Three Essays about Energy Prices and Energy Markets

Mousa Tawfeeq

Dissertation submitted to the

Davis College of Agriculture, Natural Resources and Design

at West Virginia University

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

in

Natural Resource Economics

Alan R. Collins, Ph.D., Committee Chair

Adam Nowak, Ph.D.

Stratford Douglas, Ph.D.

Xiaoli, Etienne, Ph.D.

Levan Elbakidze, Ph.D.

Division of Resource Economics and Management

Morgantown, West Virginia

2018

Keywords: energy prices, stock markets, spatial panel data, coal consumption, CO₂ emissions

Copyright 2018 Mousa Tawfeeq
Abstract

Three Essays about Energy Prices and Energy Markets
Mousa Tafweeq

In this dissertation, three related issues concerning empirical time series and panel models for three essays in energy economics will be investigated. The overall theme of these essays is to explore the relationships between the use of and prices for fossil fuels (coal, crude oil, and natural gas) along with larger societal issues of stock market changes, resource extraction rates, and CO₂ emissions. The three main questions will be addressed: (1) how do crude oil prices affect stock markets?, (2) how have the dynamic linkages between the coal market and natural gas prices changes with the shale gas revolution?, and (3) how have urbanization and other factors impacted CO₂ emissions? For each question, it is considered proper econometric models to provide empirical answers which will contribute either to the academic literature or energy economics and policy.

In the first essay, we investigate linkages between oil price changes and stock market capitalization in Middle Eastern (ME) economies from 2000-2015. This essay examines how oil price changes influence stock market capitalization by using a variety of econometric models including structural breaks, VECM, VAR, and IRF. These results suggest the existence of positive and significant impacts of oil price on MC for all oil-exporting economies.

The second essay introduces the idea of a dynamic relationship between coal plus-natural gas prices and their impacts on coal consumption and extraction. In this essay, we adopt an Autoregressive Distributed Lag (ARDL) model to investigate the short-run and long-run characteristics of coal demand in the U.S. from 2000-2016. Methods and data in this research
contribute to the literature on how the natural gas revolution affects the coal industry in the U.S. The findings show that coal and natural gas are substitutes as energy sources in the U.S. energy market. Moreover, income elasticity reveals that coal is an inferior source of energy since economic growth has a negative impact on coal demand in the long run. Solar energy and weather both have significant negative impacts on coal consumption over the entire time period. Results provide theoretical and empirical methods of the coal market, and allow for a better understanding of its factors. Also, these results support the policies to substitute natural gas for coal due to coal-natural gas prices, efficiency, and environment friendly.

The final essay combines the issues of urbanization, energy consumption, coal use, energy prices, and CO₂ emissions regarding U.S. fossil fuel consumption. Fossil fuels make up more than 80% of the human’s energy use (U.S. EIA, 2016) and more than 75% of fossil fuels are used in urban areas. Therefore, dynamic long-run linkages between fossil fuel energy use and urbanization are of importance to energy policy makers across the globe. This essay focuses on spillover impacts of energy consumption on CO₂ emissions by using spatial autoregressive regression (SAR), spatial error model (SEM), and spatial dynamic model (SDM). Negative direct and positive indirect effects of urbanization on state level CO₂ emissions are found along with the negative effects of income along with renewable portfolio standards (RPS) on state level emissions. The income effects provide additional support for the Environmental Kuznets Curve on CO₂ emissions such that income growth reduces per capita CO₂ emissions at an own state level. For RPS, however, the indirect effects of CO₂ emission reductions are as large as the own state direct effects, indicating that cross border impacts are important. Both of these results are important due to the prominence of state and local government actions in climate change mitigation now that the federal government has withdrawn from the Paris Climate Accords.
DEDICATION

I would like dedicate this work to Sharmin and Dlovan for providing me with the unconditional love and support throughout all of my pursuits. My Mom who had dedicated her life to entire family. My younger brother who sacrificed his life for humanity and dignity.
ACKNOWLEDGEMENTS

I am very grateful for the support and guidance I have received from my family and faculty throughout my entire graduate career. I would like to extend a special thanks to my advisor, Dr. Alan Collins, for his support and guidance throughout my time at WVU. I also wish to thank Dr. Gerard D’Souza, Dr. Levan Elbakizde, Dr. Xiaoli Etienne, Dr. Stratford Douglas, and Dr. Adam Nowak for providing me with invaluable advice and assistance through this process. Finally, I would also like to thank my fellow grad students for with whom I have spent countless hours studying with and listening to my ideas on economics, energy economics, and the future of natural resource economics.
Table of Contents

Title Page ......................................................................................................................................i
Abstract .........................................................................................................................................ii
Dedication .......................................................................................................................................iv
Acknowledgements ...........................................................................................................................v
List of Figures ..................................................................................................................................vii
List of Tables ...................................................................................................................................viii

CHAPTER 1: Introduction .....................................................................................................................1

CHAPTER 2- Essay #1: Linking Crude Oil Prices and Middle East Stock Markets ..................................5
1. Introduction ...................................................................................................................................5
2. Literature Review ..........................................................................................................................9
3. Theory .........................................................................................................................................11
4. Methodology ...............................................................................................................................14
5. Data and Descriptive Statistics ....................................................................................................20
6. Results and Discussion ................................................................................................................22
7. Conclusions .................................................................................................................................28
8. References ...................................................................................................................................32

Chapter 3- Essay #2: The Dynamic Response of Coal Consumption to Energy Prices and GDP: An ARDL Approach to the U.S..............................................................................................52
1. Introduction ..................................................................................................................................52
2. Literature Review ..........................................................................................................................57
3. Methodology ...............................................................................................................................59
4. Data ............................................................................................................................................64
5. Results .........................................................................................................................................67
6. Conclusions .................................................................................................................................74
7. References ...................................................................................................................................77

Chapter 4- Essay #3: The Spillover Impacts of Urbanization and Energy Usage on CO₂ Emissions Patterns in the U.S ..................................................................................................................88
1. Introduction ..................................................................................................................................88
2. Literature Review ..........................................................................................................................93
3. Theoretical Model .........................................................................................................................98
4. Empirical Model ..........................................................................................................................102
5. Data ............................................................................................................................................106
6. Results and Discussion .................................................................................................................109
7. Conclusions .................................................................................................................................114
8. References ...................................................................................................................................116

Chapter 5: Conclusions ......................................................................................................................127
List of Figures

Essay #1:

Figure 2.1: Mechanisms of crude oil price changes on an oil importer stock market, increase (left arrow) and decrease (right arrow)………………………………………………………………………………13

Figure 2-2 Market capitalization (MC in billion USD) and crude oil price (OP in USD)……..47

Figure 2-3: Impulse response function (IRF) for each Middle East economy………………..48

Essay #2:

Figure 3-1: U.S. power generation from coal and natural gas.............................................54

Figure 3-2: Coal producing states and change in coal production.....................................55

Figure 3-3: Economic growth (US$)..........................................................83

Figure 3-4: Coal consumption (MBTU)...........................................................................83

Figure 3-5: Coal–natural gas prices ($/MBTU).................................................................83

Figure 3-6: Heating degree-days.....................................................................................83

Figure 3-7: Cooling degree-days....................................................................................84

Figure 3-8: Solar consumption (TBTU)..........................................................................84

Figure 3-9: Plot of the cumulative sum of recursive residuals (CUSUM).......................84

Figure 3-10: Impulse response function (combined).......................................................85

Panel A: 2000 to 2016 Data.........................................................................................85

Panel B: 2000 to 2008 Data.........................................................................................86

Panel C: 2009 to 2016 Data........................................................................................86

Essay #3:

Figure 4-1: Urban and rural populations of the world, 1950–2050...................................88

Figure 4-2: CO₂ emissions from 1750 and projected to 2050.........................................88
List of Tables

**Essay #1:**

Table 2-1: Macroeconomic overview of ME region…………………………………………..37
Table 2-2: Descriptive statistics for MC (billion USD) and for OP (USD per barrel)…………..37
Table 2-3: Tests for unit root hypothesis for market capitalization and oil price………………..38
Table 2-4: Selection of optimum lag according AIC and BIC criteria…………………………39
Table 2-5: Tests of cointegration for MC and OP………………………………………………39
Table 2-6: The result of VECM for MC and OP, 2001-2008……………………………………40
Table 2-7: The result of VECM for MC and OP 2009-2015……………………………………41
Table 2-8: The result of VECM for MC and OP 2001-2015……………………………………42
Table 2-9: The result of VAR for MC and OP 2001-2008………………………………………43
Table 2-10: The result of VAR for MC and OP 2009-2015……………………………………44
Table 2-11: The result of VAR for MC and OP 2001-2015……………………………………45
Table 2-12: Summary of results………………………………………………………………….46

**Essay #2:**

Table 3-1: Descriptive statistics…………………………………………………………………80
Table 3-2: Tests for unit root test……………………………………………………………..80
Table 3-3: Results of F bound test…………………………………………………………81
Table 3-4: Estimated long run coefficients…………………………………………………81
Table 3-5: Estimated short run coefficients…………………………………………………82
Table 3-6: Diagnostic tests statistics……………………………………………………………82

**Essay #3:**

Table 4-1: Descriptive statistics………………………………………………………………123
Table 4-2: Post diagnostic tests …………………………………………………………….123
Table 4-3: Estimation results of non-spatial- panel data models……………………………..124
Table 4-4: Estimation results of spatial- panel data models…………………………………125
Table 4-5: Estimation results of direct, indirect and total effect of the SDM…………………126
Chapter 1: Introduction

Economists consider profoundly about energy price volatilities, stock markets, and energy consumption as vital parts of an economy. In the face of challenges and opportunities in the energy market, crude oil prices and, energy markets and urbanization have become concerns of the global community (Hamilton, 2003; Kilian, 2006; Poumanyvong and Shinji, 2010; Teulon, 2015; Xu and Boqiang, 2015). To address these concerns, many economists are paying attention to the dynamic linkages between oil price and the stock market value, between coal and natural gas prices and coal consumption, and among urbanization, energy factors, and CO₂ emissions (Sadorsky, 2006; York, 2003; Sadorsky, 2014; Burnett et.al, 2013). The three research studies cover the impact of economic and policy analysis on energy and environmental economics. They include crude oil prices and stock market in the Middle Eastern economies, the link between coal or natural gas prices and coal consumption in the U.S., and urbanization, energy use and CO₂ emissions in the U.S. In these essays, a variety of non-spatial and spatial econometric models are utilized to identify links among variables.

Chapter two focuses on dynamic linking between crude oil prices and stock market capitalization. The economy of the Middle East (ME) region is diverse in nature, composed of 17 countries, which includes both oil exporting and importing countries. The economies of this region depend heavily on hydrocarbon reserves. Major oil exporting economies in this region include Iraq, Saudi Arabia, Qatar, Bahrain, Oman, Yemen, UAE, Kurdistan region, Kuwait and Iran. In this essay, we investigate nine ME economies including oil-exporting and oil-importing since there is no accurate data for other eight countries and authorities. In the oil-exporting economies, a significant portion of total GDP comes from hydrocarbon exports. More importantly, ME oil exporters rely heavily on oil prices as shown by fluctuations in economic
performance when oil prices change. The negative impact of oil price volatility has affected structural reform initiatives in ME economies to achieve stability and sustainability of economic growth (Dasgupta et al., 2002). Moreover, increasing the role of stock market and crude oil have been the vital part of in the ME economies and stock market activities in this region. Given these facts, it is reasonable to suspect that crude oil prices have economic and financial effects within countries of the ME, yet little is known about these impacts in the ME region. Therefore, the main objective of this study is to show whether the oil prices changes impact stock market capitalization in economies across this region.

Chapter three examines U.S. coal consumption and the impacts of prices (coal and natural gas), weather, and economic growth on the use of this resource. In recent years, the coal industry has faced decreasing production with the resultant negative effects on job markets and incomes in the coal producing states. While arguably cleaner than coal, natural gas is a possible substitute to coal since it is more efficient and less emits environment than coal. Moreover, coal industry is in a crisis condition in the U.S. is receiving significant attention recently. Since fuel choice for power generation relies on fuel prices, using ARDL model, we investigate the dynamic short and long-run correlations between coal consumption, natural gas price, price of coal and other factors over time period 2000-2016. The findings show that there exist long-run relationships among variables. Coal and natural gas prices have negative and positive linkages with consumption of coal, respectively.

Chapter four examines spillover effects of urbanization and energy factors on explaining state level CO₂ emissions. A number of previous studies have noted the geographic neighborhood and location in CO₂ emissions (Conley and Ligon, 2002) Chuai, et al., 2012; Yu, 2012; Burnett et al., 2013). Thus, we model the issue of CO₂ emissions using spatial
econometric methods. Given the potential spatial correlations between CO\textsubscript{2} emissions and state level independent variables, we utilize models: Spatial Autoregressive Regression (SAR), Spatial Error (SEM), and Spatial Durbin (SDM). Results suggest that coal consumption and energy use have larger effects on CO\textsubscript{2} emissions than energy prices, urbanization, GSP and RPS. Further, urbanization has a negative direct impact on own state CO\textsubscript{2} emissions, but positive indirect effects on emissions in neighboring states.

This dissertation takes a comprehensive view of the energy prices, energy market and CO\textsubscript{2} emissions examining both U.S and ME economies. Historically, coal extraction and its consumption have been dominant in the electricity energy market. The final essay attempts to better understand what roles coal and natural gas prices have played in the decline of coal production, while controlling for weather and renewable energy. Here, we model coal consumption and evaluate how the dynamic relationships have changed since the development of shale gas\textsuperscript{1}. In the fourth chapter, we focus on the spillover effects that state level energy factors and GSP have on per capita CO\textsubscript{2} emissions. With this dissertation, researchers and policymakers will have a better understanding of the U.S energy market and the ME stock market value, and can better address the challenges and opportunities offered energy prices and energy prices.

