determine if a combination of independent variables collectively has a 'significant' influence on a model.



In Japan's industrial sector context, the short-run and long-run own-price elasticities regarding the demand for natural gas are 0.079 and 0.101, respectively. These estimates do not exhibit statistical significance at a 10% significance level, which implies that the demand for natural gas remains relatively inelastic concerning changes in the price of natural gas, both in the short term and the long term. The short- and long-run cross-price elasticities of natural gas demand in relation to coal are 0.049 and -0.083, respectively. The short- and long-run cross-price elasticity of natural gas demand concerning electricity are -0.803 and -1.034, respectively. While only the long-run cross-price elasticity for coal is significant at a 10% level, the cross-price elasticities for electricity are statistically significant at a 1% level. The income elasticity of natural gas demand is 0.433 in the short-run, while in the long-run, the elasticity is -0.437. Both income elasticities are significant at a 5% level.

Analysing the results for natural gas demand in the Japanese industrial sector over the past two decades reveals several key findings. Firstly, variations in natural gas prices are inelastic to its demand, mainly because natural gas plays a vital role in Japan's energy mix following the 2011 nuclear disaster. Its demand is a critical buffer against energy shortages, regardless of price fluctuations. Moreover, it is noteworthy that coal prices exhibit an inelastic effect on natural gas demand, reflecting a broader trend of reducing the use of more polluting fuels and promoting cleaner alternatives like natural gas.

Secondly, the findings reveal a significant negative correlation when examining the electricity price elasticity of natural gas demand, suggesting that electricity and natural gas function more as complements than substitutes in the industrial sector. This outcome may be due to the flexibility of using electricity as an alternative to coal-based power generation despite natural gas's higher fuel quality.

Finally, regarding income elasticity, short-term growth in Industrial Value Added (*IVA*) positively stimulates natural gas demand. However, long-term income elasticity turns negative, reflecting the Japanese government's commitment to reducing greenhouse gas emissions by incorporating renewable energy sources into the energy mix in the coming years.

Considering the industrial sector in Korea, the short- and long-run own-price elasticities for the demand for natural gas are -0.245 and -3.298, respectively. However, only the long-run own-price elasticity is significant at a 10% level. The short-run cross-price elasticity of natural gas demand in regard to electricity is -0.072 and not significant. On the other hand, the long-run elasticity is -6.115 and statistically significant at a 1% level. The short- and long-run income elasticities are 0.751 and 4.750, respectively. While the long-run income elasticity is significant at 1%, the short-run is significant at 10%.

The rising price of natural gas has a negative effect on its demand within the Korean industrial sector, particularly over the long term. In this sector, a complementary relationship exists between natural gas and electricity consumption. As the industrial value-added (*IVA*) in Korea continues to grow, the capacity for industrial power generation has expanded. Consequently, this growth has increased the demand for natural gas across various sectors, resulting in a relevant and positive income elasticity of natural gas demand.

### **10.5.3 Analysis of the Natural gas Demand in Residential Sector: Japan and Korea**

We incorporate the real price of natural gas consumption in the residential sector, the real prices of residential liquefied petroleum gas and electricity, and the real total GDP as an income variable of urban residents into formulating the natural gas demand equation. Cross-price variables for residential liquefied petroleum gas (LPG) and electricity in the assessment of natural gas demand are warranted because these energy sources can act as substitutes for natural gas in household applications. Moreover, since LPG and electricity are commonly used for heating and cooking, their inclusion allows for a more comprehensive understanding of how price shifts affect residential natural gas demand, shedding light on consumer behaviour and energy market dynamics. Additionally, incorporating the total Gross Domestic Product (GDP) variable as an income assessment is justified because it reflects urban residents' economic well-being and can influence household energy consumption patterns. Therefore, Equations (10.10) and (10.11) represent the natural gas demand function of the residential sector in Japan and Korea, respectively.

Based on the significance of the parameters and concerns related to collinearity, the LPG price is not included in the demand model for Japan's residential sector. Consequently, the resulting equations are as follows:

$$\begin{aligned} \text{Ln}D_{\text{Res } t}^{\text{Jpn}} = & \phi_{\text{Res}}^{\text{Jpn}} + \gamma_0 \text{Ln}D_{\text{Res } t-1}^{\text{Jpn}} + \gamma_1 \text{Ln}P_{\text{Res } t}^{\text{Jpn}} + \gamma_2 \text{Ln}E_{\text{Res } t}^{\text{Jpn}} \\ & + \gamma_3 \text{Ln}E_{\text{Res } t-1}^{\text{Jpn}} + \gamma_4 \text{LnGDP}_{\text{Res } t}^{\text{Jpn}} + \varepsilon_{\text{Res}}^{\text{Jpn}} \end{aligned} \quad (10.10)$$

$$\begin{aligned} \text{Ln}D_{\text{Res } t}^{\text{Kor}} = & \phi_{\text{Res}}^{\text{Kor}} + \delta_0 \text{Ln}D_{\text{Res } t-1}^{\text{Kor}} + \delta_1 \text{Ln}D_{\text{Res } t-2}^{\text{Kor}} + \delta_2 \text{Ln}P_{\text{Res } t}^{\text{Kor}} \\ & + \delta_3 \text{Ln}P_{\text{Res } t-1}^{\text{Kor}} + \delta_4 \text{Ln}P_{\text{Res } t-2}^{\text{Kor}} + \delta_5 \text{Ln}E_{\text{Res } t}^{\text{Kor}} \\ & + \delta_6 \text{Ln}L_{\text{Res } t}^{\text{Kor}} + \delta_7 \text{Ln}L_{\text{Res } t-1}^{\text{Kor}} \\ & + \delta_8 \text{LnGDP}_{\text{Res } t}^{\text{Kor}} + \delta_9 \text{LnGDP}_{\text{Res } t-1}^{\text{Kor}} + \delta_{10} \text{LnGDP}_{\text{Res } t-2}^{\text{Kor}} \\ & + \varepsilon_{\text{Res}}^{\text{Kor}} \end{aligned} \quad (10.11)$$

The coefficient results for the ARDL demand Equations (10.10) and (10.11) are presented in Table 10.4.

Table 10.4<sup>24</sup> Coefficient Estimation of Natural Gas Demand in Japan and Korea: Residential Sector.

Japan – Residential Sector				
Coefficient	ARDL estimated value	Standard Error	t-Statistic	P-value
$\phi_{\text{Res}}^{\text{Jpn}}$	21.609***	5.798	3.727	0.002
$\gamma_0$	0.086	0.197	0.439	0.667
$\gamma_1$	-0.213**	0.072	-2.937	0.011
$\gamma_2$	0.190	0.141	1.348	0.199
$\gamma_3$	-0.251	0.138	-1.815	0.091
$\gamma_4$	-0.312	0.162	-1.917	0.076

Adjusted R<sup>2</sup> = 0.75 , F-statistic = 5.17 , AIC = -4.88

<sup>24</sup> Table 10.4 continues on the next page.

Table 10.4 Coefficient Estimation of Natural Gas Demand in Japan and Korea:  
Residential Sector, Continued.

<b>Korea – Residential Sector</b>				
Coefficient	ARDL estimated value	Standard Error	t-Statistic	P-value
$\phi_{Res}^{Kor}$	-14.861**	4.434	-3.352	0.012
$\delta_0$	-0.720	0.308	-2.335	0.052
$\delta_1$	-1.024***	0.268	-3.825	0.007
$\delta_2$	0.012	0.150	0.080	0.939
$\delta_3$	-0.104	0.160	-0.648	0.538
$\delta_4$	-0.724***	0.203	-3.565	0.009
$\delta_5$	0.729**	0.230	3.165	0.016
$\delta_6$	0.137	0.089	1.543	0.167
$\delta_7$	0.135	0.117	1.155	0.268
$\delta_8$	0.061	0.534	1.115	0.912
$\delta_9$	-0.172	0.555	-0.309	0.766
$\delta_{10}$	1.222***	0.558	2.190	0.065

Adjusted  $R^2 = 0.84$ , F-statistic = 9.43, AIC = -4.16

Notes: \*  $p < 0.1$ , \*\*  $p < 0.5$ , \*\*\*  $p < 0.01$ .

When examining Japan's residential sector, we find that short- and long-run own-price elasticities concerning natural gas demand are statistically significant at the 1% level, with values of -0.213 and -0.233, respectively. In contrast, the short- and long-run cross-price elasticities of natural gas demand concerning residential electricity, at 0.199 and -0.068, respectively, are not statistically significant, indicating an inelastic relationship with natural gas demand. Lastly, the income elasticity of natural gas demand, both in the short and long run, stands at -0.312 and -0.341, respectively, which are significant at a 10% level.

Within the Japanese residential sector, natural gas holds a different position in the energy mix than in the industrial sector. Consequently, an increase in its price typically results in a negative response in demand, following the usual economic pattern. Notably, variations in the price of residential electricity, which serves as a competing energy source, have no influence (inelastic) on the demand for natural gas. As was observed in the industrial sector, income elasticity in the residential sector is negative. This trend can also be attributed to the transition toward renewable energy, which gains increasing feasibility as the population's income (GDP) rises.