The dissertation consists of five chapters including the introduction. Chapter 2 offers the Essay 1 that examines dynamic linking between oil price changes and stock market capitalization in the ME. Chapter 3 presents Essay 2 that provides an empirical analysis of how energy prices, weather, and income influence coal consumption in the U.S’ energy market. Chapter 4 represents Essay 3 that examines whether the energy factors of state-level affect carbon dioxide emission

\textsuperscript{1} We explore structural changes in energy markets in all three essays due to recent increases in shale oil and gas production using hydraulic fracturing technology.
level is convergence in the U.S. This dissertation concludes in Chapter 5 with directions for future research.
Chapter 2 - Essay #1: Linking Crude Oil Prices and Middle East Stock Markets

1. Introduction

Due to the vital role of crude oil in the global economy, there has been a vast amount of studies intended to determine how crude oil price changes impact world economies and financial markets. A large number research studies have found that there exists statistically significant impacts of crude oil price on economic growth and stock market indicators (Driesprong et al., 2008; Jones and Kaul, 1996; Narayan and Sharma, 2011; Lee et al., 2012; Sadorsky, 1999; Park and Ratti, 2008; Scholtens and Yurtsever, 2012; Apergis and Miller, 2009; Güntner, 2013; Kilian and Park, 2009).

Fewer studies have focused on the impact of crude oil price changes on where large reserves and production of crude oil exist - Middle East (ME) stock markets. Moreover, there is no research that links crude oil prices and market capitalization (as a measure of stock market changes) in the ME economies. The economic and financial impacts of crude oil prices have to be captured by stock market value, for they are bound to influence cash flows, outstanding shares, and the price of shares in the stock markets. Findings from previous research indicate that crude oil price changes do have a significant impacts on the market performances (Jones et al., 2004; Basher and Sadorsky, 2006; Miller and Ratti, 2009; Filis et al., 2011; Arouri and Roult, 2012; Le and Chang, 2015; Ewing and Malik, 2016). In a recent study, Ghalayini (2011) emphasize that crude oil price is an essential element of ME economies. The ME economies have experienced with growth and crisis related to variability in crude oil prices. Considering that there exists, in fact, an economic effect of the crude oil price volatility should also be captured by market capitalization in the ME stock markets.

Previous empirical studies on crude oil price and stock market activities use stock market
returns or stock prices specifications (i.e., Driesprong et al., 2008; Jones and Kaul, 1996; Narayan and Sharma, 2011; Lee et al., 2012; Sadorsky, 1999; Park and Ratti, 2008; Scholtens and Yurtsever, 2012; Park and Ratti, 2008; Apergis and Miller, 2009; Güntner, 2013; Kilian and Park, 2009). We argue that there are problems with how stock indexes are calculated that might lead to disadvantages. For instance, the TA-35 index in Israel is a price-weighted index which is calculated by taking the sum of the prices of all 35 stocks in the index. Stocks with higher prices have a larger impact on movements in the index as compared to lower-priced stocks. However, stock market capitalization (MC) provides the total value of a company indicating its value in the market and economy. By looking at MC, a proper evaluation is provided for the current situation of financial and economic development. Furthermore, time series data of MC movements could provide trend to investors as to how the value of companies in the stock markets. In this view, MC is a macroeconomic indicator which may allow for policy makers and investors to make better decisions in an economy.

ME stock markets have been affected by crude oil price changes before and after global financial crisis. The ME countries experienced strong economic growth during 2000–2008 as a result of higher oil prices (Selvik and Stenslie, 2011). However, the global financial crisis has transformed into a severe financial and economic crisis in most of the ME economies. Although, since 2002, the world’s crude oil prices have risen rapidly, they have fallen sharply to an even $35 per barrel recently. In November 2007, prices of both WTI’s OP and the Brent’s OP went beyond $90 per barrel and the recent record peak of US$145 per barrel reached in July 2008. This resulted in great fluctuations in the OP market, causing the ME stock markets to major changes, which in turn changed the financial markets and altered the entire economies of region into crisis. Since many of the ME economies represent the majority of oil-trade, ME policymakers
must not only take into consideration how their decisions affect energy and oil prices, but also the impact that oil price shocks have on their own stock markets. Hence, this empirical study investigates a dynamic linkages between OP and MC in the ME stock markets. 

In this context, the objective of this study is to examine the impact of changes in crude oil prices (OP) on MC in the stock markets of nine ME economies including a mixture of oil-exporting and oil-importing economies (United Arab Emirates, Bahrain, Egypt, Israel, Jordan, Kuwait, Qatar, Saudi Arabia, and Turkey). Most research referring to crude oil prices and stock markets concentrates on how various indications of changes of crude oil price could influence stock market. Several studies have examined the impact of the changes of crude oil price on stock markets. However, taking account the essential role of crude oil as a major input in production and supply of goods, it is crucial to examine how the prices of crude of oil impacts market capitalization.

In this paper, the principal focus is to examine the dynamic long-run linkages between stock MC and OP across nine ME economies from 2001 to 2015 using a Vector Autoregressive Regression (VAR) and Vector Error Correction Model (VECM). This study offers the following contributions to the literature: (1) a consideration of and testing for time series (i.e. unit root, structural break tests, misspecification tests and LM test, test for normality, and test for stability condition) and econometric models exploring the relationship between OP and stock MC in the ME region; and (2) using daily stock market capitalization which combines all companies in each country not just subset of index.

Our findings suggest four conclusions about the relationship between OP and MC. First, the Johansen’s test in the period 2001-2008 shows that there are cointegration between OP and MC in Bahrain, Kuwait, Saudi Arabia, and Turkey and no-cointegration the remaining five
economies. During the 2009-2015, there are cointegration OP and MC relationship for United Arab Emirates, Israel, Jordan, and Qatar and no cointegration for the other five economies. In the total period 2001-2015, the results suggest that the existence of a cointegration relationship among the OP and MC variables in seven of nine countries (Egypt and Jordan are exceptions). The significant linkages between OP and MC imply some degree of predictability, dependency on oil prices and association between stock market values and crude oil in the ME region. Governments, investors and policy makers could have a view to the stock markets in the ME, thus, the results may assist them to make a better decision.

Second, the empirical results of VECM indicate that a dynamic long–run response of MC to an OP increase in all countries except for Israel and Turkey, suggesting that all oil-exporting stock markets examined have long-run dependence on oil prices in all periods. Third, the evidence of VAR system show that an increase in OP is associated with a significant increase in the short-run MC for all economies except Israel, Jordan, and Egypt in the three time periods. This result emphasizes that oil-exporting countries are more likely depend on oil price increase in short and long –run periods.

Finally, the IRFs of response of MC to OP shocks confirm positive relationships between OP and MC for the majority of ME stock markets. These research findings will inform policy makers in making decisions about the structure and reform of economies in the ME. For instance, detecting a particular pattern in changes might assist investors in making decisions based on both MC and OP. In addition, the optimal decision for international investors that invest in the ME and seek to maximize the expected profit of their stocks, could be to invest only when OP has positive effect on MC. Therefore, these findings not only will be informative for portfolio investors in building a stock portfolio but also help us understand better the linkages
between OP and MC providing us with a clearer picture in the ME economies. As a result, observing whether the changes in crude oil prices are transmitted to the ME stock markets will reconsider attention from the policy makers of the region and the global investors.

The remainder of this paper is organized as follows. Section 2 surveys relevant literature on oil price and market capitalization. Section 3 provides a theoretical framework for this research. Section 4 is explanation of empirical related to models time series. The next section describes data and presents descriptive statistics. Section 6 presents the results and reports empirical findings in the paper. Section 7 concludes.

2. Literature Review

Oil prices and stock markets have been a subject of economic and financial research in the global economy, especially during the past decade (Jones et al., 2004; Maghyereh, 2006; Basher and Sadorsky, 2006; Yu and Hassan, 2008; Miller and Ratti 2009; Lardic and Mignon, 2008; Filis et al., 2011; Arouri and Roult 2012; Le and Chang, 2015; Ewing and Malik, 2016). Jones et al., (2004) address the theoretical and empirical insights of the macroeconomic consequences of oil price shocks, while other studies examine how oil price shocks affect the performance of stock markets in different economies. Using data from different markets and applying different methodologies, many of these studies conclude that there exists a strong linkage between the oil prices and stock market activities.

The linkage between oil prices and stock markets has also become the subject matter of scholarly studies that address a variety of topics within developed economies. Sadorsky (1999) examines the dynamic interdependence between oil prices and economic activities, including inflation adjusted stock returns in the US. Using monthly data from 1947-1996, he showed that stock returns, represented by S&P 500 index had been affected by oil prices and their volatility. Positive shocks to oil prices were found to have negative effects on stock returns. On the other
hand, no effect of stock market returns on oil prices was detected perhaps because oil prices are
formed by global supply and demand forces.

Using time series models and a VECM approach, Miller and Ratti (2009) examine the
long run relationship between oil prices and stock markets from 1971 to 2008 in six OECD
countries. Their results suggest negative linkages between increasing oil prices and stock market
indices for all six economies. Abhyankar et al. (2013) conducted a study on oil price shocks and
stock market activities for Japan’s economy from 1999 to 2011. They concluded with a VAR
model that there was a negative effect on the Japanese stock market as an oil-importing country.
Additional research has examined the effect of oil price changes on stock market returns in
emerging countries. Basher and Sadorsky (2006) studied the effects of crude oil price changes on
emerging stock market returns in 21 emerging stock markets around the world, including
countries in Middle East, Asia, South America, Europe, and Africa. Both oil-exporting and oil-
importing countries were included in their study. They utilized a multi-factor model that
incorporated both unconditional and conditional risk factors in an approach similar to a capital
asset pricing model. Six models investigated the relationships among stock market returns,
market risk, oil price, exchange rate risk, squared market price risk, and total risk. In general,
they found strong evidence that oil price was found to have a statistically significant positive
effect on stock market returns in most emerging markets.

There is some evidence supporting the link between oil prices and stock markets of the
Gulf Cooperation Council (GCC) within the ME economies. Fayyad and Daly (2011) assessed
the effect of oil price shocks on stock markets in GCC countries compared with the U.S. and
United Kingdom (UK). The time series data applied in this research came from weighted equity
market indices of seven stock markets - Kuwait, Oman, United Arab Emirates (UAE), Bahrain,
and Qatar along with the UK and U.S. They applied a VAR analysis to examine linkages between oil price and stock market returns from September 2005 to February 2010. They found that oil price had statistically significant, positive impacts on stock markets in the GCC economies compared to negative impacts on stock markets in the U.S and UK. These empirical results suggest that the forecasted influence of oil on stock markets rose with an increase in oil prices.

Arouri and Roult (2012) investigate a long-run relationship between oil prices and stock markets in GCC using bootstrap panel cointegration and seemingly unrelated regression (SUR) methods. They use monthly data from four of the six GCC stock markets (Kuwait, Bahrain, Oman, and Saudi Arabia) between January 1996 and December 2007. Their findings showed the existence of cointegration and causality links between oil prices and stock returns. They also concluded that there exists a positive effect of oil price on stock markets in each country except in Saudi Arabia.

Nwosa (2014) studied oil prices and stock market prices in Nigeria using multivariate cointegration test and VECM with data from 1985 to 2010. His results showed that oil prices had a significant connection with stock market indices in the long run. The total impacts of the oil price changes accounted for about 47% of the variation in stock market capitalization in the long run. These findings highlight the key role that oil price shocks have on the volatility of stock markets in several developing economies.

3. Theory

A well-developed stock market plays an important role in the mobilization of capital inflows, which in turn, boosts economic growth. For example, stock returns affect interest rates, market value, industrial production and encomia growth (Sadorsky, 1999). One measure of changes in stock market value is market capitalization, which is calculated by multiplying the
per-share price by the number of outstanding shares. MC provides a measure of the total value for companies and entire economy condition. For companies in a given economy, stock market capitalization correlates with business development, capital accumulation, energy and oil prices, and other variables (Maghyereh, 2006).

In theory, oil price shocks can affect economic activity and economic growth (Brown and Yücel, 2002; Lardic and Mignon, 2008; Ghalayini, 2011; Gadea et.al, 2016). On the one hand, stock market activities and MC are affected through both consumption and investment to extract oil. Consumption of goods and services are influenced indirectly by crude oil price related to income and output (Ewing and Thompson, 2007). OP changes can affect output through production cost effect and then it will impact income (Figure 2-1).

We argue that oil prices have affected income in the world economy in different ways. For a given level of world GDP, for instance, previous studies found that oil prices have a negative impact on income in oil-importing countries and positive effects in oil-exporting economies (Lescaroux and Mignon, 2009; Filis et.al, 2011; Gadea et.al., 2016). Therefore, oil price is an essential factor to determine firm cash flows and stock market performance (Figure 2-1). When the price of oil increases, income transfers take place from oil-importing economies to oil-exporting economies and consumption is reduced in oil-importing economies (Lardic and Mignon, 2008, 2006). On the other hand, crude oil is a significant input in production and GDP (Jimenez-Rodriguez and Sánchez, 2006; Lardic and Mignon, 2006). Increases in the price of oil could decrease the demand for petroleum, reducing productivity of other inputs which force companies to lower output. Therefore, oil price influences economic performance and markets and thus affects stock market activity and market capitalization.

Oil price volatilities lead to changes in stock market returns, while stock market returns
and cash flows cause changes in market capitalization (Arouri and Roult, 2012; Fayyad and Daly, 2011; Abhyankar et al., 2013; Ha and Chang, 2015; Teulon and Guesmi, 2014). Changing oil prices impact the prices of a wide range of commodities, including inputs to industrial (e.g., chemicals and electricity) and transportation sectors. The overall net impact of oil prices on stock market capitalization relies on a complex combination of cash flow impacts and then market capitalization in stock market. As shown in Figure 2-1, oil price changes can effect stock market capitalization through multiple pathways. In this figure, oil-importing countries experience decreased output with from production cost increases due to oil price decreases. Increased crude oil prices also raises the inflation rate via higher prices throughout the economy. The consequences of decreasing output decreases income levels, increases cost of living, and reduces consumption. Additional impacts of increased inflation include higher discount rates in the economy. The end result of these changes in the economy is that firm cash flows will decrease, thereby leading to declining market capitalization.

Figure 2.1: Mechanisms of crude oil price changes on an oil importer stock market, increase (left arrow) and decrease (right arrow)
Thus, more formally expressed, stock market capitalization is strictly a function of discounted cash flows, and discounted cash flows are function of oil prices, macroeconomic activities, and socio-political variables:

\[ MC_t = M(DCF_t) \quad (2-1) \]
\[ DCF_t = O(OP_t, X_t) \quad (2-2) \]

Where, \( MC_t \) represent the value of stock market capitalization, \( DCF_t \) is the discounted sum of expected future cash flows, \( X_t \) represents macroeconomic activities (i.e. interest rate or shale oil production), and \( OP_t \) is oil prices. Oil price fluctuations impact corporate output and earnings, domestic prices, and stock market share prices.

4. Methodology

A variety econometric methods have been utilized in the past decade to examine the impact of oil price shocks on the stock market. In order to comprehensively explore the linkages between OP and MC in ME economies, we use time series methods, including unit root tests, multivariate time series VECM, and VAR analysis. A VAR model is utilized to examine all variables as jointly endogenous and does not inflict restrictions on the structural relationship. VECM is a restricted VAR that has cointegration restrictions. Thus, the VECM specification restricts the long-run trend of the endogenous variables to converge to their cointegrating linkages while allowing a wide range of short-run dynamics.

4.1 Unit Root Testing

The first step of any time series analysis is to examine the order of integration for all variables used. In this analysis, Augmented Dickey Fuller (ADF) (Dickey and Fuller, 1981) and Phillips and Perron (1988) (PP) tests were applied to examine the unit roots and the stationary
properties of oil price and stock market capitalization variables. Equation (3) was used to calculate the test statistics for both the OP and MC variables.

$$\Delta Y_t = \alpha + \beta t + \gamma y_{t-1} + \sum_{i=1}^{p} \delta \Delta y_{t-i} + \epsilon_t$$ (2-3)

In Equation (2-3), $y_t$ stands for either the OP or MC variable, $\alpha$ is the constant term, $\beta$ the coefficient of time trend, $\epsilon_t$ denotes the error term and $p$ is the number of lags determined by the auto regressive process. From this equation, the $H_0$: $\gamma = 0$, i.e., $y$ has a unit root against the alternative hypothesis of $H_1$: $\gamma < 0$, i.e. $y_t$ is stationary. Observations from the majority of economic time series data studies are non-stationary at levels but when first difference is considered, the series become stationary (Engle and Granger, 1987). After confirmation of the stationarity of variables, we examine whether there exists any long run relationship with the Johansen and Juselius (JJ) cointegration test.

In this study, a Zivot-Andrews (ZA) test was utilized. Zivot and Andrews (1992) describe a unit root test for a time series that allows for one structural break in the series. Structural breaks might appear in the trend, intercept, or both. The ZA test detects unknown structural breaks in the time series. A break date is chosen where the t-statistic from the ADF test of a unit root is at a minimum level or negative. Thus, a break date can be selected where the evidence is less suitable for the null of a unit root. According to the process described in Zivot and Andrews (1992), a breakpoint can be endogenously verified.

### 4.2 Johansen’s Cointegration Method

One of the principal objectives in a time series approach is to estimate a long-run equilibrium, using a systems-based method (Johansen, 1988, 1991, 1992, and 1995; Dolado et al., 1990; and Johansen and Juselius, 1990). The Johansen’s estimation method is represented by Equation (2-4).
\[ \Delta y_t = A_0 + \Pi y_{t-1} + \sum_{i=1}^{r} \Gamma_i \Delta y_{t-i} + \varepsilon_t \]  \hspace{1cm} (2-4)

Where \( y_t \) denotes 2x1 vector containing the OP and MC variables. Long-run relationships are captured through \( \Pi \) matrix (\( n \times r \)), in which \( n \) is the number of variables and \( r \) is the number of cointegrating vectors. The equation can be written as \( \Pi = \alpha \beta \), where \( \beta \) is matrix of cointegrating vectors, \( \alpha \) is adjustment coefficients (Johansen, 1991). \( \Gamma_i \) is a matrix with coefficients associated to short-run dynamic effects. The matrix \( \alpha \) consists of error correction coefficients and \( r \) is the number of cointegrating relationships in the variable (\( 0 < r < n \)), which is also known as speed of adjustment parameter. \( \beta \) is a matrix of \( r \) cointegrating vectors, which represent the long run relationship between variables. The optimum lag is selected on the basis of Akaike Information Criteria (AIC) (Akaike, 1974). In general, VAR estimates discussed in the next sub-section are sensitive to the number of lags included.

The cointegration rank is tested using maximum eigenvalue and trace statistics proposed by Johansen (1988). The long run information of the series is taken into account when analyzing short run change and the resulting model is a short run error correction model. The rank of a matrix is equal to the number of its characteristic roots that differ from zero. The test for the number of characteristics roots that are not significantly different from unity can be conducted using the following test statistics:

\[
\lambda_{trace} = -T \sum_{i=r+1}^{K} \ln(1 - \lambda_i^0) \hspace{1cm} (2-5)
\]

\[
\lambda_{max} = -T \ln(1 - \lambda_{r+1}^0) \hspace{1cm} (2-6)
\]

Here \( \lambda_i^0 \) represents the estimated values of the characteristics roots (called eigenvalues) obtained from the estimated \( \Pi \) matrix, and \( T \) is the number of usable observations. The first, called the trace test, tests the hypothesis that there are at most \( r \) cointegrating vectors. In this test, the trace statistic has a nonstandard distribution because the null hypothesis places restrictions on the
coefficients on \( y_{t-1} \), which is assumed to have \( K - r \) random-walk components. The farther the estimated characteristic roots are from zero, the more negative is \( \ln (1 - \lambda_i^\wedge) \) and the larger is the \( \lambda_{\text{trace}} \) statistic. The testing sequence terminates and the corresponding cointegrating rank in the null hypothesis is selected when the null hypothesis cannot be rejected for the first time. In case the first null hypothesis in the sequence cannot be rejected, then it means that there is no cointegrating relationship involving the \( K I(1) \) variables, and hence a VAR process in first difference is then considered for studying relationships involving the \( K \) variables.

The second test, called the maximum eigenvalue test, utilizes the hypothesis that there are \( r \) cointegrating vectors versus the hypothesis that there are \( r+1 \) cointegrating vectors. This means if the value of characteristic root is close to zero, then the \( \lambda_{\text{max}} \) will be small. Vecrank (Max rank) is the command for determining the number of cointegrating equations. (Osterwald-Lenum, 1992; Johansen, 1995). If Max rank is zero and no cointegration, we conduct VAR otherwise VECM will be used.

### 4.3 VAR Model

A VAR model of order \( p \) that includes \( k \) variables can be expressed as:

\[
y_t = A_0 + \sum_{i=1}^{p} A_i y_{t-i} + \epsilon_t, \quad (2-7)
\]

Where \( p \) is the number of lags, \( y_t = [y_{1t} \ldots y_{kt}]' \) is a column vector of all the variables in the model (short-run oil price and market cap); \( A_0 \) is a column vector of constant terms; \( A_i \) is \( k \times k \) matrix of unknown coefficients; and \( \epsilon_t \) is a column of errors with the following properties:

\[
E(\epsilon_t) = 0 \quad \forall t,
\]

\[
E(\epsilon_t \epsilon'_t) = \Omega \quad \text{if } s = t, \quad (2-8)
\]

\[
E(\epsilon_t \epsilon'_t) = 0 \quad \text{if } s \neq t,
\]

where \( \Omega \) is the variance–covariance matrix with non-zero off-diagonal elements. If the variables
considered in this study are not cointegrated, then the unrestricted VAR is used. If the variables are cointegrated and I(1), a vector error correction model can be employed. This method of analysis permits us to test for the direction of causality. Cointegration analysis is conducted to shed light on the long-run relations that may exist among the series of nine ME economies. Following Johansen (1995), Harris and Sollis (2003), Lutkepohl (2005), Miller and Ratti (2009), Greene (2007), and Nwosa (2014), we estimated VECM for series cointegrated with OP and VAR for series which are not cointegrated with OP.

To investigate whether oil price shocks influence ME stock market capitalization, we utilized a restricted VAR which is VECM and unrestricted VAR with a linear specification of oil price and stock market capitalization. The error terms were assumed to be zero-mean independent white noise processes. Based on related literature (Cologni and Manera, 2008; Apergis and Miller 2009; Filis and Chatziantoniou, 2014) we constructed a VAR model to examine the influence of oil prices on market capitalization.

\[
MC_{it} = \varphi + \sum_{i=1}^{p} \alpha_i MC_{t-i} + \sum_{i=1}^{p} \beta_i OP_{t-i} + \epsilon_t \quad (2-9)
\]

Where \(MC_{it}\) is the market capitalization for country \(i\) at time \(t\) and \(OP_t\) is the average oil price at time \(t\). \(\varphi\) is constant and \(a_i, \beta_i, i = 1, \ldots, p\) denote the linear relationship between variables \(MC_t\) and \(OP_t\).

We lagged the explanatory variables to account for past value influences on the MC variable and to determine whether the dependent variable is predictable or not. Indeed, the impact of oil prices on MC might not happen immediately, in which case, lagged explanatory variables are appropriate. To determine the optimal length of oil price impact, it is important to take into account the market capitalization of the previous year. The optimal length of impact
was determined by using the AIC. This criterion determines the maximized log-likelihood of the model where \( k \) is the number of parameters estimated and the model with the lowest AIC is utilized.

The models of VAR and VECM are able to capture the dynamics of the interrelationships between the variables. Impulse response functions, or IRFs, measure the effects of oil price shocks to stock market cap variable. The IRF examines the response while the model shocked by a one-standard-deviation. Further, the models are able to capture the dynamics of the interrelationships between the variables. The IRFs represent the dynamic response path of a variable due to a one-period standard deviation shock on another variable. Based on these models, the impact of oil price movement on stock market capitalization is examined using generalized impulse response functions (IRFs).

5. Data and Descriptive Statistics

In this section, we describe the variables and sample periods, and provide an economic overview of the sample countries. Nine countries in the ME considered for the study are United Arab Emirates (ARE), Bahrain (BHR), Egypt (EGY), Israel (ISR), Jordan (JOR), Kuwait (KWT), Qatar (QAT), Saudi Arabia (SAU), and Turkey (TUR). Among these nine countries, Israel, Jordan, Egypt, and Turkey are net importers, while the remaining countries are oil exporters. Table 2-1 presents selected macroeconomic indicators for the ME economies. The percentage of stock market capitalization to GDP ranges from 16.7% in Egypt to 85.4% in Kuwait, which shows variation between countries in terms of market capitalization. Saudi Arabia is the top exporter of petroleum products in the world, United Arab Emirates ranked in the 6th position, Kuwait ranked in the 13th, and Qatar ranked in the 14th position. In each of these countries, energy exports contributed the vast majority of total exports (TE), ranging from
87% to 94%. Industrial and service sectors contributed more than 90% of the GDP for every country in the ME region except for Egypt. Finally, oil rents contribute a significant portion of GDP for some ME countries like Kuwait (53%) and Saudi Arabia (38.7%), while for others (like Israel, Jordan, and Turkey), these percentages are zero or close to it.

To analyze the impact of OP on nine ME stock markets, we used daily data. Instead of analyzing stock returns, we focus on the MC within each market, where MC is defined as the total tradable value of the number of outstanding shares times the share price. Based upon results from the Zivot-Andrews test, attributed to expanded shale oil production, and the global financial crisis, it was determined to split the data between two periods. Thus, while the dataset covers the period from 2001 to 2015, and analyses were conducted for two sub-periods of 2001-2008 and 2009-2015.

Historical data on outstanding shares and the share price for the selected ME stock markets were obtained Compustat Global, through the Wharton Research Data Services (WRDS). We calculated MC for each company and merged for all companies for each country. Data were expressed in monetary units for local currencies and converted to U.S. dollars. The data files were sorted by Global Company Key (GVKEY), which is a unique six-digit number represented for each company (issue, currency, and index) in the Capital IQ Compustat Database. Given that there is a large amount of missing data before 2001, we selected daily data after January 1\textsuperscript{st}, 2001.

Oil price data were obtained from the U.S. Energy Information Administration (EIA), using the Brent oil spot price, the benchmark price in the oil market measured in U.S. dollars per

---

\textsuperscript{2} The raw data includes trading shares and prices of shares or closed day value of each share for every company in the nine ME stock markets. The share price of each company has a specific code. Outstanding shares are traded not always on a given day for each company in the ME stock markets. GVKEY is a unique six-digit number key assigned to each company (issue, currency, and index) in stock market within each country. Market capitalization values were created for each country by aggregating over GVKEYs.
barrel. Oil prices show fluctuating patterns occur over time, particularly after 2008. Some events creating these fluctuations include: Venezuela cutting off oil sales to ExxonMobil in February 2008, production from Iraqi oil fields not recovering from wartime damage, the US shale gas and oil production boom, and international sanctions against Iran’s oil-dominated economy. These events led rising oil prices from 2004 until the financial crisis in 2008 driven primarily by surging demand and supply shocks in the oil market (Smith, 2009). Then in 2009, oil prices fell sharply due to a surplus of oil production and the US tight oil production. OPEC, the largest player in the oil market in 2009, failed to reach an agreement on production curbs, resulting in plummeting prices.

The two main variables are analyzed using natural logarithms for several reasons: (1) coefficients for variables on a natural-log scale are directly interpretable as proportional differences, and (2) following previous studies (Maghayereh and Al-Kandari, 2007; Arouri et. al, 2012, Filis and, Chatziantoniou, 2014), we utilize logged data to examine the impacts of OP on the MC. (Maghayereh and Al-Kandari, 2007; Arouri et. al, 2012). Thus, logarithmically transforming variables is a proper way to handle stock and oil market where there exists a non-linear relationship exists between the oil prices and stock market capitalization. In addition, logarithmic forms are a suitable way to transfer skewed variables to a log-normal distribution form. Results of skewness in Table 2-2 show that MC and OP are skewed and have asymmetric distributions.