Analysing natural gas demand in Korea's residential sector, both short- and long-run own-price elasticities for natural gas demand are estimated at 0.012 and -0.297, respectively, with only the long-run elasticity proving statistically significant at the 5% level, suggesting a more substantial impact over time. Conversely, cross-price elasticities related to LPG and residential electricity in both the short and long run fail to reach statistical significance, indicating an inelastic relationship with natural gas demand. Additionally, income elasticities stand at 0.061 for the short-run and 0.405 for the long-run, with the latter being statistically significant at the 1% level, while the former does not achieve statistical significance.

In the Korean residential market, the elasticity of natural gas demand exhibits significance only in the long term, with a negative own-price elasticity, aligning with typical economic patterns, much like in Japan. There is no influence (demand is inelastic) of substitute prices, such as electricity and LPG, on natural gas demand. In contrast to Japan, Korea demonstrates a positive income elasticity in the long run, which implies that as the population's GDP increases, there is a corresponding rise in natural gas demand within the residential sector.

## **10.6 Concluding Remarks**

This chapter has explored the dynamics of natural gas demand within Japan and Korea's industrial and residential sectors, deploying an ARDL approach to capture demand elasticities over short- and long-term horizons. The analysis deepens our understanding of the economic dynamics affecting natural gas demand, highlighting the need for evidence-based, strategic energy policies in both countries.

In Japan, the demand for natural gas in the industrial sector is inelastic to price changes, underscoring the indispensable role of natural gas in ensuring energy security, especially post-2011 nuclear disaster. The transition away from coal and the significant negative elasticity with respect to electricity prices illustrate a shift towards cleaner energy sources, with natural gas and electricity emerging as complements rather than substitutes. This scenario highlights the strategic importance of maintaining a stable and diversified energy supply, promoting energy efficiency, and supporting the integration of renewable energy sources.

The residential sector reflects a negative response to own-price changes, a distinct position within the energy mix compared to the industrial sector. Consequently, an increase in its price generally leads to a decrease in demand, consistent with standard economic behaviour. The negative income elasticity reveals a shift towards renewable energies as income levels rise. This shift indicates a broader commitment to sustainable energy consumption and the need for policies that support renewable energy adoption and the development of energy infrastructures that can accommodate these changes.

Conversely, in Korea, the industrial sector exhibits a significant negative own-price elasticity, suggesting that demand reduces as natural gas prices increase. This finding implies a need for energy diversification and the development of more resilient energy infrastructures that can adapt to fluctuating energy markets. Moreover, the positive income elasticity indicates a growing demand for natural gas with economic growth, emphasizing the importance of enhancing energy efficiency and investing in alternative energy sources to meet this increasing demand without exacerbating environmental impacts.

The residential sector in Korea, like Japan, shows a negative own-price elasticity, reflecting a significant response to price changes over time. However, the positive income elasticity highlights the rising demand for natural gas as household income increases, underscoring the need for policies that encourage energy efficiency and the adoption of renewable energy as an alternative source that households could access.

The findings from both countries underscore the need for targeted energy policies that address the specific dynamics of natural gas demand within different sectors. For Japan, the emphasis should be on further integrating renewable energy sources and strengthening the energy infrastructure to support a gradual transition away from fossil fuels. For Korea, there is a clear imperative to diversify energy sources, promote renewable energy sources, and prepare for the increasing demand for natural gas in the face of economic growth, with a keen focus on environmental sustainability. Both countries' experiences highlight the importance of strategic energy planning incorporating the evolving landscape of global energy markets, technological advancements in renewable energy, and the socioeconomic factors influencing energy demand. As Japan and Korea navigate their future energy pathways, the insights from

this chapter can assist the development of flexible and sustainable energy policies that meet the current demand and anticipate and adapt to future energy needs.

# CHAPTER 11

## CONCLUSIONS AND POLICY IMPLICATIONS

### 11.1 Introduction

This thesis explores the integration of global gas markets and their relationship with crude oil prices, considering major natural gas benchmarks from North America, Europe, and Asia and the Brent crude oil price. It employs different analytical methods, including the convergence test, linear and asymmetrical bounds cointegration tests, and VAR analysis. It also touches upon the demand drivers for natural gas in Japan and South Korea, acknowledging their significant roles as major importers in the global natural gas and LNG markets. Amidst the expanding trade and consumption of natural gas, the findings from this research make a meaningful contribution to the ongoing academic and industry discussion concerning the integration of global gas prices and the strategic direction forward.

### 11.2 Discussion of the Results and Policy Implications

For the global gas market integration analysis, this thesis uses a dataset spanning the years 2001 to 2020, deliberately omitting the period affected by the COVID-19 pandemic. This exclusion is strategic, recognizing that the pandemic-induced demand and supply shocks did not impact all regions uniformly, with effects considered to be transient rather than enduring over the long term. Additionally, there was a notable surge in the total interconnectedness of energy commodity markets following the onset of the pandemic, significantly enhancing the spillover effects of other energy commodities on the natural gas market (Lin and Su 2021).

The primary findings of our convergence analysis suggest a lack of evidence indicating integration among transoceanic gas markets, and a diminished impact of crude oil prices on gas prices in the most recent sample period.

The extent of LNG trading is unlikely to match that of crude oil due to the significantly higher proportion of transportation costs in LNG final prices. While LNG shipping contributes to a growing share of global gas consumption, it remains relatively small,



with only about a third of total LNG shipped available to spot and short-term markets. Therefore, rather than envisioning a singular global natural gas market, it is more accurate to consider distinct regional markets such as East Asia, Europe, and North America. In these markets, gas prices are constrained by an upper limit set by the oil market and a lower limit established by potential arbitrage opportunities between the different regions.

Furthermore, in our initial convergence assessment of the latest sample period, it is observed that regional gas markets exhibit divergence amongst them, except for the three European prices. However, there is a noticeable convergence trend between the Asian spot LNG price and European prices after 2015. The North American Henry Hub price appears isolated and does not converge with any other price in the model. Additionally, the oil price loses its convergence pattern with all gas prices except for the Japanese average import gas price.

The cointegration and causality assessments reveal a growing integration among global natural gas markets, with reduced reliance on oil prices over time. European and Asian markets are increasingly interconnected, with Russia playing a dominant role in pricing dynamics while the US transitions into a net exporter. Although fewer causality pairs are identified comparing the cointegrated bivariate models and the VAR model, the findings corroborate the diminishing influence of oil prices and the increased level of integration between European and Asian spot gas prices. They also highlight the independent nature of the North American gas market.

The US natural gas market is predominantly influenced by domestic elements, with its gas prices showing little sensitivity to fluctuations in crude oil prices. Additionally, it underscores the North American gas market's disconnection from other regional markets, such as those in Europe and Asia, emphasizing its distinct market dynamics. This observation aligns with the research findings of Aruga (2016), Zhang et al. (2018), Zhang and Ji (2018), and Szafranek and Rubaszek (2023).

The distinct dynamics of the US natural gas market can be traced back to two key developments. Firstly, the market became the first to be fully deregulated after the Natural Gas Policy Act 1978, leading to market forces solely determining natural gas prices from the mid-1990s (Joslow 2013). Secondly, the shale gas revolution in the

2010s emerged as a pivotal factor in decoupling US natural gas prices from oil prices and prices in other regional markets. Corbeau and Ledesma (2016) attribute this price divergence to abundant, low-cost supply, high Brent oil prices (above USD 100/barrel), the European recession, and the Fukushima crisis in Japan.

The surge in U.S. natural gas production, primarily from the shale gas revolution, transformed the country from a net importer to a net exporter of natural gas. By February 2016, LNG exports commenced, escalating to 8 billion cubic feet (Bcf) per day by 2020 and reaching 10-11 Bcf/day by 2022. This shift led U.S. exporters to adopt a pricing strategy tied to the Henry Hub, reflecting the opportunity cost of overseas sales. U.S. LNG export agreements favour a Henry Hub-linked pricing model over traditional oil-indexed formulas, allowing buyers to adjust to global spot market prices (U.S. Department of Energy 2022). Nonetheless, the "shale gas boom" also saturated the domestic natural gas market due to inadequate liquefaction facilities, constraining export volumes. Consequently, North American natural gas prices have not yet converged with those in key LNG-importing regions such as Europe and Asia, a finding supported by this thesis and corroborated by existing literature.

North American LNG export capacities are rapidly expanding, with 2023 marking its ascent as the world's largest LNG provider, exceeding Qatar and levelling with Australia. A new wave of LNG export infrastructure is expected to come online around 2025. According to Fulwood (2023), an estimated addition of up to 100 million tonnes per annum (MTPA) could augment the existing global capacity of 470 MTPA between 2025 and 2027. Consequently, the projected surge in U.S. LNG exports to European and Asian markets in the coming years is expected to gradually shift natural gas markets from regionally segmented to a more integrated global market framework. This shift will likely prompt other players in the market to increase the share of flexible pricing structures and shorter-term delivery terms in their conventional long-term LNG contracts to sustain market share, further fostering global gas market integration.

Another important finding revealed throughout this thesis is that despite this apparent disconnection with the North American market, the European and Asian spot gas markets are becoming increasingly correlated, and their dependence on oil prices has diminished, with the exemption of the Japanese average import gas price which is still highly dependent on the oil-indexation price mechanism.

These results can be attributed to the fact that Europe's gas market has undergone a liberalisation process over the last two decades. By 2020, Europe saw a major shift from oil-linked to hub-based natural gas pricing, decreasing oil indexation from 91% to 19% since 2005, as suppliers, including Russian exporter Gazprom, adopted gas hub pricing, which rose from 7% to 81% (International Gas Union 2021). On the Asian side, Japan, heavily reliant on LNG imports and historically tied to JCC-based contracts, led the efforts for contract flexibility post-Fukushima, highlighted by the Fair Trade Commission's scrutiny in 2017, leading to the removal of destination restrictions and the establishment of the Japan OTC exchange, reflecting a broader shift in Asia away from JCC pricing and towards longer, more adaptable LNG contracts (Carriere 2018). Also, the Platts Japan Korea Marker (JKM), which is highly correlated with the Asian LNG price used in this thesis (ALNG), has recently emerged as a significant independent benchmark price for LNG. Despite a notable increase in spot LNG transactions in the broader Asia-Pacific region since 2017, only a quarter of total LNG imports in 2020 were conducted through spot or short-term contracts (International Gas Union 2021).

Amidst increasing flexibility in destination restrictions and pricing mechanisms in LNG contracts and rising trading volumes, a competitive landscape is emerging between European and Asian importers for LNG exports, potentially bolstering convergence and causality in these regional spot natural gas prices. For instance, Qatar's role as a swing supplier, catering to long-term contracts in Asia and short-term and spot contracts to Asia and Europe based on price dynamics, reinforces pricing linkages between these regions (Kim, et al. 2020). Therefore, it is imperative to assess supply security in these LNG importer markets within broader market contexts, prompting the need for bilateral policies to manage demand and supply shocks effectively. Policy implications may involve enhanced collaboration between regions to share crucial information on LNG trade flows, production levels, and demand forecasts, facilitating informed decision-making and mitigating market uncertainties.

The decreasing influence of oil prices on natural gas prices in Europe and Asia, as highlighted in this thesis, resonates with recent research. In a study by Hupka et al. (2023), an in-depth literature review reveals a shift in the understanding of the oil-gas relationship. While earlier studies indicated a strong cointegrated relationship between

natural gas and crude oil, subsequent research using updated data and refined methodologies suggests this relationship has diminished following structural breaks. Employing Engle-Granger cointegration with structural breaks, the authors identify a pivotal break in August 2008, mirroring the primary structural break examined in this thesis, and conclude a strong cointegrated relationship between natural gas and oil prices before the break, transitioning to no relationship thereafter.

The Asian LNG and European gas hub prices exhibit a notable shift following the 2022 European supply shock, providing further evidence of this trend. As illustrated in Figure 11.1, sourced from Stern (2023), the surge in European LNG prices was followed by the Asian LNG price JKM, underscoring their divergence from the Japan oil-indexed gas price.

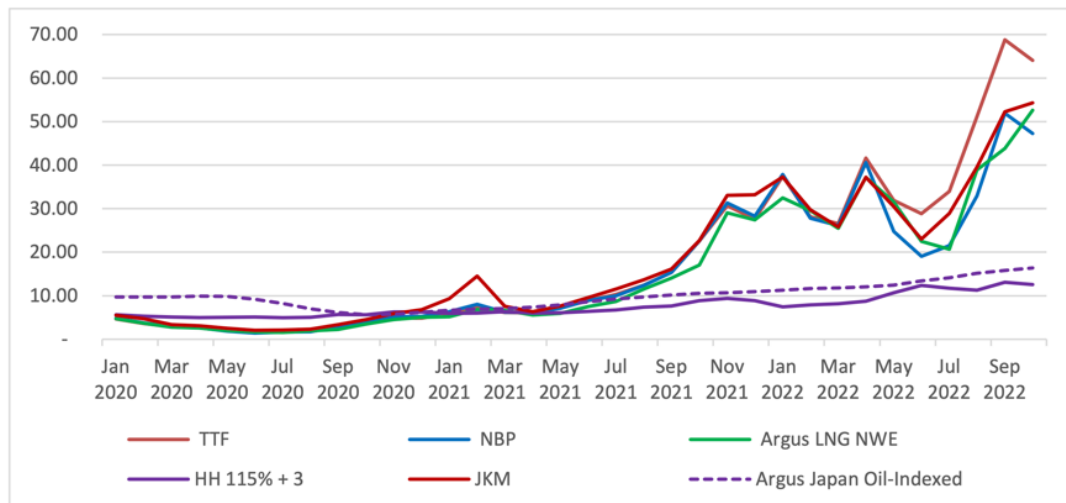


Figure 11.1 European Gas Hub and Asian LNG Prices 2020-22 (\$/MMBtu). Source: Stern (2023)

Nevertheless, it remains notable that oil prices continue to significantly influence Asian and European gas prices. Despite the diminishing impact of oil on spot prices in Asia and Europe, a considerable portion of natural gas transactions still adhere to long-term contracts indexed to oil prices.

Long-term contracts with oil-indexation remain advantageous for natural gas buyers with consistent gas demand, such as base load electricity facilities and industrial processes requiring clean fuel. With the recent developments, these contracts typically offer flexibility, allowing buyers to adjust volumes, typically around 10 %. Importers

can strategically respond to fluctuations in spot LNG prices by adjusting contract volumes, leveraging LNG storage to mitigate demand shocks and prevent prolonged periods of elevated prices compared to oil-indexed rates. Additionally, during periods of declining LNG prices, importers can decrease purchases under long-term contracts and compensate by increasing spot market purchases, ensuring cost-effectiveness and flexibility in procurement strategies. As such, policymakers may encourage importers to adopt strategies that optimise the benefits of long-term contracts while remaining responsive to market dynamics, promoting stability and efficiency in LNG procurement.

In contrast, importer markets, particularly in Asia, must carefully consider the benchmark for indexation in their long-term contracts. According to a recent study by Zhang et al. (2023), Asian gas purchasing contracts indexed to the JKM (Asian spot LNG benchmark) are more adept at risk sharing than oil-indexed contracts, despite potential liquidity constraints. Policymakers should thus advocate for diversifying natural gas suppliers and incorporating spot LNG benchmarks into purchasing agreements to enhance risk-sharing dynamics between importers and exporters.

The evolving dynamics of the global gas market are particularly pertinent to Australia, given its significant role as the world's second-largest LNG exporter, which necessitates strategic adjustments in export regulations and negotiation tactics. With Australia's predominant LNG exports directed towards the Asian market, policymakers should adapt to the emergence of LNG spot benchmarks and anticipate potential changes in regional natural gas price dynamics, fostering resilience and competitiveness in the evolving market landscape.

Another significant discovery from this research is the strong integration of European spot gas prices, where Russian export gas prices largely lead the pricing dynamics. This correlation is unsurprising given that Eurostat data indicates the EU's substantial reliance on gas imports from Russia, reaching nearly 40% in 2020. Furthermore, reinforcing this observation, the notable surge in European natural gas prices from the latter half of 2021 to August 2022, attributed to reduced gas exports from Russia following the Russia-Ukraine war, substantially disrupted the European economy.

This underscores the vital role of LNG as a flexible alternative for Europe's gas supply amid its heavy reliance on Russian pipeline imports. The EU actively pursues diversification measures to address this dependency, including constructing new LNG terminals and improving connections between existing regasification facilities in Spain and Central Europe (European Commission 2022). By late 2022, Europe's increased LNG imports caused a global supply shortage, which can only be alleviated by major new LNG projects, primarily from Qatar and the US, expected to come online in 2026-27 (Stern 2023). Expanding US LNG exports and upgrading to European LNG import infrastructure could improve the connection between these geographically separated gas markets and lessen the Russian impact on Europe's gas market. Therefore, European policymakers should prioritise diversifying their gas supply, considering Europe's demonstrated low gas price elasticity of natural supply, anticipating a shift in gas price dynamics as reliance on Russia decreases as LNG imports grow.

The dynamics of natural gas demand in Japan and Korea's industrial and residential sectors indicate that Japan's industrial sector exhibits inelasticity in gas demand response to price changes, contrasting with the negative response observed in the residential sector. Similarly, demand for gas in Korea's industrial sector demonstrates a negative own-price elasticity, while both industrial and residential sectors show positive income elasticity.

As prices in Japan's industrial sector do not influence the demand, the Japanese economy could suffer from rising natural gas import prices. Japan may consider further regulatory measures to ensure fair competition and transparency in LNG contracts, including continued oversight by regulatory bodies to prevent anti-competitive practices and promote market efficiency. Additionally, policymakers should promote market diversification by exploring new trade agreements, partnerships, and infrastructure investments to reduce dependence on specific LNG suppliers and enhance energy security while supporting the region's establishment and growth of LNG trading hubs. This can be achieved by providing regulatory support, infrastructure investment, and market incentives to facilitate transparent price discovery and enhance market efficiency.