Table 2-2 presents the descriptive statistics related to OP and MC. Nominal mean oil price between 2001 and 2015 was $67.14 USD per barrel, the minimum and maximum varied between $17.5 USD and $145.4 USD during the study period. The degree of correlation between MC and daily OP among the ME countries was strongly positive with all correlation coefficients
above 0.70. Saudi Arabia ranked highest in terms of mean MC at $311 billion USD, followed by UAE at $121 billion USD, and Kuwait with $111 billion USD.

Figure 2.2 presents charts of oil prices and market capitalization by country. Oil prices show fluctuating patterns over time and are particularly volatile after 2008. Some events creating these fluctuations include: Venezuela cutting off oil sales to ExxonMobil in February 2008, production from Iraqi oil fields not recovering from wartime damage, the U.S. shale gas and oil production boom starting around 2008, and international sanctions against Iran’s oil-dominated economy. Oil prices rose consistently from 2004 until the financial crisis in 2008 driven primarily by surging demand and supply shocks in the oil market (Smith, 2009). Then in 2009, oil prices fell sharply due to a surplus of oil production in the U.S. and around the world. OPEC, the largest player in the oil market in 2009, failed to reach an agreement on production curbs, resulting in plummeting prices. For market capitalization, these data show dramatic changes from 2001 through 2005 for the stock markets in SAU, ARE, and BHR. While showing a general upward trend, the other six countries don’t show such large changes in MC during the early time periods.

6. Results and Discussion

For each selected country, we tested for unit roots to check for stationary in OP and MC variables. We estimated ADF and PP unit root tests for each of the variables and their first differences. These tests were based on null hypotheses of unit root for all variables. Values from ADF and PP show that MC and OP variables were non-stationary at level, since the null hypothesis of unit root could not be rejected for most of the countries, particularly for OP (Table 2-3). Therefore, the first difference of each series was taken, and both the ADF and the PP test statistics were once again computed with these first difference values. The null hypothesis of the
unit root test was rejected when their first differences were considered. Based on Table 2-3 results, it was concluded that both OP and MC are integrated by an order of one, I(1), i.e. stationary at first difference for all nine economies, and therefore suitable to investigate the presence of a long-run relationship between these two variables.

Moreover, the Zivot-Andrews (ZA) test was added to ADF and PP tests for stationarity in order to take into consideration structural breaks that can occur in the intercept and/or the trend of a data series (Zivot and Andrews, 1992). The results of the ZA unit root test with trend are summarized in Table 2-3. Crude oil prices were non-stationary and the MC variables were non-stationary in every country except for Bahrain, Jordan, and Qatar - which were stationary in level. However, both variables were stationary in first differences for all nine countries.

By country, the break dates from the ZA tests were consistent between series. The ZA test results of oil price show that the day of the structural break was Oct 22, 2007. However, the date of December 31, 2008 was utilized as the date for splitting the OP and MC data in order to correspond with the financial crisis an in 2008 and shale oil revolution. Thus, the influence of shale oil production and global financial crisis conditions are measured by dividing the sample period into two balanced sub-periods: pre-shale (20001-2008) and post-shale (2009-2015). For each period, the lag length is selected according to the Akaike information criterion (AIC). Table 2-4 shows that the optimum lengths of lags for the two variables for each country in different time periods (2001-2008, 2009-2015, and 2001-2015): United Arab Emirates (6,3,6), Bahrain (8,6,8), Egypt (7,7,8), Israel (8,5,6), Jordan (6,7,6), Kuwait (8,8,8), Qatar (5,6,6), Saudi Arabia (6,3,6), and Turkey (6,4,7).

After confirming that all the variables contain a unit root, we examine whether cointegration changed significantly across the three periods. This stage involves testing for the
existence of a long-run equilibrium linkage between OP and MC using Johansen’s cointegration test. Results for both the trace and maximum eigenvalue tests are reported in Table 2-5. Based on the cointegration test, long-run linkages for MC and OP differ for in all nine countries during the three periods: 2001–2008, 2009-2015, and 2001–2015. The value of the $\lambda_{\text{trace}}$ test statistic and eigenvalue, $\lambda_{\text{max}}$, in the period 2001-2008 show that there exists cointegration for BHR, KWT, SAU, and TUR. Hence, the null hypotheses of no-cointegration ($r = 0$) are rejected in favor of a co-integrating relationship between the OP and MC variables. Conversely, there are no-cointegration and long-run linkages between MC and OP for ARE, EGY, JOR, ISR, and QAT in that time period. During the 2009-2015, there are cointegration relationships in ME stock markets of ISR, JOR, and QAT, and no cointegration in the remaining countries. The existence of a cointegrating relationship among OP and MC suggests that there is causality in at least one direction, however, it does not determine the direction of causality between the variables.

Bahrain, Saudi Arabia, Kuwait, and Turkey have cointegration relationships in the period of 2001-2008. After the global financial crisis, our results indicate that only ISR, JOR and QAT have cointegration relationship. We can argue that the global financial crisis affected all oil-exporting countries except Qatar during the second period of the study. These results are possibly related to reforms in the Qatar economy and financial markets since the Qatar government has been making efforts to diversify its economy away from crude oil income.

In 2008, natural gas overtook crude oil as the largest contributor to the economy of Qatar. Crude oil and natural gas together accounted for 46.2 percent of the overall GDP of Qatar (2009), for the first time overtaken by non-oil and gas sectors in the Qatari Markets. Performance Indicators of the Financial Sector didn’t change in Qatar prior to the end of 2008. It also resulted in an accumulation of non-oil sectors and foreign assets. In fact, Qatar's stock market has not
been affected as much as other oil-producing states (Abdelbaki, 2010). Most oil-exporting countries in the ME, particularly Qatar, experienced significant increases in banking system credit to the private sector and external funding for the banking system was strongly affected by oil price changes, except for Qatar (Khamis, et al., 2010). As a result of the financial crisis and decreases of crude oil prices did not changed Qatar’s stock market value.

Overall in the total period 2001-2015, the results suggest that the value of $\lambda_{\text{trace}}$ test statistic and eigenvalue, $\lambda_{\text{max}}$, under the null hypothesis of no cointegration ($r = 0$) rejects the null hypothesis and show the existence of cointegration relationships among the variables at a 5% level of significance in seven of nine countries. There is no cointegration for EGY and JOR. If there is no cointegration for OP and MC variables in a country, we estimate VAR. Otherwise we utilize the VECM which is restricted VAR.

Results from VECM model suggest that the existence of a long-run association between the variables of OP and MC. All ME economies with cointegrated data showed statistically significant error correction terms (Tables 2-6, 2-7, and 2-8). Based upon cointegrating tests, VECM, and ECM results in Table 2-6, there exists long-run relationships of OP on MC for BHR, KWT, SAU, and TUR during the period of 2001-2008. In addition, the findings of VECM, cointegration test, and ECM show that MC and OP have long-run associations for JOR, QAT and ISR in the 2009-2015 period (Table 2-7). Moreover, the results of VECM show that during the period of 2001-2015 there are long-run relationships between OP and MC for seven ME economies except EGY and JOR (Table 2-8). The findings of VECM, also, indicate that MC and OP have long-run linkages for only three countries in the pre-shale and four countries in the post-shale period.

The VAR models in Table 2-9 and cointegration tests show that there exists short-run
causality running from OP to MC for ARE, EGY, ISR, JOR, and QAT in the 2001-2008 period. Moreover, cointegration tests and VAR results in Table 2-10 provide evidence that there exists short-run linkages for ARE, BHR, EGY, KWT, and SAU, but not TUR in the 2009-2015. In comparison, there are five economies with short-run and four countries with long-run linkages between OP and MC in the 2001-2008. In the 2001-2015 period, EGY and JOR are the only countries where OP and MC were not cointegrated. Thus, we applied the unrestricted VAR model (Table 2-11). The results of the unrestricted VAR model provide evidence that there is a short-run causality running from OP to MC in these two economies.

The ME stock markets consist of a variety of business and economic sectors. Firms and stock markets in oil exporting have more reliance on oil income. While the stock market of United Arab Emirates has been diversifying in different categories including financial and banking, insurance, industrial sector, and energy companies, its economy remains extremely reliant on crude oil as more than 85% of the United Arab Emirates’ economy was based on the oil exports in 2009. In Bahrain, listed companies in stock market are mostly in banking, insurance, and financial sector. As an oil-exporter, Bahrain, depends upon oil revenue and shareholders’ investment relies on money and salaries from oil prices. In the Saudi Arabia stock market, listed companies include energy, materials, capital goods, transportation, commercial and professional services, insurance, banking and health care. Likewise the stock markets in Kuwait and Qatar consist of energy, industrial, financial, insurance and commercial services. In the Egyptian stock market, there are more diversified business and companies in industrial, energy, health care, financial, tourism and hotels. The results imply that most oil-exporting countries have affected by changing crude oil prices since the companies and shareholders are more influenced by crude oil price changes when compared to oil-importing countries in the ME.
Oil-importing countries such as Israel, Jordan, and Turkey have more non-petroleum companies and less dependency on oil income in the stock market. In Israel, its stock market is based on high tech companies, energy sector, real estate and industrial companies. Jordan’s stock market is based upon the performance 100 of the largest companies of financial, industrial, services and tourism sectors. These sectors represent around 90 percent of the aggregate market capitalization of the listed companies in Jordan. The Turkish stock market consists more of companies from industrial and manufacturing, agricultural, real estate, and commercial services.

Finally, we examine the effect of OP shocks on MC by examining the impulse response functions (IRFs). The IRFs map out the response of a one standard deviation shock of OP or MC to the shock in the error term in VAR system which include unrestricted VAR and VECM over several time periods. An IRF measures the effect of a shock to an endogenous variable on itself or on another exogenous variable (Lutkepohl, 2005). Based on these models, the impact of OP movement on MC is examined using cumulative IRFs. Hence, the IRF of the VAR and VEC models enable us to examine how each of the variable responds to innovations from other variables in the system. The orthogonal innovations, denoted by $\varepsilon_t$, are obtained by the errors terms in Equation (2-7) which indicate that error values are uncorrelated with each other. Results for nine countries and 95% confidence bounds around each orthogonalized impulse response appear in Figure 2-2.

We make three observations. First, we find a significant change in how MC reacts to shocks in OP changes before versus after 2008. For the period of 2001-2008, the response of MC to a shock in OP seems to be small and mainly increasing over time. Second, the IRF results in the 2009-2015 period show that OP has negative impact on MC in ME economies except for EGY, JOR, and KWT. These negative effects are possibly due to some combination of the global
financial crisis and expanded shale oil production, both of which seem to have affected the ME economies. Third, the IRFs show that the responses of all nine stock markets to one standard deviation crude oil price shock displays fluctuations with both positive and negative effects.

7. Conclusions

In this paper, we analyze whether or not crude oil price changes have dynamic long-run associations with stock market capitalization in nine ME economies using daily data for the period 2001–2015 by means of estimating multivariate VECM and VAR models. In order to estimate the effect of OP shocks on MC, we include cointegration tests and IRF’s. The inclusion of these tools help us understand the linkages between OP and MC in the ME region.

The empirical analyses provide some important characteristics of OP and MC for nine ME economies in three different periods (Table 2-12). Among the nine countries considered in this study, OP and MC were cointegrated in each country except for EGY and JOR during the entire time period of 2001-2015. This suggests that a dynamic long-run causality linkage exists between the two variables of OP and MC. Oil exporting countries (ARE, BHR, KWT, QAT, and SAU) had long-run relationships while oil importing countries (EGY, ISR, JOR, and TUR) had short-run relationships. The findings show that OP has statistically significant, short and long run effects on MC across all nine economies. These results suggest that stock market capitalization have reacted to the changes of crude oil prices in the oil-exporting which have more dependency on oil prices and income from oil.

Prior to the financial crisis and the shale oil revolution (2001-2008), the results show that there existed long-run causality between OP on MC for BHR, KWT, SAU, and TUR. During this time period, VAR models show that there exists short-run causality running from OP to MC for ARE, EGY, ISR, JOR, and QAT in the 2001-2008 period. The empirical results show that before the global financial crisis and expanded shale oil production, crude oil prices increased
sharply during the 2001-2008 period. In other words, MC reacts to changes in OP over a short time period (i.e. days). These results suggest that the OP affects MC in day(s) and the linkage between variables is predictable within short time horizons.

For the time period 2009-2015, a summary of research results show that short-run causality between OP and MC existed in six countries, with exceptions being ISR, JOR, and QAT (Table 2-12). These results provide evidence that OP decreases have not affected MC in half of ME oil-importing stock markets after global financial crisis and shale oil revolution. We can argue that these economies are not as reliant on oil prices as much as the oil-exporting economies in the ME. During the second period, all oil-exporting markets had short run linkages except QAT. These results suggest that the global financial crisis and oil shale production have affected most of oil-exporting economies in the ME region.

As a final observation from Table 2-12, only the stock market in Egypt showed the same results for OP and MC in terms of no cointegration and a short-run relationship throughout all three time periods. Every other country switched between short and long-run from 2001-2008 to 2009-2015, with the exception of ARE, which had a consistent short-run relationship for both sub-periods, but a long-run relationship over the entire time period. This result may indicate that Egypt has had the most stable stock market among ME countries during 2001 to 2015.

As a result, we can argue that there exist a significant inter-relationship between most of the ME economies. The empirical results obtained in this study are consistent with previous studies. Similarities between our results and previous findings are: (1) there is evidence for cointegration between oil prices and stock markets in the ME countries (Fayyad and Daly 2011); (2) there exist not only long-run relationships between oil prices and stock markets for oil-importing and oil-exporting economies but also a positive causality between two variables in
majority of the ME stock markets (Filis et al., 2011; Arouri and Roult, 2012); (3) crude oil prices could predict stock market capitalization in the ME economies; (4) the IRFs of a response of MC to OP shocks verify the positive relationship between oil prices and stock market value for most of ME markets; and (5) stock markets in both oil-importing and oil-exporting countries tend to react to OP shocks either positive or negative (Yu et al., 2008; Apergis and Miller, 2009; Filis et al., 2011). These common findings prove a significant influence of OP on stock markets in the region.