In Korea, the rise in gas prices negatively affects natural gas demand within the industrial sector, yet demand increases with rising incomes in both industrial and residential areas. To address this, the Korean government should prioritise energy diversification efforts to reduce dependence on natural gas and fortify the nation's energy infrastructure. This strategy could involve advocating for advancing and utilising renewable energy sources, allocating resources to develop energy storage technologies, and investigating alternative fuel alternatives. Implementation may include offering financial incentives, simplifying regulatory procedures, and fostering research and development in renewable energy technologies.

In conclusion, the projected increase in global gas consumption by ten per cent by 2030, estimated by the Oxford Institute of Energy Studies, particularly concentrated in Asia and the Middle East, underscores the necessity of strategic energy planning. While there is momentum towards renewable energy sources, their widespread adoption will take time, and natural gas remains a crucial transitional fuel, especially in regions heavily reliant on coal, like South and East Asia. Efforts to restrict natural gas production and exports overlook the long-term benefits, as natural gas can serve as a cleaner alternative to coal in electricity generation. Encouraging the substitution of natural gas for coal and stringent carbon emission controls in the natural gas supply chain can help manage the transition towards a more sustainable energy future. Additionally, LNG is pivotal in providing affordable gas to regions lacking access to pipeline infrastructure, offering a viable alternative to coal while renewable capacities ramp up.

## REFERENCES

- Aguilera, Roberto F., Julian Inchauspe, and Ronald D. Ripple. "The Asia Pacific Natural Gas Market: Large Enough for All?" *Energy Policy* 65 (2014): 1-6. <https://doi.org/10.1016/j.enpol.2013.10.014>.
- Alim, Abdullahi, Peter R. Hartley, and Yihui Lan. "Asian Spot Prices for LNG and other Energy Commodities." *The Energy Journal* 39, no. 1 (2018): 123-142. <https://doi.org/10.5547/01956574.39.1.aali>.
- Alquist, Ron, Saroj Bhattarai, and Olivier Coibion. "Commodity-Price Comovement and Global Economic Activity." *Journal of Monetary Economics* 112 (2020): 41-56. <https://doi.org/10.1016/j.jmoneco.2019.02.004>.
- Alquist, Ron, Lutz Kilian, and Robert Vigfusson. "Forecasting the Price of Oil." In: G. Elliott and A. Timmermann (eds.), *Handbook of Economic Forecasting*, 2A, Amsterdam: North-Holland, 2013.
- Alsamara, Mouyad, Zouhair Mrabet, Karim Barkat, and Mohamed Elafif. "The Impact of Trade and Financial Developments on Economics Growth in Turkey: ARDL Approach with Structural Break." *Emerging Markets Finance and Trade* 55, no. 8 (2019): 1671-1680. <https://doi.org/10.1080/1540496X.2018.1521800>.
- Andersen, Trude Berg, Odd B. Nilsen, and Ragnar Tveteras. "How is Demand for Natural Gas Determined Across European Industrial Sectors?" *Energy Policy* 39, no. 9 (2011): 5499-5508. <https://doi.org/10.1016/j.enpol.2011.05.012>.
- Aruga, Kentaka. "The U.S. Shale Gas Revolution and its Effect on International Gas Markets." *Journal of Unconventional Oil and Gas Resources* 14 (2016): 1-5. <https://doi.org/10.1016/j.juogr.2015.11.002>.
- Asche, Frank, Odd B. Nilsen, and Ragnar Tveteras. "Natural Gas Demand in the European Household Sector." *The Energy Journal* 29, no. 3 (2008): 27-46. <https://doi.org/10.5547/ISSN0195-6574-EJ-Vol29-No3-2>.
- Baumeister, Christiane, Dimitris Korobilis, and Thomas K. Lee. "Energy Markets and Global Economics Conditions." *The Review of Economics and Statistics* 104, no. 4 (2022): 828-844. [https://doi.org/10.1162/rest\\_a\\_00977](https://doi.org/10.1162/rest_a_00977).
- Baumeister, Christiane, and Lutz Kilian. "Real-Time Forecast of the Real Price of Oil." *Journal of Business & Economic Statistics* 30, no. 2 (2012): 326-336. <https://doi.org/10.1080/07350015.2011.648859>.
- Baumeister, Christiane, and James D. Hamilton. "Structural Interpretation of Vector Autoregressions with Incomplete Identification: Revisiting the Role of Oil Supply and Demand Shocks" *American Economic Review* 109, no. 5 (2019): 1873-1910. <https://doi.org/10.1257/aer.20151569>.



- Ben-David, Dan, and David H. Papell. "Slowdowns and Meltdowns: Postwar Growth Evidence from 74 Countries." *The Review of Economics and Statistics* 80, no. 4 (1998): 561-571. <https://doi.org/10.1162/003465398557834>.
- Bennet, Gordon. *LNG Trading, Liquidity and Hedging: A New Landscape for Natural Gas Benchmarks*. The Oxford Institute for Energy Studies, 2019. <https://www.oxfordenergy.org/wpcms/wp-content/uploads/2019/09/OEF-119.pdf>.
- Bentzen, Jan, and Tom Engsted. "A Revival of the Autoregressive Distributed Lag Model in Estimating Energy Demand Relationships." *Energy* 26, no. 1 (2001): 45-55. [https://doi.org/10.1016/S0360-5442\(00\)00052-9](https://doi.org/10.1016/S0360-5442(00)00052-9).
- Bentzen, Jan, and Tom Engsted. "Short- and Long-Run Elasticities in Energy Demand: A Cointegration Approach." *Energy Economics* 15, no. 1 (1993): 9-16. [https://doi.org/10.1016/0140-9883\(93\)90037-R](https://doi.org/10.1016/0140-9883(93)90037-R).
- Box, George E. P., and Gwilym M. Jenkins. *Time Series Analysis, Forecasting and Control*. San Francisco: Holden Day, 1970.
- BP. *BP Statistical Review of World Energy: June 2010*. BP, 2010. <http://www.antjeschupp.de/files/bpstatisticreview.pdf>.
- BP. *BP Statistical Review of World Energy: June 2018*. BP, 2018. <https://www.bp.com/content/dam/bp/business-sites/en/global/corporate/pdfs/energy-economics/statistical-review/bp-stats-review-2018-full-report.pdf>.
- BP. *Statistical Review of World Energy: 2020 69<sup>th</sup> edition*. BP, 2020. <https://www.bp.com/content/dam/bp/business-sites/en/global/corporate/pdfs/energy-economics/statistical-review/bp-stats-review-2020-full-report.pdf>.
- BP. *BP Statistical Review of World Energy: 2022 71<sup>st</sup> edition*. BP, 2022. <https://www.bp.com/content/dam/bp/business-sites/en/global/corporate/pdfs/energy-economics/statistical-review/bp-stats-review-2022-full-report.pdf>.
- Brandt, Patrick T., and John T. Williams. *Multiple Time Series Models*. Thousand Oaks: SAGE Publications, 2007.
- Breusch, T. S. "Testing for Autocorrelation in Dynamic Linear Models\*." *Australian Economic Papers* 17, no. 31 (1978): 334-355. <https://doi.org/10.1111/j.1467-8454.1978.tb00635.x>.
- Broadstock, David C., Raymond Li, and Linjin Wang. "Integration Reforms in the European Natural Gas Market: A Rolling-Window Spillover Analysis." *Energy Economics* 92 (2020): 104939. <https://doi.org/10.1016/j.eneco.2020.104939>.

- Brooks, Chris. *Introductory Econometrics for Finance*. Cambridge: Cambridge University Press, 2008.
- Campbell, John Y., and Pierre Perron. "Pitfalls and Opportunities: What Macroeconomists Should Know about Unit Roots." *NBER Macroeconomics Annual* 6 (1991): 141-201. <https://doi.org/10.2307/3585053>.
- Caporin, Massimiliano, and Fulvio Fontini. "The Long-Run Oil-Natural Gas Price Relationship and the Shale Gas Revolution." *Energy Economics* 64 (2017): 511-519. <https://doi.org/10.1016/j.eneco.2016.07.024>.
- Carriere, Chantal. "The Effects of Japan's Push for Greater LNG Market Flexibility on LNG Pricing and Destination Restrictions." *The Journal of World Energy Law & Business* 11, no. 2 (2018): 136-144. <https://doi.org/10.1093/jwelb/jwy002>.
- Chai, Jian, Zhaohao Wei, Yi Hu, Siping Su, and Zhe G. Zhang. "Is China's Natural Gas Market Globally Connected?" *Energy Policy* 132 (2019): 940-949. <https://doi.org/10.1016/j.enpol.2019.06.042>.
- Chiappini, Raphaël, Yves Jégourel, and Paul Raymond. "Towards a Worldwide Integrated Market? New Evidence on the Dynamics of U.S., European, and Asian Natural Gas Prices." *Energy Economics* 81 (2019): 545-565. <https://doi.org/10.1016/j.eneco.2019.04.020>.
- Christiano, Lawrence J. "Searching for a Break in GNP." *Journal of Business and Economics Statistics* 10, no. 3 (1992): 237-250. <https://doi.org/10.2307/1391540>.
- Corbeau, Anne-Sophie, and David Ladesma. *LNG Markets in Transition: The Great Reconfiguration*. Oxford: University of Oxford, 2016.
- Center for Strategic & International Studies. *When Natural Gas Prices Converge*. Center for Strategic & International Studies, 2020. <https://www.csis.org/analysis/when-natural-gas-prices-converge>.
- Cuddington, John T., and Leila Dagher. "Estimating Short and Long-Run Demand Elasticities: A Primer with Energy-Sector Applications." *The Energy Journal* 36, no. 1 (2015): 185-209. <https://doi.org/10.5547/01956574.36.1.7>.
- Dagher, Leila. "Natural Gas Demand at the Utility Level: An Application of Dynamics Elasticities." *Energy Economics* 34, no. 4 (2012): 961-969. <https://doi.org/10.1016/j.eneco.2011.05.010>.
- Dickey, David A., and Wayne A. Fuller. "Distribution of the Estimators for Autoregressive Time Series with a Unit Root." *Journal of the American Statistical Association* 74, no. 366 (1979): 427-431. <https://doi.org/10.2307/2286348>.

- Doane, Michael J., and Daniel F. Spulber. "Open Access and the Evolution of the U.S. Spot Market for Natural Gas." *The Journal of Law & Economics* 37, no. 2 (1994): 477-517. <https://doi.org/10.1086/467321>.
- Dong, Kangyin, Renjin Sun, and Gal Hochman. "No Natural Gas and Renewable Energy Consumption Lead to Less CO<sub>2</sub> Emission? Empirical Evidence from a Panel of BRICS Countries." *Energy* 141 (2017): 1466-1478. <https://doi.org/10.1016/j.energy.2017.11.092>.
- Donnelly, William A. *The Econometrics of Energy Demand: A Survey of Applications*. New York: Praeger Publishers, 1987.
- Dukhanina, Ekaterina, Olivier Massol, and François Lévêque. "Policy Measures Targeting a More Integrated Gas Market: Impact of a Merger of Two Trading Zones on Prices and Arbitrage Activity in France." *Energy Policy* 132 (2019): 583-593. <https://doi.org/10.1016/j.enpol.2019.05.044>.
- Energy Working Group. *Investments in Natural Gas Supply Chain under the Low Price Environment*. Energy Working Group, 2018. [https://aperc.or.jp/file/2018/7/2/Aperc\\_8834\\_Natural\\_Gas\\_web.pdf.pdf](https://aperc.or.jp/file/2018/7/2/Aperc_8834_Natural_Gas_web.pdf.pdf).
- Erdős, Péter. "Have Oil and Gas Prices got Separated?" *Energy Policy* 49 (2012): 707-718. <https://doi.org/10.1016/j.enpol.2012.07.022>.
- Erdogdu, Erkan. "Natural Gas Demand in Turkey." *Applied Energy* 81, no. 1 (2010): 211-219. <https://doi.org/10.1016/j.apenergy.2009.07.006>.
- Erias, Antonio F., and Emma F. Iglesias. "Price and Income Elasticity of Natural Gas Demand in Europe and the Effects of Lockdowns due to COVID-19." *Energy Strategy Reviews* 44 (2022): 100945. <https://doi.org/10.1016/j.esr.2022.100945>.
- Erias, Antonio F., and Emma F. Iglesias. "The Daily Price and Income Elasticity of Natural Gas Demand in Europe." *Energy Reports* 8 (2022): 14595-14605. <https://doi.org/10.1016/j.egy.2022.10.404>.
- European Commission. *Communication REPowerEU Plan COM(2022)230 (europa.eu)*. European Commission, 2022. [https://commission.europa.eu/publications/key-documents-repowerEU\\_en](https://commission.europa.eu/publications/key-documents-repowerEU_en).
- Ernst & Young. *Global LNG: Will New Demand and New Supply Mean New Pricing?* Ernst & Young, 2013. [https://www.ourenergypolicy.org/wp-content/uploads/2013/03/Global\\_LNG\\_New\\_pricing\\_ahead\\_DW0240.pdf](https://www.ourenergypolicy.org/wp-content/uploads/2013/03/Global_LNG_New_pricing_ahead_DW0240.pdf).
- Falk, Barry. "Further Evidence on the Asymmetric Behavior of Economic Time Series Over the Business Cycle." *Journal of Political Economy* 94, no. 5 (1986): 1096-1109. <https://doi.org/10.1086/261423>.
- Farhani, Sahbi, Muhammad Shahbaz, Mohamed Arouri, and Frédéric Teulon. "The Role of Natural Gas Consumption and Trade in Tunisia's Output." *Energy Policy* 66 (2014): 677-684. <https://doi.org/10.1016/j.enpol.2013.11.040>.

- Foster, Andrew J. "Price Discovery in Oil Markets: A Time Varying Analysis of the 1990-1991 Gulf Conflict." *Energy Economics* 18, no. 3 (1996): 231-246. [https://doi.org/10.1016/0140-9883\(96\)00020-5](https://doi.org/10.1016/0140-9883(96)00020-5).
- Fulwood, Mike. *A New Global Gas Order? (Part 1): The Outlook to 2030 after the Energy Crisis*. The Oxford Institute for Energy Studies, 2023. <https://www.oxfordenergy.org/wpcms/wp-content/uploads/2023/07/NG-184-A-New-Global-Gas-Order-Part-1.pdf>.
- Fulwood, Mike. *Are Asian LNG Spot Prices Finally Decoupling from Oil?* The Oxford Institute for Energy Studies, 2019. <https://www.oxfordenergy.org/wpcms/wp-content/uploads/2019/05/Are-Asian-LNG-Spot-Prices-Finally-Decoupling-from-Oil.pdf>.
- Furuoka, Fumitaka. "Natural Gas Consumption and Economic Development in China and Japan: An Empirical Examination of the Asian Context." *Renewables and Sustainable Energy Reviews* 56 (2016): 100-115. <https://doi.org/10.1016/j.rser.2015.11.038>.
- Geng, Jiang-Bo, Qiang Ji, and Ying Fan. "The Impact of the North American Shale Gas: Revolution on Regional Natural Gas Markets: Evidence from the Regime-Switching Model." *Energy Policy* 96 (2016): 167-178. <https://doi.org/10.1016/j.enpol.2016.05.047>.
- GIIGNL. *The LNG Industry: GIIGNL Annual Report*. GIIGNL, 2021. [https://giignl.org/wp-content/uploads/2021/11/GIIGNL\\_Annual\\_Report\\_November2021.pdf](https://giignl.org/wp-content/uploads/2021/11/GIIGNL_Annual_Report_November2021.pdf).
- GIIGNL. *The LNG Industry: GIIGNL Annual Report*. GIIGNL, 2023. <https://giignl.org/wp-content/uploads/2023/07/GIIGNL-2023-Annual-Report-July20.pdf>.
- Glynn, John, Nelson Perera, and Reetu Verma. "Unit Root Tests and Structural Breaks: A Survey with Applications." *Revista de Métodos Cuantitativos para la Economía y la Empresa* 3, no. 1 (2007): 63-79.
- Godfrey, L. G. "Testing Against General Autoregressive and Moving Average Error Models when the Regressors Include Lagged Dependent Variables." *Econometrica* 46, no. 6 (1978): 1293-1301. <https://doi.org/10.2307/1913829>.
- Granger, C. W. J., and P. Newbold. "Spurious Regressions in Econometrics." *Journal of Econometrics* 2, no. 2 (1974): 111-120. [https://doi.org/10.1016/0304-4076\(74\)90034-7](https://doi.org/10.1016/0304-4076(74)90034-7).
- Gujarati, Damodar N. *Basic Econometrics*. New York: McGraw-Hill, 2003.
- Hamilton, James D. "Measuring Global Economic Activity." *Journal of Applied Econometrics* 36, no. 3 (2019): 293-303. <https://doi.org/10.1002/jae.2740>.
- Haldrup, Niels. "An Econometrics Analysis of I(2) Variables." *Journal of Economic Surveys* 12, no. 5 (2002): 595-650. <https://doi.org/10.1111/1467-6419.00069>.