In line with expectation, most oil-exporting economies have positive short-run impacts of OP on stock MC. Among oil importing economies, OP movement affects negatively on MC for Israel. Most of oil exporting economies considered in this study are heavily dependent on oil price movements in particular in full sample. Increase in oil price strengthens their economic growth, which in turn transmitted to higher MC. These results suggest that a dynamic long run causality among oil price and stock market capitalization for seven ME economies in total period (Table 2-8).

This similarity of the findings for majority of ME economies can be explained by several facts. First, since the ME countries having common land borders with each other, they have close economic relationships and increasing crude oil prices have impacts on the most economies through increasing exports of goods and services from oil-importing to oil-exporting economies. Thus, a higher crude oil price has benefits for both oil-importing and oil-exporting countries in the region. Since market capitalization determines the dynamic impact of oil price changes, it is argued that higher oil prices are needed to stimulate stock markets and economic performance in the ME. However, consequences of this dependency in particular for oil-exporting have had economic instability. Since majority of oil-exporting rely on oil income, oil
price shocks have affected their economies negatively. Therefore, a high priority for policy
makers in ME economies is to diversify their macro economies by improving the shares of the
non-oil industry and additional sectors to their economies.

There are several limitations worth noting in this study. First, there is lack of available
and accurate data for all ME economies due to political instability in majority of countries
comparison to developed economies plus most of ME economies have thin markets for stocks
(Sowers et al., 2011). While there are 17 counties in this region, data could be obtained from
only nine economies. In addition, our research lacks considerations of how other
macroeconomic factors, such as interest and exchange rates as well as future expectations for
cash flows, may impact MC. Future research could explore the influences of macroeconomic
activities and other factors contributing to the effect of oil prices on stock market activities in the
Middle Eastern economies such as interest rates, shale oil production, currency rates, renewable
energy, political instability, and rates of conflict.
8. References


Tables and Figures

Table 2-1: Macroeconomic overview of ME region

<table>
<thead>
<tr>
<th>Stock market capitalization of listed domestic companies (% of GDP)</th>
<th>ARE</th>
<th>BHR</th>
<th>EGY</th>
<th>ISR</th>
<th>JOR</th>
<th>KWT</th>
<th>QAT</th>
<th>SAU</th>
<th>TUR</th>
</tr>
</thead>
<tbody>
<tr>
<td>52.9</td>
<td>59.7</td>
<td>16.7</td>
<td>82.4</td>
<td>67.8</td>
<td>56.7</td>
<td>85.4</td>
<td>65.2</td>
<td>26.3</td>
<td></td>
</tr>
</tbody>
</table>

| Energy exports (% of TE) | 90 | 45 | 24 | 1 | 62 | 94 | 88 | 87 | 4 |
| Oil exports global rank | 6 | 33 | 40 | 0 | 0 | 13 | 14 | 1 | 0 |

<table>
<thead>
<tr>
<th>Composition% of GDP:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agriculture</td>
</tr>
<tr>
<td>Industry</td>
</tr>
<tr>
<td>Service</td>
</tr>
</tbody>
</table>

| GDP per Capita (USD) | 42,000 | 21,200 | 6,000 | 31,000 | 5,300 | 40,700 | 102,900 | 20,400 | 16,067 |
| Oil rents as a (% of GDP) | 21.6 | 15.3 | 5.8 | 0 | 0 | 53 | 19.5 | 38.7 | 0.2 |


Table 2-2: Descriptive statistics for MC (billion USD) and for OP (USD per barrel)

<table>
<thead>
<tr>
<th>MC</th>
<th>ARE</th>
<th>BHR</th>
<th>EGY</th>
<th>ISR</th>
<th>JOR</th>
<th>KWT</th>
<th>QAT</th>
<th>SAU</th>
<th>TUR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>121</td>
<td>22.9</td>
<td>6.1</td>
<td>132</td>
<td>24.5</td>
<td>111</td>
<td>107</td>
<td>311</td>
<td>177</td>
</tr>
<tr>
<td>Min</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3.74</td>
<td>.0255</td>
<td>2.76</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Max</td>
<td>273</td>
<td>55.3</td>
<td>175</td>
<td>236</td>
<td>58.6</td>
<td>225</td>
<td>231</td>
<td>796</td>
<td>376</td>
</tr>
<tr>
<td>SD</td>
<td>75.4</td>
<td>9.11</td>
<td>40.5</td>
<td>57.5</td>
<td>11.5</td>
<td>45.2</td>
<td>54.6</td>
<td>176</td>
<td>103</td>
</tr>
<tr>
<td>Skewness</td>
<td>0.089</td>
<td>0.29</td>
<td>-0.109</td>
<td>-0.12</td>
<td>-0.21</td>
<td>-0.078</td>
<td>-0.03</td>
<td>-0.17</td>
<td>0.011</td>
</tr>
<tr>
<td>Correlation with oil</td>
<td>0.72</td>
<td>0.7194</td>
<td>0.84</td>
<td>0.84</td>
<td>0.76</td>
<td>0.76</td>
<td>0.701</td>
<td>0.72</td>
<td>0.83</td>
</tr>
</tbody>
</table>

| Observations | 3580 | 3153 | 3673 | 3674 | 3389 | 3140 | 2919 | 3808 | 3923 |

Oil price | 67.14 | 17.5 | 27.98 | 0.012 | - |
Table 2-3: Tests for unit root hypothesis for market capitalization and oil price

<table>
<thead>
<tr>
<th>Country</th>
<th>Variables</th>
<th>Variables in level</th>
<th>Variables in first difference</th>
<th>Minimum t-statistics&lt;sup&gt;1&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>ADF</td>
<td>PP</td>
<td>ADF</td>
</tr>
<tr>
<td>ARE</td>
<td>LMC</td>
<td>-5.134***</td>
<td>-4.779***</td>
<td>-64.889***</td>
</tr>
<tr>
<td>BHR</td>
<td>LMC</td>
<td>-10.627***</td>
<td>-5.504***</td>
<td>-70.095***</td>
</tr>
<tr>
<td></td>
<td>LOP</td>
<td>-1.414</td>
<td>-1.874</td>
<td>-97.672***</td>
</tr>
<tr>
<td>EGY</td>
<td>LMC</td>
<td>-2.005</td>
<td>-1.395</td>
<td>-35.094***</td>
</tr>
<tr>
<td></td>
<td>LOP</td>
<td>-1.916</td>
<td>-1.844</td>
<td>-64.657***</td>
</tr>
<tr>
<td>ISR</td>
<td>LMC</td>
<td>-2.005</td>
<td>-1.395</td>
<td>-35.094***</td>
</tr>
<tr>
<td></td>
<td>LOP</td>
<td>-1.916</td>
<td>-1.844</td>
<td>-64.657***</td>
</tr>
<tr>
<td>JOR</td>
<td>LMC</td>
<td>-2.561</td>
<td>-2.419</td>
<td>-55.905***</td>
</tr>
<tr>
<td></td>
<td>LOP</td>
<td>-1.983</td>
<td>-1.906</td>
<td>-62.182***</td>
</tr>
<tr>
<td></td>
<td>LOP</td>
<td>-2.055</td>
<td>-1.989</td>
<td>-56.156***</td>
</tr>
<tr>
<td>QAT</td>
<td>LMC</td>
<td>-5.380***</td>
<td>-4.971***</td>
<td>-55.234***</td>
</tr>
<tr>
<td></td>
<td>LOP</td>
<td>-2.317</td>
<td>-2.252</td>
<td>-57.234***</td>
</tr>
<tr>
<td>SAU</td>
<td>LMC</td>
<td>-1.712</td>
<td>-4.695***</td>
<td>-60.218***</td>
</tr>
<tr>
<td></td>
<td>LOP</td>
<td>-3.842***</td>
<td>-3.796***</td>
<td>-63.890***</td>
</tr>
<tr>
<td>TUR</td>
<td>LMC</td>
<td>-1.648</td>
<td>-1.140</td>
<td>-78.339***</td>
</tr>
<tr>
<td></td>
<td>LOP</td>
<td>-1.892</td>
<td>-1.817</td>
<td>-64.755***</td>
</tr>
</tbody>
</table>

Notes: All variables are measured as natural logs and *** p<0.01 and z(t)= -3.430, ** p<0.05 and z(t)= -2.860, p<0.1 and * z(t)= -2.570. Critical values for ZA test: 1%(−4.93), 5%(−4.42), and 10%(−4.11). Based on Zivot-Andrews unit test root tests for market capitalization and oil price in all economies except BHR, JOR, and QAT are non-stationary in level, while all countries are stationary in the first difference at 5% significance.

<sup>1</sup> The oil price break date is Oct 22, 2007.
Table 2-4: Selection of optimum lag according AIC and BIC criteria

<table>
<thead>
<tr>
<th>Country</th>
<th>ARE</th>
<th>BHR</th>
<th>EGY</th>
<th>ISR</th>
<th>JOR</th>
<th>KWT</th>
<th>QAT</th>
<th>SAU</th>
<th>TUR</th>
</tr>
</thead>
<tbody>
<tr>
<td>2001-2008</td>
<td>6</td>
<td>8</td>
<td>7</td>
<td>8</td>
<td>6</td>
<td>8</td>
<td>5</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>2009-2015</td>
<td>3</td>
<td>6</td>
<td>7</td>
<td>5</td>
<td>7</td>
<td>8</td>
<td>6</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>2001-2015</td>
<td>6</td>
<td>8</td>
<td>8</td>
<td>6</td>
<td>6</td>
<td>8</td>
<td>6</td>
<td>6</td>
<td>7</td>
</tr>
</tbody>
</table>

Total Obs. 3580 3153 3673 3674 3389 3140 2919 3808 3923

Note: Akaike information criterion (AIC), the Bayesian information criterion (BIC).

Table 2-5: Tests of cointegration for MC and OP

<table>
<thead>
<tr>
<th>Country</th>
<th>CT</th>
<th>ARE</th>
<th>BHR</th>
<th>EGY</th>
<th>ISR</th>
<th>JOR</th>
<th>KWT</th>
<th>QAT</th>
<th>SAU</th>
<th>TUR</th>
</tr>
</thead>
<tbody>
<tr>
<td>2001-2008</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\lambda_{trace}$</td>
<td>17.94**</td>
<td>3.35**</td>
<td>6.98**</td>
<td>9.04**</td>
<td>10.86**</td>
<td>21.23</td>
<td>10.438</td>
<td>14.77**</td>
<td>0.88**</td>
<td></td>
</tr>
<tr>
<td>$\lambda_{max}$</td>
<td>13.2*</td>
<td>3.35**</td>
<td>6.82**</td>
<td>8.60**</td>
<td>7.94**</td>
<td>14.69</td>
<td>7.84</td>
<td>9.87**</td>
<td>0.88**</td>
<td></td>
</tr>
<tr>
<td>Max.Rank</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>2009-2015</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\lambda_{trace}$</td>
<td>10.132</td>
<td>13.46**</td>
<td>14.7**</td>
<td>6.49</td>
<td>0.00</td>
<td>13.36</td>
<td>1.89**</td>
<td>12.5**</td>
<td>13.54**</td>
<td></td>
</tr>
<tr>
<td>$\lambda_{max}$</td>
<td>8.601</td>
<td>10.46**</td>
<td>12.14</td>
<td>12.06</td>
<td>0.00</td>
<td>10.67</td>
<td>1.89</td>
<td>9.58**</td>
<td>11.71**</td>
<td></td>
</tr>
<tr>
<td>Max.Rank</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>2001-2015</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\lambda_{trace}$</td>
<td>4.74</td>
<td>5.57**</td>
<td>8.268**</td>
<td>1.54**</td>
<td>13.08**</td>
<td>4.77</td>
<td>2.65**</td>
<td>6.79**</td>
<td>1.89**</td>
<td></td>
</tr>
<tr>
<td>$\lambda_{max}$</td>
<td>4.74</td>
<td>5.57**</td>
<td>6.72**</td>
<td>1.54**</td>
<td>8.97**</td>
<td>4.77</td>
<td>2.65</td>
<td>6.79**</td>
<td>1.89**</td>
<td></td>
</tr>
<tr>
<td>Max.Rank</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>

Total Obs. 3580 3153 3673 3674 3389 3140 2919 3808 3923

Note: All variables are measured as natural logs and ** p<0.05. Max rank of r=0 represents that there is no cointegration and max rank of r>=1 shows that there exists one and more cointegrations. CT represents the cointegration tests.
Table 2-6: The result of VECM for MC and OP, 2001-2008

<table>
<thead>
<tr>
<th>Country</th>
<th>BHR</th>
<th>KWT</th>
<th>SAU</th>
<th>TUR</th>
</tr>
</thead>
<tbody>
<tr>
<td>ECM(-1)</td>
<td>-0.031**(.004)</td>
<td>-0.003**(.00)</td>
<td>-0.003**(.001)</td>
<td>-0.02**(.003)</td>
</tr>
<tr>
<td>ΔInMC(-1)</td>
<td>-0.22**(.025)</td>
<td>-0.075**(.03)</td>
<td>0.014(.02)</td>
<td>-0.235**(.022)</td>
</tr>
<tr>
<td>ΔInMC(-2)</td>
<td>-0.054(.026)</td>
<td>0.035(.03)</td>
<td>0.008(.02)</td>
<td>-0.128**(.023)</td>
</tr>
<tr>
<td>ΔInMC(-3)</td>
<td>-0.078**(.025)</td>
<td>-0.021(.03)</td>
<td>0.007(.02)</td>
<td>-0.025(.023)</td>
</tr>
<tr>
<td>ΔInMC(-4)</td>
<td>-0.04246(.025)</td>
<td>0.023(.03)</td>
<td>0.005(.02)</td>
<td>-0.041(.023)</td>
</tr>
<tr>
<td>ΔInMC(-5)</td>
<td>0.21**(0.03)</td>
<td>0.036(.03)</td>
<td>-0.005(.02)</td>
<td>-0.073**(.023)</td>
</tr>
<tr>
<td>ΔInMC(-6)</td>
<td>-0.0705**(.026)</td>
<td>-0.043(.03)</td>
<td>-0.006(.02)</td>
<td>-0.0065(.022)</td>
</tr>
<tr>
<td>ΔInMC(-7)</td>
<td>-0.062**(.025)</td>
<td>-0.053**(.03)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ΔInMC(-8)</td>
<td>-0.0312(0.25)</td>
<td>0.023(.03)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ΔInMC(-9)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ΔInOP(-1)</td>
<td>0.0131(.15)</td>
<td>0.031**(.015)</td>
<td>-0.149**(.07)</td>
<td>0.036(.041)</td>
</tr>
<tr>
<td>ΔInOP(-2)</td>
<td>0.0050(.15)</td>
<td>0.017(.015)</td>
<td>0.09(.07)</td>
<td>0.003(.041)</td>
</tr>
<tr>
<td>ΔInOP(-3)</td>
<td>0.1660(.15)</td>
<td>0.013(.015)</td>
<td>0.018(.07)</td>
<td>-0.001(.041)</td>
</tr>
<tr>
<td>ΔInOP(-4)</td>
<td>0.0085(.15)</td>
<td>-0.014(.015)</td>
<td>-0.44**(.07)</td>
<td>0.043(.042)</td>
</tr>
<tr>
<td>ΔInOP(-5)</td>
<td>0.0458(.15)</td>
<td>0.035**(.015)</td>
<td>-0.129(.07)</td>
<td>0.037(.042)</td>
</tr>
<tr>
<td>ΔInOP(-6)</td>
<td>0.1693(.15)</td>
<td>0.008(.015)</td>
<td>-0.053(.07)</td>
<td>-0.045(.042)</td>
</tr>
<tr>
<td>ΔInOP(-7)</td>
<td>0.821**(.15)</td>
<td>-0.0009(.015)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ΔInOP(-8)</td>
<td>0.697**(.16)</td>
<td>0.018(.015)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| Obs. | 1476 | 1463 | 1998 | 1998 |

*Note:* Standard errors are in parentheses. **significant at the 5% level.
Table 2-7: The result of VECM for MC and OP 2009-2015

<table>
<thead>
<tr>
<th>Country</th>
<th>ISR</th>
<th>JOR</th>
<th>QAT</th>
</tr>
</thead>
<tbody>
<tr>
<td>ECM(-1)</td>
<td>-.005**(.001)</td>
<td>-.004**(.001)</td>
<td>-.005**(.001)</td>
</tr>
<tr>
<td>ΔlnMC(-1)</td>
<td>-.034(.24)</td>
<td>.06**(.02)</td>
<td>-.0378(.024)</td>
</tr>
<tr>
<td>ΔlnMC(-2)</td>
<td>-.001(.25)</td>
<td>-.0301(.02)</td>
<td>.0285(.024)</td>
</tr>
<tr>
<td>ΔlnMC(-3)</td>
<td>-.11**(.24)</td>
<td>.061(.02)</td>
<td>.0395(.025)</td>
</tr>
<tr>
<td>ΔlnMC(-4)</td>
<td>.002(.25)</td>
<td>-.050(.02)</td>
<td>-.0015(.025)</td>
</tr>
<tr>
<td>ΔlnMC(-5)</td>
<td>.034(.25)</td>
<td>-.035(.02)</td>
<td>-.0190(.024)</td>
</tr>
<tr>
<td>ΔlnMC(-6)</td>
<td>-.002(.02)</td>
<td>.0165(.024)</td>
<td></td>
</tr>
<tr>
<td>ΔlnMC(-7)</td>
<td></td>
<td></td>
<td>.016(.02)</td>
</tr>
<tr>
<td>ΔlnMC(-8)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ΔlnMC(-9)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ΔlnOP(-1)</td>
<td>-.03**(.014)</td>
<td>.039**(.008)</td>
<td>.1239**(.015)</td>
</tr>
<tr>
<td>ΔlnOP(-2)</td>
<td>.018(.014)</td>
<td>.015(.008)</td>
<td>.0343**(.016)</td>
</tr>
<tr>
<td>ΔlnOP(-3)</td>
<td>.012(.014)</td>
<td>.004(.008)</td>
<td>.0533**(.016)</td>
</tr>
<tr>
<td>ΔlnOP(-4)</td>
<td>.012(.014)</td>
<td>-.003(.008)</td>
<td>.0237(.016)</td>
</tr>
<tr>
<td>ΔlnOP(-5)</td>
<td>.02(.014)</td>
<td>.031**(.008)</td>
<td>.057**(.015)</td>
</tr>
<tr>
<td>ΔlnOP(-6)</td>
<td></td>
<td>.025**(.008)</td>
<td>.0035(.016)</td>
</tr>
<tr>
<td>ΔlnOP(-7)</td>
<td></td>
<td>-.004(.008)</td>
<td></td>
</tr>
<tr>
<td>ΔlnOP(-8)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Obs.</td>
<td>1672</td>
<td>1669</td>
<td>1654</td>
</tr>
</tbody>
</table>

Note: Standard errors are in parentheses. **significant at the 5% level.
Table 2-8: The result of VECM for MC and OP 2001-2015

<table>
<thead>
<tr>
<th>Country</th>
<th>ARE</th>
<th>BHR</th>
<th>ISR</th>
<th>KWT</th>
<th>QAT</th>
<th>SAU</th>
<th>TUR</th>
</tr>
</thead>
<tbody>
<tr>
<td>ECM(-1)</td>
<td>-0.02** (.00)</td>
<td>-0.01** (.00)</td>
<td>-0.03** (.001)</td>
<td>-0.02** (.00)</td>
<td>-0.001** (.00)</td>
<td>-0.003** (.00)</td>
<td>-0.001** (.00)</td>
</tr>
<tr>
<td>ΔInMC(-1)</td>
<td>0.0134 (.016)</td>
<td>-0.80** (.01)</td>
<td>-0.06 (.01)</td>
<td>-0.03** (.02)</td>
<td>0.062** (.02)</td>
<td>0.0169 (.016)</td>
<td>-0.206** (.02)</td>
</tr>
<tr>
<td>ΔInMC(-2)</td>
<td>0.148** (.016)</td>
<td>-0.62** (.02)</td>
<td>-0.018 (.015)</td>
<td>-0.03** (.02)</td>
<td>0.019 (.02)</td>
<td>0.0089 (.016)</td>
<td>-0.101** (.02)</td>
</tr>
<tr>
<td>ΔInMC(-3)</td>
<td>-0.020 (.016)</td>
<td>-0.49** (.021)</td>
<td>-0.009 (.015)</td>
<td>-0.016 (.02)</td>
<td>0.027 (.02)</td>
<td>0.006 (.016)</td>
<td>-0.007 (.017)</td>
</tr>
<tr>
<td>ΔInMC(-4)</td>
<td>-0.0267 (.016)</td>
<td>-0.43** (.023)</td>
<td>0.01 (.015)</td>
<td>0.013 (.02)</td>
<td>0.009 (.02)</td>
<td>0.003 (.016)</td>
<td>-0.027 (.017)</td>
</tr>
<tr>
<td>ΔInMC(-5)</td>
<td>0.021 (.016)</td>
<td>0.22** (.024)</td>
<td>-0.003 (.015)</td>
<td>0.036** (.02)</td>
<td>-0.20 (.02)</td>
<td>-0.005 (.02)</td>
<td>-0.06** (.02)</td>
</tr>
<tr>
<td>ΔInMC(-6)</td>
<td>-0.0096 (.015)</td>
<td>0.06** (.023)</td>
<td>-0.006 (.013)</td>
<td>-0.02** (.02)</td>
<td>0.004 (.01)</td>
<td>-0.007 (.016)</td>
<td>-0.001 (.016)</td>
</tr>
<tr>
<td>ΔInMC(-7)</td>
<td>-0.038 (.021)</td>
<td>-0.05** (.02)</td>
<td>0.005 (.01)</td>
<td>0.03 (.016)</td>
<td>0.002 (.01)</td>
<td>-0.06 (.016)</td>
<td>-0.001 (.016)</td>
</tr>
<tr>
<td>ΔInMC(-8)</td>
<td>-0.114** (.01)</td>
<td>0.027** (.01)</td>
<td>0.015 (.01)</td>
<td>0.03 (.016)</td>
<td>0.002 (.01)</td>
<td>-0.06 (.016)</td>
<td>-0.001 (.016)</td>
</tr>
<tr>
<td>ΔInMC(-9)</td>
<td>0.05** (.015)</td>
<td>0.099** (.04)</td>
<td>-0.04** (.02)</td>
<td>0.036** (.009)</td>
<td>0.077** (.01)</td>
<td>0.118** (.042)</td>
<td>0.036 (.025)</td>
</tr>
<tr>
<td>ΔInOP(-1)</td>
<td>0.019 (.016)</td>
<td>0.150** (.041)</td>
<td>-0.003 (.016)</td>
<td>-0.02** (.009)</td>
<td>0.050** (.01)</td>
<td>0.091** (.042)</td>
<td>0.027 (.025)</td>
</tr>
<tr>
<td>ΔInOP(-2)</td>
<td>0.033** (.016)</td>
<td>0.153** (.041)</td>
<td>0.04 (.08)</td>
<td>0.02** (.009)</td>
<td>0.045** (.01)</td>
<td>0.0199 (.042)</td>
<td>0.015 (.025)</td>
</tr>
<tr>
<td>ΔInOP(-3)</td>
<td>0.0098 (.016)</td>
<td>0.05 (.041)</td>
<td>0.003 (.016)</td>
<td>0.02 (.009)</td>
<td>0.025 (.01)</td>
<td>0.265** (.043)</td>
<td>0.038 (.025)</td>
</tr>
<tr>
<td>ΔInOP(-4)</td>
<td>-0.0010 (.016)</td>
<td>0.172** (.041)</td>
<td>-0.05** (.016)</td>
<td>-0.00 (.009)</td>
<td>0.022 (.01)</td>
<td>-0.06 (.043)</td>
<td>0.031 (.025)</td>
</tr>
<tr>
<td>ΔInOP(-5)</td>
<td>-0.0055 (.015)</td>
<td>-0.052 (.041)</td>
<td>-0.02 (.02)</td>
<td>-0.004 (.009)</td>
<td>0.006 (.01)</td>
<td>-0.028 (.043)</td>
<td>-0.031 (.025)</td>
</tr>
<tr>
<td>ΔInOP(-6)</td>
<td>0.015 (.041)</td>
<td>-0.02 (.02)</td>
<td>0.011 (.009)</td>
<td>0.006 (.01)</td>
<td>-0.028 (.043)</td>
<td>-0.031 (.025)</td>
<td>0.015 (.025)</td>
</tr>
<tr>
<td>ΔInOP(-7)</td>
<td>-0.007 (.041)</td>
<td>0.024 (.009)</td>
<td>-0.007 (.016)</td>
<td>0.01 (.009)</td>
<td>-0.06 (.016)</td>
<td>-0.007 (.016)</td>
<td>0.006 (.01)</td>
</tr>
</tbody>
</table>

| Obs. | 3580 | 3153 | 3674 | 3140 | 2919 | 3808 | 3923 |

Note: Standard errors are in parentheses. **significant at the 5% level.
<table>
<thead>
<tr>
<th>Country</th>
<th>ARE</th>
<th>EGY</th>
<th>ISR</th>
<th>JOR</th>
<th>QAT</th>
</tr>
</thead>
<tbody>
<tr>
<td>ΔlnMC(-1)</td>
<td>1.01**(.02)</td>
<td>.15**(.02)</td>
<td>.43**(.02)</td>
<td>1.1**(.02)</td>
<td>1.09**(.03)</td>
</tr>
<tr>
<td>ΔlnMC(-2)</td>
<td>.15**(.032)</td>
<td>.158**(.02)</td>
<td>.25**(.02)</td>
<td>-.05 (.035)</td>
<td>-.086**(.04)</td>
</tr>
<tr>
<td>ΔlnMC(-3)</td>
<td>-.17**(.032)</td>
<td>.065**(.01)</td>
<td>.09**(.02)</td>
<td>.03(.035)</td>
<td>.018(.04)</td>
</tr>
<tr>
<td>ΔlnMC(-4)</td>
<td>-.037(.033)</td>
<td>.14**(.016)</td>
<td>.15**(.02)</td>
<td>-.03(.035)</td>
<td>-.021(.039)</td>
</tr>
<tr>
<td>ΔlnMC(-5)</td>
<td>.0482(.030)</td>
<td>.69**(.02)</td>
<td>-.006(.02)</td>
<td>.007(.035)</td>
<td>-.005(.024)</td>
</tr>
<tr>
<td>ΔlnMC(-6)</td>
<td>-.0134(.020)</td>
<td>-.12**(.02)</td>
<td>-.008(.02)</td>
<td>-.01(.035)</td>
<td></td>
</tr>
<tr>
<td>ΔlnMC(-7)</td>
<td>-.08**(.02)</td>
<td>-.013(.02)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ΔlnMC(-8)</td>
<td></td>
<td>.09**(.02)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ΔlnMC(-9)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ΔlnOP(-1)</td>
<td>.045**( .026)</td>
<td>.16**(.06)</td>
<td>.05(.03)</td>
<td>.04**(.01)</td>
<td>.052**( .025)</td>
</tr>
<tr>
<td>ΔlnOP(-2)</td>
<td>-.0412(.036)</td>
<td>-.01(.09)</td>
<td>-.00(.04)</td>
<td>-.03(.017)</td>
<td>.0221(.034)</td>
</tr>
<tr>
<td>ΔlnOP(-3)</td>
<td>.03745(.037)</td>
<td>.07(.09)</td>
<td>.003(.04)</td>
<td>.008(.017)</td>
<td>-.029(.035)</td>
</tr>
<tr>
<td>ΔlnOP(-4)</td>
<td>-.0223(.036)</td>
<td>-.17(.09)</td>
<td>-.007(.04)</td>
<td>-.01(.017)</td>
<td>-.005(.035)</td>
</tr>
<tr>
<td>ΔlnOP(-5)</td>
<td>-.0458(.036)</td>
<td>.13(.09)</td>
<td>-.07(.04)</td>
<td>.005(.017)</td>
<td>-.043(.026)</td>
</tr>
<tr>
<td>ΔlnOP(-6)</td>
<td>.0260(.026)</td>
<td>-.27**(.09)</td>
<td>.03(.04)</td>
<td>-.01(.017)</td>
<td></td>
</tr>
<tr>
<td>ΔlnOP(-7)</td>
<td></td>
<td>.102(.06)</td>
<td>.06(.04)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ΔlnOP(-8)</td>
<td></td>
<td></td>
<td>-.06**(.03)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ΔlnOP(-9)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Obs.</td>
<td>1903</td>
<td>1996</td>
<td>1996</td>
<td>1711</td>
<td>1258</td>
</tr>
</tbody>
</table>