- Harvey, Andrew C. *Forecasting, Structural Time Series Models and the Kalman Filter*. Cambridge: Cambridge University Press, 1989.
- Heather, Patrick, and Beatrice Petrovich. *European Traded Gas Hubs: An Updated Analysis on Liquidity, Maturity and Barriers to Market Integration*. The Oxford Institute for Energy Studies, 2017.  
<https://www.oxfordenergy.org/wpcms/wp-content/uploads/2017/05/European-traded-gas-hubs-an-updated-analysis-on-liquidity-maturity-and-barriers-to-market-integration-OIES-Energy-Insight.pdf>.
- Heather, Patrick. *European Traded Gas Hubs: Their Continued Relevance*. The Oxford Institute for Energy Studies, 2023.  
<https://www.oxfordenergy.org/wpcms/wp-content/uploads/2023/06/European-Traded-Gas-Hubs-their-continued-relevance-NG183.pdf>.
- Hodrick, Robert J., and Edward C. Prescott. "Postwar U.S. Business Cycles: An Empirical Investigation." *Journal of Money, Credit and Banking* 29, no. 1 (1997): 1-16. <https://doi.org/10.2307/2953682>.
- Hulshof, Daan, Jan-Pieter van der Maat, and Machiel Mulder. "Market Fundamentals, Competition and Natural-Gas Prices." *Energy Policy* 94 (2016): 480-491. <https://doi.org/10.1016/j.enpol.2015.12.016>.
- Hunt, Lester, and Neil Manning. "Energy Price- and Income-Elasticities of Demand: Some Estimates for the UK Using the Cointegration Procedure." *Scottish Journal of Political Economy* 36, no. 2 (1989): 183-193.  
<https://doi.org/10.1111/j.1467-9485.1989.tb01085.x>.
- Hupka, Yuri, Ivilina Popova, Betty Simkins, and Thomas Lee. "A Review of the Literature on LNG: Hubs Development, Market Integration, and Price Discovery." *Journal of Commodity Markets* 31 (2023): 100349.  
<https://doi.org/10.1016/j.jcomm.2023.100349>.
- International Gas Union. *World LNG Report 2021*. International Gas Union, 2021.  
<https://igu.org/resources/world-lng-report-2021>.
- International Energy Agency. *Japan 2021: Energy Policy Review*. International Energy Agency, 2021. [https://iea.blob.core.windows.net/assets/3470b395-cfdd-44a9-9184-0537cf069c3d/Japan2021\\_EnergyPolicyReview.pdf](https://iea.blob.core.windows.net/assets/3470b395-cfdd-44a9-9184-0537cf069c3d/Japan2021_EnergyPolicyReview.pdf).
- International Energy Agency. *Korea 2020: Energy Policy Review*. International Energy Agency, 2020. [https://iea.blob.core.windows.net/assets/90602336-71d1-4ea9-8d4f-efeeb24471f6/Korea\\_2020\\_Energy\\_Policy\\_Review.pdf](https://iea.blob.core.windows.net/assets/90602336-71d1-4ea9-8d4f-efeeb24471f6/Korea_2020_Energy_Policy_Review.pdf).
- International Energy Agency. *The Future of Hydrogen: Seizing Today's Opportunities*. International Energy Agency, 2019.  
[https://iea.blob.core.windows.net/assets/9e3a3493-b9a6-4b7d-b499-7ca48e357561/The\\_Future\\_of\\_Hydrogen.pdf](https://iea.blob.core.windows.net/assets/9e3a3493-b9a6-4b7d-b499-7ca48e357561/The_Future_of_Hydrogen.pdf)

- Johansen, Soren. *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models*. Oxford: Oxford University Press, 1995.
- Joslow, Paul L. "Natural Gas: From Shortages to Abundance in the United States." *The American Economic Review* 103, no. 3 (2013): 338-43. [10.1257/aer.103.3.338](https://doi.org/10.1257/aer.103.3.338).
- Kalman, R. E. "A New Approach to Linear Filtering and Prediction Problems" *Journal of Basic Engineering* 82, no. 1 (1960): 35-45. <https://doi.org/10.1115/1.3662552>.
- Kalman, R. E., and R. S. Bucky. "New Results in Linear Filtering and Prediction Theory". *Journal of Basic Engineering* 83, no. 1 (1961): 95-108. <https://doi.org/10.1115/1.3658902>.
- Kim, Man-Keun, and Deo-Wook Kim. "Leading and Lagging Natural Gas Markets Between Asia and Europe." *OPEC Energy Review* 43, no. 3 (2019): 383-395. <https://doi.org/10.1111/opec.12163>.
- Kilian, Lutz. "Not All Oil Price Shocks are Alike: Disentangling Demand and Supply Shocks in the Crude Oil Market." *American Economic Review* 99, no. 3 (2009): 1053-1069. <https://doi.org/10.1257/aer.99.3.1053>.
- Kim, Sang-Hyun, Yeon-Yi Lim, Dae-Wook Kim, and Man-Keun Kim. "Swing Suppliers and International Natural Gas Markets Integration." *Energies* 13, no. 18 (2020): 4661. <https://doi.org/10.3390/en13184661>.
- Kisswani, Khalid M. "(A)symmetric Time-Varying Effects of Uncertainty Fluctuations on Oil Price Volatility: A Nonlinear ARDL Investigation." *Resources Policy* 73 (2021): 102210. <https://doi.org/10.1016/j.resourpol.2021.102210>.
- Kisswani, Khalid M, Michel Zaitouni, and Omar Moufakkir. "*Current Issues in Tourism* 23, no. 4 (2020): 500-522. <https://doi.org/10.1080/13683500.2019.1629578>.
- Kotek, Péter, Borbála T. Tóth, and András Mezösi. *REKK Policy Brief. What Caused the 2018 March Price Spike on TTF?* REKK, 2018. <https://rekk.hu/research-paper/82/what-caused-the-2018-march-price-spike-on-ttf>.
- Kutcherov, Vladimir, Maria Morgunova, Valery Bessel, and Alexey Lopatin. "Russian Natural Gas Exports: An Analysis of Challenges and Opportunities." *Energy Strategy Reviews* 30 (2020): 100511. <https://doi.org/10.1016/j.esr.2020.100511>.
- Kwiatkowski, Denis, Peter C. B. Phillips, Peter Schmidt, and Yongcheol Shin. "Testing the Null Hypothesis of Stationarity Against the Alternative of a Unit Root." *Journal of Econometrics* 54 (1992): 159-178.

- Lütkepohl, Helmut, and Markus Krätzig. *Applies Time Series Econometrics*. Cambridge: Cambridge University Press, 2004.
- Leamer, Edward E. "Vector Autoregressions for Causal Inferences?" *Carnegie-Rochester Conference Series on Public Policy* 22 (1985): 255-304. [https://doi.org/10.1016/0167-2231\(85\)90035-1](https://doi.org/10.1016/0167-2231(85)90035-1).
- Lee, Seo-Young, Ju-Hee Kim, and Seung-Hoon Yoo. "Role of Natural Gas Supply Sector in the National Economy: A Comparative Analysis Between South Korea and Japan." *Applied Science* 13, no. 3 (2023): 1689. <https://doi.org/10.3390/app13031689>.
- Li, Raymond, Roselyne Joyeux, and Ronald D. Ripple. "International Natural Gas Markets." *The Energy Journal* 35, no. 4 (2014): 159-179. <https://doi.org/10.5547/01956574.35.4.7>.
- Liao, Ruth, and Patrick Sykes. *Going Global: The TTF as an LNG Benchmark*. Independent Commodity Intelligence Service (ICIS), 2019. [https://s3-eu-west-1.amazonaws.com/cjp-rbi-icis/wp-content/uploads/sites/7/2019/09/13114453/ICIS\\_market\\_insight\\_going\\_global\\_the\\_ttf\\_as\\_a\\_global\\_benchmark.pdf](https://s3-eu-west-1.amazonaws.com/cjp-rbi-icis/wp-content/uploads/sites/7/2019/09/13114453/ICIS_market_insight_going_global_the_ttf_as_a_global_benchmark.pdf).
- Lin, Boqiang, and Tong Su. "Does COVID-19 Open a Pandora's Box of Changing the Connectedness in Energy Commodities?" *Research in International Business and Finance* 56 (2021): 101360. <https://doi.org/10.1016/j.ribaf.2020.101360>.
- Lin, Boqiang, and Zhensheng Li. "Analysis of the Natural Gas Demand and Subsidy in China: A Multi-Sectoral Perspective." *Energy* 202 (2020): 117786. <https://doi.org/10.1016/j.energy.2020.117786>.
- Liu, Guixian, Xiucheng Dong, Qingzhe Jiang, Cong Dong, and Jiaman Li. "Natural Gas Consumption of Urban Households in China and Corresponding Influencing Factors." *Energy Policy* 122 (2018): 17-26. <https://doi.org/10.1016/j.enpol.2018.07.016>.
- Loureiro, Jose R., Julian Inchauspe, Roberto F. Aguilera. "World Regional Natural Gas Prices: Convergence, Divergence or What? New Evidence." *Journal of Commodity Markets* 32 (2023): 100368. <https://doi.org/10.1016/j.jcomm.2023.100368>.
- MacKinnon, James G. "Critical Values for Cointegration Tests". In *Long-Run Economic Relationships*, edited by R. F. Engle and C. W. J. Granger, 267-276. Oxford: Oxford University Press, 1991.
- McKinsey & Company. *The Energy Transition: A Region-by-Region Agenda for Near-term Action*. McKinsey & Company, Accessed May 13, 2024. <https://www.mckinsey.com/industries/electric-power-and-natural-gas/our-insights/the-energy-transition-a-region-by-region-agenda-for-near-term-action>.