Note: Standard errors are in parentheses. **significant at the 5% level.
<table>
<thead>
<tr>
<th>Country</th>
<th>ARE</th>
<th>BHR</th>
<th>EGY</th>
<th>KWT</th>
<th>SAU</th>
<th>TUR</th>
</tr>
</thead>
<tbody>
<tr>
<td>ΔInMC(-1)</td>
<td>1.05**( .024)</td>
<td>.54**( .02)</td>
<td>.50**( .02)</td>
<td>1.07**( .02)</td>
<td>1.02**( .02)</td>
<td>.966**( .025)</td>
</tr>
<tr>
<td>ΔInMC(-2)</td>
<td>.0427( .035)</td>
<td>.13**( .02)</td>
<td>.22**( .02)</td>
<td>-.09**( .03)</td>
<td>.0024( .03)</td>
<td>.053( .035)</td>
</tr>
<tr>
<td>ΔInMC(-3)</td>
<td>-.0940**( .024)</td>
<td>.18**( .02)</td>
<td>.015( .02)</td>
<td>.01(.035)</td>
<td>-.027( .02)</td>
<td>.028( .035)</td>
</tr>
<tr>
<td>ΔInMC(-4)</td>
<td>-.047( .02)</td>
<td>.12**( .02)</td>
<td>-.03(.035)</td>
<td></td>
<td></td>
<td>-.055**( .025)</td>
</tr>
<tr>
<td>ΔInMC(-5)</td>
<td>.51( .02)</td>
<td>.54**( .02)</td>
<td>.06(.035)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ΔInMC(-6)</td>
<td>-.34**( .02)</td>
<td>-.35**( .02)</td>
<td>.019( .035)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ΔInMC(-7)</td>
<td>-.06**( .02)</td>
<td>-.12( .035)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ΔInMC(-8)</td>
<td></td>
<td>.07( .02)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ΔInMC(-9)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ΔInOP(-1)</td>
<td>.0719**( .012)</td>
<td>-.028( .03)</td>
<td>.016( .03)</td>
<td>.048(.011)</td>
<td>.065**( .01)</td>
<td>.003(.022)</td>
</tr>
<tr>
<td>ΔInOP(-2)</td>
<td>-.0374**( .017)</td>
<td>.11**( .04)</td>
<td>.08( .04)</td>
<td>-.017(.015)</td>
<td>.006 (.02)</td>
<td>.040(.030)</td>
</tr>
<tr>
<td>ΔInOP(-3)</td>
<td>-.0329( .012)</td>
<td>-.0003( .03)</td>
<td>.014( .04)</td>
<td>.002(.015)</td>
<td>-.07**( .01)</td>
<td>-.022(.030)</td>
</tr>
<tr>
<td>ΔInOP(-4)</td>
<td>-.035( .04)</td>
<td>.014( .05)</td>
<td>-.001( .015)</td>
<td></td>
<td></td>
<td>-.018 (.022)</td>
</tr>
<tr>
<td>ΔInOP(-5)</td>
<td>-.036( .04)</td>
<td>-.056( .04)</td>
<td>-.012(.015)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ΔInOP(-6)</td>
<td>-.009( .03)</td>
<td>-.13**( .04)</td>
<td>-.031(.015)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ΔInOP(-7)</td>
<td></td>
<td>.06(.03)</td>
<td>.05**( .015)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ΔInOP(-8)</td>
<td></td>
<td>-.034(.011)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ΔInOP(-9)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Obs.</th>
<th>1677</th>
<th>1669</th>
<th>1670</th>
<th>1669</th>
<th>1669</th>
<th>1669</th>
</tr>
</thead>
</table>

Note: Standard errors are in parentheses. **significant at the 5% level.
Table 2-11: The result of VAR for MC and OP 2001-2015

<table>
<thead>
<tr>
<th>Country</th>
<th>EGY</th>
<th>JOR</th>
</tr>
</thead>
<tbody>
<tr>
<td>ΔInMC(-1)</td>
<td>.19**(.01)</td>
<td>1.06**(.016)</td>
</tr>
<tr>
<td>ΔInMC(-2)</td>
<td>.17**(.02)</td>
<td>-.065**(.024)</td>
</tr>
<tr>
<td>ΔInMC(-3)</td>
<td>.11**(.02)</td>
<td>.046(.024)</td>
</tr>
<tr>
<td>ΔInMC(-4)</td>
<td>.14**(.01)</td>
<td>-.052**(.024)</td>
</tr>
<tr>
<td>ΔInMC(-5)</td>
<td>.67**(.01)</td>
<td>.013(.024)</td>
</tr>
<tr>
<td>ΔInMC(-6)</td>
<td>-.15**(.02)</td>
<td>-.005(.024)</td>
</tr>
<tr>
<td>ΔInMC(-7)</td>
<td>-.08**(.02)</td>
<td></td>
</tr>
<tr>
<td>ΔInMC(-8)</td>
<td>-.05**(.02)</td>
<td></td>
</tr>
<tr>
<td>ΔInMC(-9)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ΔInOP(-1)</td>
<td>.11**(.04)</td>
<td>.035**(.007)</td>
</tr>
<tr>
<td>ΔInOP(-2)</td>
<td>.0489(.06)</td>
<td>-.025**(.010)</td>
</tr>
<tr>
<td>ΔInOP(-3)</td>
<td>.0167</td>
<td>.003(.010)</td>
</tr>
<tr>
<td>ΔInOP(-4)</td>
<td>-.11(.06)</td>
<td>-.012(.010)</td>
</tr>
<tr>
<td>ΔInOP(-5)</td>
<td>.10(.06)</td>
<td>.014(.010)</td>
</tr>
<tr>
<td>ΔInOP(-6)</td>
<td>-.23**(.06)</td>
<td>-.015**(.007)</td>
</tr>
<tr>
<td>ΔInOP(-7)</td>
<td>.07(.06)</td>
<td></td>
</tr>
<tr>
<td>ΔInOP(-8)</td>
<td>-.004(.06)</td>
<td></td>
</tr>
<tr>
<td>ΔInOP(-9)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Obs.</td>
<td>3673</td>
<td>3389</td>
</tr>
</tbody>
</table>

Note: Standard errors are in parentheses. **significant at the 5% level.
Table 2-12: Summary of results

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Country</td>
<td>r</td>
<td>Linkage</td>
<td>r</td>
<td>Linkage</td>
<td>r</td>
<td>Linkage</td>
</tr>
<tr>
<td>United Arab Emirates (ARE)</td>
<td>No</td>
<td>SRR</td>
<td>Yes</td>
<td>SRR</td>
<td>Yes</td>
<td>LRR</td>
</tr>
<tr>
<td>Bahrain (BHR)</td>
<td>Yes</td>
<td>LRR</td>
<td>No</td>
<td>SRR</td>
<td>Yes</td>
<td>LRR</td>
</tr>
<tr>
<td>Egypt (EGY)</td>
<td>No</td>
<td>SRR</td>
<td>No</td>
<td>SRR</td>
<td>No</td>
<td>SRR</td>
</tr>
<tr>
<td>Israel (ISR)</td>
<td>No</td>
<td>SRR</td>
<td>Yes</td>
<td>LRR</td>
<td>Yes</td>
<td>SRR</td>
</tr>
<tr>
<td>Jordan (JOR)</td>
<td>No</td>
<td>SRR</td>
<td>Yes</td>
<td>LRR</td>
<td>No</td>
<td>SRR</td>
</tr>
<tr>
<td>Kuwait (KWT)</td>
<td>Yes</td>
<td>LRR</td>
<td>No</td>
<td>SRR</td>
<td>Yes</td>
<td>LRR</td>
</tr>
<tr>
<td>Qatar (QAT)</td>
<td>No</td>
<td>SRR</td>
<td>Yes</td>
<td>LRR</td>
<td>Yes</td>
<td>LRR</td>
</tr>
<tr>
<td>Saudi Arabia (SAU)</td>
<td>Yes</td>
<td>LRR</td>
<td>No</td>
<td>SRR</td>
<td>Yes</td>
<td>LRR</td>
</tr>
<tr>
<td>Turkey (TUR)</td>
<td>Yes</td>
<td>LRR</td>
<td>No</td>
<td>SRR</td>
<td>Yes</td>
<td>SRR</td>
</tr>
</tbody>
</table>

Note: Cointegration (Yes, r>=1), no-Cointegration (No, r=0), ECM (Long-run linkage LRR) and VAR (Short-run linkage SRR)
Figure. 2-2 Market capitalization (MC in Billion USD) and crude oil price (OP in USD) by Country

The United Arab Emirates (ARE)

Bahrain (BHR)

Egypt (EGY)

Israel (ISR)

Jordan (JOR)

Kuwait (KWT)

Qatar (QAT)

Saudi Arabia (SAU)

Date

MC

OP

MC

OP

MC

OP

MC

OP

MC

OP

MC

OP

MC

OP

MC

OP
Figure.2-3: Impulse response function (IRF) for each Middle East economy


BHR
Graphs by irfname, impulse variable, and response variable.
Chapter 3 - Essay #2: The Dynamic Response of Coal Consumption to Energy Prices and GDP: An ARDL Approach to the U.S.

1. Introduction

A variety of energy sources and technologies are used in the United States to generate electricity. These sources and technologies have changed over time based upon price, efficiency and environmental impacts. The U.S. Energy Information Administration (EIA) lists the three main categories of energy sources for the 4,015 billion kilowatt-hours (kWh) of electricity generated in 2017: fossil fuels (coal, natural gas, and petroleum), nuclear energy, and renewable energy sources. About 63% of this electricity generation is from fossil fuels (coal, natural gas, petroleum, and other gases), of which just under ½ of this power supply originates from coal. Coal has served as a crucial source of energy to generate electricity for many decades. As much of the U.S. gravitates towards more efficient and sustainable energy sources for generating electricity, natural gas is seen as a transition fossil fuel, able to produce less polluting electricity than coal. The possibility for substitution between energy sources and the sensitivity of power utilities to changes in the prices of energy sources are essential for understanding energy markets as policy makers seek to promote cleaner energy and enhance energy efficiency. Therefore, coal-to-natural gas substitution in the power plants is expected to drive consumption down for coal in energy market.

Rising concerns about climate change have resulted in increased consumption of clean energy for electricity in the U.S. as well as in nations around the world. Fuel switching, or the substitution of low-carbon energy sources for fossil fuels with higher carbon content is one of the principal methods suggested to reduce carbon dioxide (CO₂) emissions (Hayhoe et al., 2002).
According to the U.S. EIA, the amount of CO\textsubscript{2} produced when coal (228)\textsuperscript{4} is burned is almost twice that of natural gas (117) due to the carbon content of the coal. Thus, utilization of natural gas in place of coal leads to a significant reduction in CO\textsubscript{2} emissions when generating electricity. Natural gas is a possible alternative to coal also improves conversion efficiency of energy to electricity, 62% for natural gas (rising to over 90%) compared to a 25% to 35% efficient coal-fired plant (Audus, 1999; Nakicenovic, 2000; Hayhoe et al., 2002). Considering that the majority of CO\textsubscript{2} emissions originate from coal consumption, the ability to switch coal to natural gas in the power plants become a main target and instrument for energy and environmental policy makers.

Furthermore, choices of fuel and factor alternatives are critical issues in energy policy. A natural gas price shock could have significant impacts on the costs and utilization of coal in energy market and power plant relatively (Atkeson and Kehoe, 1999). Since natural gas combustion has lower amounts of CO\textsubscript{2}, sulfur dioxide (SO\textsubscript{2}), nitrogen oxides (NOx) emissions per megawatt hour (MWh), fuel displacing from coal to natural gas reduces environmental emissions. We could argue that increased natural gas usage by power plants is not only related to efficiency and reducing pollution, but also leads to enhanced links between natural gas and coal prices.

Natural gas development has increased gradually in the U.S.’ energy industry. Coal has long been the U.S. preferred source of cheaper energy, although falling natural gas prices that stabilized below $2.50/MBTU make natural gas a competitive energy source for power plants (Slingsby, 2015). During the past three decades, the natural gas market has seen significant price volatility, and the U.S. energy market has shifted away from coal-fired power plants toward natural gas–fired power generation (U.S. EIA, 2016). Moreover, the commercialization of

\footnote{Pounds of CO\textsubscript{2} emitted per million British thermal units (Btu) of coal.}
fracking technology starting in the 2008-2009 led to dramatically lower natural gas prices. At the same time, U.S. electricity generation from coal has decreased along with domestic coal consumption (Figure 3-1). As a result, the production of coal has decreased over past two decades. As shown in Figure 3-2, decreasing coal extraction has affected labor market and income for local and state economies in the most producing coal states in the U.S. (EIA, 2017). Moreover, efficiency and lower natural gas prices have led to more consumption of natural gas and less production of coal due to the development of shale gas resources in the U.S. (U.S. EIA, 2016).

![U.S. power generation in billions of megawatt-hours](image)

Figure 3-1: U.S. power generation from coal and natural gas (source: U.S. EIA (2016))

---

5 There are 11 states including Wyoming, West Virginia, Pennsylvania, Illinois, Kentucky, Texas, Montana, Indiana, North Dakota, Colorado, and Ohio.
Fuel choice for power generation relies on fuel prices. While the elasticity of demand for energy as a whole is low (Platchkov and Pollitt, 2011), own-price demand elasticity for coal should be positive because of the presence of demand side substitutes for coal, such as natural gas and renewable energy sources. Thus, there will be a shift from coal to natural gas if natural gas prices remain lower compared to coal, as both coal and natural gas are consumed in electricity generation. A study conducted by U.S. EIA (2012) showed that coal consumption and natural gas price cross elasticity was 0.17 and own price elasticity was -0.11 for coal. If coal and natural gas are substitutes, it is hypothesized that there is a positive relationship between consumption of coal and price of natural gas.