- Ministry of Trade, Industry and Energy. *9<sup>th</sup> Basic Plan for Power Supply and Demand*. Ministry of Trade, Industry and Energy, Accessed May 13, 2024. <https://policy.asiapacificenergy.org/node/4314>.
- Mu, Xiaoyi, and Haichun Ye. "Towards an Integrated Spot LNG Market: An Interim Assessment". *The Energy Journal* 39, no. 1 (2018): 211-234. <https://dx.doi.org/10.2139/ssrn.2967797>.
- Murshed, Muntasir. "LPG Consumption and Environmental Kuznets Curve Hypothesis in South Asia: a Time-series ARDL Analysis with Multiple Structural breaks." *Environmental Science and Pollution Research* 28 (2020): 8337-8372. <https://doi.org/10.1007/s11356-020-10701-7>.
- Nelson, Charles R, and Charles R Plosser. "Trends and Random Walks in Macroeconomic Time Series: Some Evidence and Implications." *Journal of Monetary Economics* 10, no. 2 (1982): 139-162. [https://doi.org/10.1016/0304-3932\(82\)90012-5](https://doi.org/10.1016/0304-3932(82)90012-5).
- Netfçi, Salih N. "Are Economics Time Series Asymmetric Over the Business Cycle?". *Journal of Political Economy* 92, no. 2 (1984): 307-328. <https://doi.org/10.1086/261226>.
- Neumann, Anne, Boriss Siliverstovs and Christian von Hirschhausen. "Convergence of European Spot Market Prices for Natural Gas? A Real-Time Analysis of Market Integration Using the Kalman Filter." *Applies Economics Letters* 13, no. 11 (2006): 727-732. <https://doi.org/10.1080/13504850500404258>.
- Neumann, Anne. "Linking Natural Gas Markets – Is LNG Doing Its Job?" *The Energy Journal* 30, no. 1 (2009): 187-199. <https://doi.org/10.5547/ISSN0195-6574-EJ-Vol30-NoSI-12>.
- Newell, Richard G., and Daniel Raimi. "Implications of Shale Gas Development for Climate Change." *Environmental Science & Technology* 48, no. 15 (2014): 8360-8388. doi: 10.1021/es4046154.
- Nick, Sebastian, and Benjamin Tischler. "The Law of One Price in Global Gas Markets – A Threshold Cointegration Analysis". EWI Working Papers 2014-16. Energiewirtschaftliches Institut an der Universitaet zu Koeln (EWI), 2014.
- Nkoro, Emeka, and Aham K. Uko. "Autoregressive Distributed Lag (ARDL) Cointegration Technique: Application and Interpretation." *Journal of Statistical and Econometric Methods*. 5, no. 4 (2016): 63-91.
- OzForex Limited, "Historical Exchange Rates" Accessed May 13, 2024. <https://www.ofx.com/en-au/forex-news/historical-exchange-rates/monthly-average-rates/>.



- Oglend, Atle, Tore S. Kleppe, and Petter Osmundsen. "Trade with Endogenous Transportation Costs: The Case of Liquefied Natural Gas." *Energy Economics* 59, (2016): 138-148. <https://doi.org/10.1016/j.eneco.2016.08.013>.
- Olsen, Kyle K., James W. Mjelde, and David A. Bessler. "Price Formulation and the Law of One Price in Internationally Linked Markets: An Examination of the Natural Gas Markets in the USA and Canada." *The Annals of Regional Science* 54 (2015): 117-142. <https://doi.org/10.1007/s00168-014-0648-7>.
- Pata, Ugur Korkut, and Abdullah E. Caglar. "Investigating the EKC Hypothesis with Renewable Energy Consumption, Human Capital, Globalization and Trade Openness for China: Evidence from Augmented ARDL Approach with Structural Break." *Energy* 216 (2021): 119220. <https://doi.org/10.1016/j.energy.2020.119220>.
- Perifanis, Theodosios, and Athanasios Dagoumas. "Price and Volatility Spillovers Between the US Crude Oil and Natural Gas Wholesale Markets" *Energies* 11, no. 10 (2018): 2757. <https://doi.org/10.3390/en11102757>.
- Perron, Pierre, and Timothy J. Vogelsang. "Nonstationarity and Level Shifts with an Application to Purchasing Power Parity" *Journal of Business & Economic Statistics* 10, no. 3 (1992): 301-320. <https://doi.org/10.2307/1391544>.
- Perron, Pierre. "Further Evidence on Breaking Trend Functions in Macroeconomic Variables." *Journal of Econometrics* 80, no. 2 (1997): 355-385. [https://doi.org/10.1016/S0304-4076\(97\)00049-3](https://doi.org/10.1016/S0304-4076(97)00049-3).
- Perron, Pierre. "The Great Crash, the Oil Price Shock, and the Unit Root Hypothesis." *Econometrica* 57, no. 6 (1989): 1361-1401. <https://doi.org/10.2307/1913712>.
- Pesaran, M. Hashem, Yongcheol Shin, and Richard J. Smith. "Bounds Testing Approach to the Analysis of Level Relationships." *Journal of Applied Econometrics* 16, no. 3 (2001): 1924-1996. <https://doi.org/10.1002/jae.616>.
- Pesaran, M. Hashem, and Yongcheol Shin. *An Autoregressive Distributed Lag Modelling Approach to Cointegration Analysis*. In: *Econometrics and Economic Theory in the 20<sup>th</sup> Century the Ragnar Frisch Centennial Symposium*, edited by Storm, S. 371-413. Cambridge: Cambridge University, 1999.
- Phillips, Peter C. B., and Donggyu Sul. "Economic Transition and Growth" *Journal of Applied Econometrics* 24, no. 7 (2009): 1153-1185. <https://doi.org/10.1002/jae.1080>.
- Phillips, Peter C. B., and Donggyu Sul. "Transition Modeling and Econometrics Convergence Tests." *Econometrica* 75, no. 6 (2007): 1771-1855. <https://doi.org/10.1111/j.1468-0262.2007.00811.x>.

- Phillips, Peter C., and Pierre Perron. "Testing for a Unit Roots in Time Series Regression." *Biometrika* 75, no. 2 (1988): 335-346. <https://doi.org/10.1093/biomet/75.2.335>.
- Pustišek, Andrej, and Michael Karasz. *Natural Gas: A Commercial Perspective*. Cham: Springer International Publishing, 2017.
- Ravazzolo, Francesco, and Joaquin Vespignani. "World Steel Production: A New Monthly Indicator of Global Real Economic Activity." *Canadian Journal of Economics* 53, no. 2 (2020): 743-766. <https://doi.org/10.1111/caje.12442>.
- Renou-Maissant, Patricia. "Toward the Integration of European Natural Gas Markets: A Time-Varying Approach" *Energy Policy* 51 (2012): 779-790. <https://doi.org/10.1016/j.enpol.2012.09.027>.
- Rodríguez, Ramón Yepes. "Real Option Valuation of Free Destination in Long-Term Liquefied Natural Gas Supplies." *Energy Economics* 30, no. 4 (2008): 1909-1932. <https://doi.org/10.1016/j.eneco.2007.03.006>.
- Runkle, David E. "Vector Autoregressions and Reality." *Journal of Business & Economics Statistics* 5, no. 4: 437-442. <https://doi.org/10.2307/1391992>.
- Satterthwaite, F. E. "An Approximate Distribution of Estimates of Variance Components." *Biometrics Bulletin* 2, no. 6 (1946): 110-114. <https://doi.org/10.2307/3002019>.
- Schwarz, Thomas V., and Andrew C. Szakmary. "Price Discovery in Petroleum Markets: Arbitrage, Cointegration, and the Time Interval of Analysis." *The Journal of Futures Markets* 14, no. 2 (1994): 147-167. <https://doi.org/10.1002/fut.3990140204>.
- Shahati, Mohamed El, Haidar Khadadeh, and Fajer Al-Aradah. "The Global LNG Price Trend and the Role of LNG in Balancing the Gas Demand in MENA Region." Paper presented at the *SPE Kuwait Oil & Gas Show and Conference*, Mishref, Kuwait, October 2019. doi: <https://doi.org/10.2118/198054-MS>
- Shi, Xupeng, and Hari M. P. Variam. "Gas and LNG trading hubs, hub indexation and destination flexibility in East Asia." *Energy Policy* 96 (2016): 587-596. <https://doi.org/10.1016/j.enpol.2016.06.032/>
- Shi, Xupeng, and Yifan Shen. "Macroeconomics uncertainty and natural gas prices: Revisiting the Asian Premium." *Energy Economics* 94 (2021): 1-13. doi: [10.1016/j.eneco.2020.105081](https://doi.org/10.1016/j.eneco.2020.105081).
- Shin, Yongcheol, Byungchul Yu, and Matthew Greenwood-Nimmo. *Modelling Asymmetric Cointegration and Dynamic Multipliers in a Nonlinear ARDL Framework*. In: *Festschrift in Honor of Peter Schmidt*, edited by R. Sickles and W. Horrace. New York: Springer, 2014.

- Silverstovs, Boriss, Guillaume L'Hégaret, Anne Neumann, and Christian von Hirschhausen. "International Market Integration for Natural Gas? A Cointegration Analysis of Prices in Europe, North America and Japan." *Energy Economics* 27, no. 4 (2005): 603-615. <https://doi.org/10.1016/j.eneco.2005.03.002>.
- Sims, Christopher A. "Macroeconomics and Reality." *Econometrica* 48, no. 1 (1980): 1-48. <https://doi.org/10.2307/1912017>.
- Sims, Christopher A. "Money, Income, and Causality." *The American Economic Review* 62, no. 4 (1972): 540-552.
- Stern, Jonathan. *A comparative History of Oil and Gas Markets and Prices: Is 2020 Just an Extreme Cyclical Event or an Acceleration of the Energy Transition?* The Oxford Institute for Energy Studies, 2020. <https://www.oxfordenergy.org/wpcms/wp-content/uploads/2020/04/Insight-68-A-Comparative-History-of-Oil-and-Gas-Markets-and-Prices.pdf>.
- Stern, Jonathan. *Continental Europe Long-Term Gas Contracts: Is a Transition Away From Oil Product-linked Pricing Inevitable and Imminent?* The Oxford Institute for Energy Studies, 2009. <https://www.oxfordenergy.org/publications/continental-european-long-term-gas-contracts-is-a-transition-away-from-oil-product-linked-pricing-inevitable-and-imminent-2/>.
- Stern, Jonathan. *Is There a Rationale for the Continuing Link to Oil Product Prices in Continental European Long-Term Gas Contracts?* The Oxford Institute for Energy Studies, 2007. <https://www.oxfordenergy.org/publications/is-there-a-rationale-for-the-continuing-link-to-oil-product-prices-in-continental-european-long-term-gas-contracts-2/>.
- Stern, Jonathan, and Howard Rogers. *The Dynamics of a Liberalized European Gas Market: Key determinants of hub prices, and roles and risks of major players.* The Oxford Institute for Energy Studies, 2014. <https://www.oxfordenergy.org/publications/the-dynamics-of-a-liberalised-european-gas-market-key-determinants-of-hub-prices-and-roles-and-risks-of-major-players/>.
- Stern, Jonathan. "Stabilization of Natural Gas and LNG Prices and the Supply/Demand Balance During the Transition to Net Zero Emissions" *IEEEJ Energy Journal* (2023): 56-60.
- Stigler, George J., and Sherwin, Robert A. "The Extent of the Market" *The Journal of Law and Economics* 28, no. 3 (1985): 555-585. <https://doi.org/10.1086/467101>.
- Szafranek, K., and M. Rubaszek. "Have European Natural Gas Prices Decoupled from Crude Oil Prices? Evidence from TVP-VAR Analysis." *Studies in Nonlinear Dynamics & Econometrics* (2023). <https://doi.org/10.1515/snede-2022-0051>.

- The World Bank. Accessed April 15, 2021. <https://data.worldbank.org>.
- Theissen, Erik. "Price Discovery in Floor and Screen Trading Systems." *Journal of Empirical Finance* 9, no. 4 (2002): 455-474. [https://doi.org/10.1016/S0927-5398\(02\)00005-1](https://doi.org/10.1016/S0927-5398(02)00005-1).
- Toda, Hiro Y. "Finite Sample Properties of Likelihood Ratio Tests for Cointegration Ranks when Linear Trends are Present." *The Review of Economics and Statistics* 76, no. 1 (1994): 66-79. <https://doi.org/10.2307/2109827>.
- Toda, Hiro Y., and Taku Yamamoto. "Statistical Inferences in Vector Auto-Regressions with Possibly Integrated Processes." *Journal of Econometrics* 66, no. 1-2 (1995): 225-250. [https://doi.org/10.1016/0304-4076\(94\)01616-8](https://doi.org/10.1016/0304-4076(94)01616-8).
- U.S. Energy Information Administration. *Country Analysis Brief: South Korea*. U.S. Energy Information Administration, 2023. [https://www.eia.gov/international/content/analysis/countries\\_long/South\\_Korea/south\\_korea.pdf](https://www.eia.gov/international/content/analysis/countries_long/South_Korea/south_korea.pdf).
- U.S. Energy Information Administration. *Monthly Energy Review: March 2021*. U.S. Energy Information Administration, 2021. <https://www.eia.gov/totalenergy/data/monthly/archive/00352103.pdf>
- U.S. Energy Information Administration. *Monthly Energy Review: November 2020*. U.S. Energy Information Administration, 2020. <https://www.eia.gov/totalenergy/data/monthly/archive/00352011.pdf>.
- U.S. Department of Energy. *LNG Monthly December 2022*. U.S. Department of Energy, 2022. [https://www.energy.gov/sites/default/files/2023-03/LNG%20Monthly%20December%202022\\_3.pdf](https://www.energy.gov/sites/default/files/2023-03/LNG%20Monthly%20December%202022_3.pdf).
- Villar, Jose A., and Joutz, Frederick L. *The Relationship Between Crude Oil and Natural Gas Prices*. Energy Information Administration, 2006. [http://aceer.uprm.edu/pdfs/CrudeOil\\_NaturalGas.pdf](http://aceer.uprm.edu/pdfs/CrudeOil_NaturalGas.pdf).
- Vivoda, Vlado. "LNG import diversification and energy security in Asia." *Energy Policy* 129 (2019): 967-974. <https://doi.org/10.1016/j.enpol.2019.01.073>.
- Vivoda, Vlado. "Natural Gas in Asia: Trade, Markets and Regional Institutions." *Energy Policy* 74 (2014): 80-90. <https://doi.org/10.1016/j.enpol.2014.08.004>.
- Wang, Tiantian, Dayong Zhang, Qiang Ji, and Xupeng Shi. "Market Reforms and Determinants of Import Natural Gas Prices in China" *Energy* 196 (2020): <https://doi.org/10.1016/j.energy.2020.117105>.
- Wang, Tiantian, Wan Qu, Dayong Zhang, Qiang Ji, and Fei Wu. "Time-Varying Determinants of China's Liquefied Natural Gas Import Price: A Dynamic Model Averaging Approach." *Energy* 259 (2022): 125013. <https://doi.org/10.1016/j.energy.2022.125013>.

- The Wall Street Journal. *Long promised, the global market for natural gas has finally arrived*. The Wall Street Journal, 2017.  
<https://www.wsj.com/articles/long-promised-the-global-market-for-natural-gas-has-finally-arrived-1496761392>.
- Yu, Yihua, Xinye Zheng, and Yi Han. "On the Demand for Natural Gas in Urban China." *Energy Policy* 70 (2014): 57-63.  
<https://doi.org/10.1016/j.enpol.2014.03.032>.
- Yule, G. Udny. "Why do we Sometimes get Nonsense-Correlations between Time-Series?--A Study in Sampling and the Nature of Time Series." *Journal of the Royal Statistical Society* 89, no.1 (1926): 1-63.  
<https://www.jstor.org/stable/pdf/2341482>.
- Zaklan, Aleksandar, Jan Abrell, and Anne Neumann. "Stationarity Changes in Long-Run Energy Commodity Prices." *Energy Economics* 59 (2016): 96-103.  
<https://doi.org/10.1016/j.eneco.2016.07.022>.
- Zapata, Hector O., and Alicia N. Rambaldi. "Monte Carlo Evidence on Co-integration and Causation." *Oxford Bulletin of Economics and Statistics* 59, no. 2 (1997): 285-298. <https://doi.org/10.1111/1468-0084.00065>.
- Zhang, Dayong, and Qiang Ji. "Further Evidence on the Debate of Oil-Gas Price Decoupling: A Long Memory Approach." *Energy Policy* 113 (2018): 68-75.  
<https://doi.org/10.1016/j.enpol.2017.10.046>.
- Zhang, Dayong, Min Shi, and Xupeng Shi. "Oil Indexation, Market Fundamentals, and Natural Gas Prices: An Investigation of the Asian Premium in Natural Gas Trade" *Energy Economics* 69 (2018): 33-41.  
<https://doi.org/10.1016/j.eneco.2017.11.001>.
- Zhang, Dayong, Tiantian Wang, and Xupeng Shi. "Is Hub-Based Pricing a Better Choice than Oil Indexation for Natural Gas? Evidence from a Multiple Bubble Test." *Energy Economics* 76 (2018): 495-503.  
<https://doi.org/10.1016/j.eneco.2018.11.001>.
- Zhang, Lingge, Dong Yang, Shining Wu, and Meifeng Luo. "Revisiting the Pricing Benchmarks for Asian LNG – An Equilibrium Analysis." *Energy* 262(A) (2023): 125426. <https://doi.org/10.1016/j.energy.2022.125426>.
- Zhang, Yi, Qiang Ji, and Ying Fan. "The Price and Income Elasticity of China's Natural Gas Demand: A Multi-Sectoral Perspective." *Energy Policy* 113 (2018): 332-341. <https://doi.org/10.1016/j.enpol.2017.11.014>.
- Zivot, Eric, and Donald W. K. Andrews. "Further Evidence on the Great Crash, the Oil-Price Shock, and the Unit-Root Hypothesis." *Journal of Business & Economic Statistics* 10, no. 3 (1992): 251-270.  
<https://doi.org/10.2307/139154>.