The effects of natural gas prices on economic activities have been studied, but the related research have lesser insights for changes of natural gas price and related impacts on the coal industry. There is literature (Hamilton, 2005; Kilian, 2008) about the linkage between energy prices and economic activity, but there is no study specifically related to natural gas price volatility and coal consumption. Accordingly, several studies using various estimation models and samples have been applied to the analysis of energy elasticities and fuel substitution in energy markets. However, no previous research has estimated the relationship between both coal
and natural gas prices and coal consumption in the U.S. This study is the first research to examine linkage between energy prices and the consumption of coal in the U.S. energy market using an Autoregressive Distributive Lag (ARDL) model. The focus of this study is to investigate the linkages between coal and natural gas prices and the coal consumption in the U.S. and to reveal other characteristics and underlying elements in the coal market.

It is critical to examine the changing relationship between natural gas and coal since natural gas has been expanding its electricity generation market share at the expense of coal in recent years (Figure 3-1). Recent evidence shows that utilities are more likely to build natural gas–fired plants than coal-fired plants (U.S. EIA, 2012; Xie et al., 2016). Given this background, the present research investigates both the long-run and short-run dynamic responses of coal consumption in the U.S. to changes in energy prices (coal and natural gas) along with GDP and other related factors. It is important to know the potential magnitude of impacts from price based policies to determine their suitability for transitioning to clean energy sources, thereby reducing carbon emissions for the U.S. economy.

The main contribution of this study is to develop an economic framework for estimating the dynamic long–run linkages between coal-natural gas prices and coal consumption in the U.S. economy. In addition, this research contributes to the literature by establishing coal-natural gas linkages utilizing an ARDL model that allows for exploration of both short and long-run relationships among variables. A novel contribution is estimating how much of the decrease in coal use occurred as a result of declining natural gas prices in the U.S energy market. This research investigates the extent to which coal and natural gas prices, economic growth, and other factors have encouraged a switch from coal use to natural gas consumption. If coal consumption responds significantly to these factors, these changes impact policy makers towards encouraging
alternative energy sources such as natural gas. Hence, to determine the strength of relationships, we examine the short and long-run correlations between coal consumption, natural gas price, coal price, and other factors over time period 2000-2016.

The organization of the paper is as follows: Section 2 gives a summary of previous studies. Sections 3 and 4 include discussions of the theoretical approach and empirical model. Section 5 presents the empirical results.

2. Literature Review

In recent years, the crisis of coal industry in the U.S. has received increasing attention in the literature. The potential for natural gas to replace coal has promoted recommendation that natural gas could be used as a bridging fuel until close to zero carbon energy sources are more environmentally and economically viable (Levi, 2013). Burke and Lioa (2015) found evidence from provincial data in China that in the long-run coal demand has increasingly price elastic. Hence, China’s coal market is more suitable for price based approach to reduce demand.

The issue of energy demand and coal industry, have also become the subject matter of scholarly studies that address various topics. For instance, Apergis and Payne (2010) examine the renewable and non-renewable energy demand, finding that the demand of clean energy sources increase with the growth of economy. Nonetheless, the relationship between economic growth and energy consumption remain a debate (Hamilton, 2005, and Kilian, 2008; Apergis and Payne, 2012; Salim and Hassan, 2014; Ben Aissa et al, 2014). During early phase of economic development, environment degrades and pollution increases with the expansion of income. However, environment enhances with the economic growth, along with clean technology diffusion, and new approaches to pollution regulation (Dasgupta et al., 2002). Apergis and Payne (2010) examine the relationship between coal consumption and economic growth for 25
OECD countries over 1980-2005. The findings show that there exist both short and long-term bidirectional causality between coal consumption and economic growth; however, the short run causality is negative. Policy towards lessening coal consumption might not have an adverse impact on economic growth. Economic growth reduces coal consumption as evident from the fact that proportion of coal use in electricity generation has been lessening around the world (Wolde-Rufael, 2010). Using a panel of 10 hydrocarbon consuming nations, there exists short-run bidirectional causality between economic growth and energy consumption (Ghali and El-Sakka, 2004).

Other studies have also been examined regarding coal and natural gas. For example, Arezki et al. (2017) have studied the surge in the generation of shale gas in the U.S. energy markets. The study finds that the shale gas revolution resulted in change in long-term gas price in the U.S. However, the difference in natural gas prices is due to differences in its endowment of natural gas which is difficult to trade. The production of shale gas led to increase in gas supply impacted the long-run trend in gas prices. Neighbouring countries such as Canada and Mexico may directly benefit from exports of the U.S. shale gas. This is particularly boon for the U.S. economy when decreasing supply of conventional natural gas sources begun to reduce (Atkeson and Kehoe, 1999). Prices of coal are determined by long-run contracts, while natural gas prices are determined territorially and state level demand supply relations (Renou-Maissant, 2012; Caporin and Fontini, 2016; Olsen et al., 2015).

There is a considerable amount of literature highlighting the substitution of coal and natural gas in power plants. For instance, Xie et al. (2016) investigate the potential substitution between coal and natural gas in China. Results show elasticity estimates of both coal-gas substitution and substitution of coal electricity to be positive over 1985-2012, suggesting that
these two energy sources are substitute. It also indicates that there is relatively higher substitution between coal and natural gas, and less possibilities to substitute coal with other fuels. This is on the grounds that coal and natural gas are substitute as inputs and utilized in the same market as factors of production. The impacts of natural gas prices on economic activities have been studied, however the related research have lesser insights for natural gas price volatility and coal industry.

3. Methodology

On a national basis, the consumption for coal is assumed to be a function of national income, coal prices, and natural gas prices. Other exogenous factors can likewise determine coal use including weather, seasonality, efficiency and other factors related to energy demand. We assume that energy prices, income, and weather are probably going to be the most important determinants of energy consumption. The functional relationship is assumed as follows:

\[
CC_t = f(Y_t, CP_t, GP_t, GDP_t, HDD_t, CDD_t, SC_t) \tag{3-1}
\]

Where \(CC_t\) is consumption for coal in period \(t\), \(Y_t\) is the real GDP in \(t\) period, \(CP_t\) is the price of coal in \(t\) period, \(GP_t\) is the price of natural gas, \(HDD\) is heating degree-days, \(CDD\) is cooling degree-days and we add solar energy use \((SC_t)\) as a control variable for coal consumption.

Since coefficients for variables on a natural-log scale are directly interpretable as proportional differences and our data are distributed normally in log form, all the variables are converted into natural logarithm, which helps to reduce the variability of data series. Thus, logarithmic transformation of the above function takes the following form:

\[
CC_t = \alpha_0 + \alpha_1 Y_t + \alpha_2 CP_t + \alpha_3 GP_t + \alpha_4 HDD + \alpha_5 CDD + \alpha_6 SC_t + \alpha_7 d\nu_t + \epsilon_t \quad \tag{3-2}
\]

where \(\epsilon_t\) the stochastic disturbance term and \(d\nu_t\) is the dummy variable which is seasonality variable. The coefficient \(\alpha_1\) represents the income elasticity of coal demand, which is expected to
be negative as higher incomes increase demand for clean energy and substitute the use of coal. The coefficient $\alpha_2$ is expected to be negative because higher price of coal might lead to substitution with other energy sources for coal. $\alpha_3$ represents cross price elasticity for natural gas, which should be positive sign. The assumption that natural gas can be substituted for coal is reasonable based upon the discussion in the literature review. First, fossil fuels have provided over 80% of U.S. energy for many decades (U.S. EIA, 2016). Second, renewable energy (i.e. solar) is still more expensive than natural gas to compete with natural gas at utility scales without large subsidies. The coefficient sign for HDD ($\alpha_4$) is expected to be positive as higher HDD should increase electricity consumption and lead to higher coal consumption. The greater number of CDD leads to more heating demand and thus is expected to have a positive effect on coal consumption ($\alpha_5>0$). The consumption of solar as a control variable and $\alpha_6$ is expected to be negative since solar represents a substitute for coal. The coefficient $\alpha_7$ represents a dummy variable to reflect a seasonality variable in the ARDL model. It can be expected signs of the coefficients are as follows: $\alpha_1<0$, $\alpha_2<0$, $\alpha_3>0$, $\alpha_4>0$, $\alpha_5>0$, $\alpha_6<0$, $\alpha_7>0$.

Initially, the stationarity of variables is checked by unit root test. If the variables are found to be non-stationary at level, we convert into first difference and then examine the stationarity. The long-run relationship among the variables can be estimated from equation 3-2. On the other hand, to estimate long-run dynamics this model can be estimated by bound testing procedures of an ARDL model (Pesaran et al. 2001).

3.1 Unit Root Test

Before applying any cointegration tests of time series variables it is required to check whether the variables are integrated or order one I(1). This is because estimating non-stationary data can yield spurious causality (Wu, e.t al, 2016). Following Engle and Granger, we first test
the unit roots of the variables used in the study to confirm their stationarity, which is done by the augmented Dickey-Fuller (ADF) and the Phillips-Perron (PP) tests. If any variables are found to be non-stationary, then first difference is applied to estimate the long run and short run dynamics of the variables considered in the study.

3.2 Autoregressive Distributed Lag (ARDL) Model

The empirical approach of this study is to explore the long-run and short-run relationships between coal consumption and its determinants. Short-run and long-run\(^6\) refer to estimate horizons defined with respect to a given point in time. Moreover, if variables are found to be co-integrated, the long-run elasticities are estimated from the co-integration equation and the short-run elasticities can be estimated from an Error Correction Model (ECM). Using an ARDL framework in the present study measures the income and price elasticities of coal demand in the U.S. energy market.

The rationale behind the selection of ARDL model is due to following reasons. The ARDL can be adopted whether the variables are stationary or non-stationary, or the combination of both. This helps to remove the problem of spurious regression before testing cointegration exercise (Nkoro and Uko, 2016). ARDL model avoids the problems of low power in detecting cointegrating relationship irrespective of the length of data. A dynamic ECM can be developed from the ARDL model (Banerjee et al., 1993). Using ordinary least square (OLS) estimation technique the log-linear functional form of ARDL model can be represented as follows:

\[
\Delta C = \alpha_0 + \sum_{i=1}^{n} \beta_i \Delta C_{t-i} + \sum_{i=0}^{n_1} \gamma_i \Delta CP_{t-i} + \sum_{i=0}^{n_2} \varphi_i \Delta GP_{t-i} + \sum_{i=0}^{n_3} \theta_i \Delta Y_{t-i} + \sum_{i=0}^{n_4} \theta_i \Delta H DD_{t-i} + \sum_{i=0}^{n_5} \psi_i \Delta CDD_{t-i} + \sum_{i=0}^{n_6} \psi_i \Delta SC_{t-i} + \delta_1 C_{t-1} + \delta_2 CP_{t-1} + \delta_3 GP_{t-1} + \delta_4 Y_{t-1} + \delta_5 H DD_{t-1} + \delta_6 CDD_{t-1} + \delta_7 SC_{t-1} + \delta_8 d v_{t-1} + \epsilon_t \tag{3-3}
\]

\(^6\) Equation 3-4 is long-run and equation 3-3 is short –run. We assume that a month represents short-run and years are long-run.
where $\Delta$ is the difference operator and $\varepsilon_t$ is the error term. In order to obtain the optimal lag length for each variable, the ARDL bounds testing follows estimation of regressions. The optimal lag length is determined using Akaike’s information criterion (AIC) that specifies the rule used lag selection. Equation (3-3) was estimated utilizing the ARDL approach to determine if the dependent and independent variables in each model are cointegrated. The joint $F$-test is conducted on the lagged log-level variables while the differenced variables have no direct impact on the bounds cointegration test (Pesaran, 1997). The $F$-statistic investigates the joint null hypothesis that the coefficients of the lagged log-levels are equal to zero which suggests that no cointegrating long-run relationship exists.

The Wald test or joint $F$ statistics is utilized to examine the null hypothesis of cointegrating linkages, $H_0: \delta_1 = \delta_2 = \delta_3 = \delta_4 = \delta_5 = \delta_6 = \delta_7 = \delta_8 = 0$, which is tested against the alternative of $\delta_1 \neq \delta_2 \neq \delta_3 \neq \delta_4 \neq \delta_5 \neq \delta_6 \neq \delta_7 \neq \delta_8 \neq 0$. The long run linkage can be detected by the statistical significance of $F$ statistics. The rejection of $H_0$ infers that we have a long-run relationship. When the $F$-statistics falls beneath the lower bound, the null hypothesis cannot be rejected. Thus, we would presume that the factors are I(0). If the $F$-statistics surpasses the upper bound, there exists cointegration and rejection of the null hypothesis of no cointegration can be rejected regardless of order of integration. Eventually, if the $F$-measurement falls between the limits, the test yields no conclusion.

Accepting that the bound test prompts the decision of cointegration, we can seriously evaluate the long-run connection between the variables. The long run dynamics is estimated as:

$$
\Delta C = \alpha_0 + \delta_1 C_{t-1} + \delta_2 CP_{t-1} + \delta_3 GP_{t-1} + \delta_4 Y_{t-1} + \delta_5 HDD_{t-1} + \delta_6 CDD_{t-1} + \delta_7 SC_{t-1} + \delta_8 SI_{t-1} + \varepsilon_t
$$

(3-4)

In order to determine the direction of the causal relationship using Granger causality model augmented with lagged error correction (ECM) is a cointegration relationship exist, and
Toda and Yamamoto (1995) augmented Granger causality in the models without cointegration relationship among the variables. Thus the parameters pertaining to short run dynamics are estimated by Error Correction Model (ECM) attached with the long run estimates.

\[
\Delta C = \alpha_0 + \sum_{i=1}^{m} \beta_i \Delta C_{t-i} + \sum_{i=0}^{n^1} \gamma_i \Delta CP_{t-i} + \sum_{i=0}^{n^2} \varphi_i \Delta GP_{t-i} + \sum_{i=0}^{n^3} \theta_i \Delta Y_{t-i} + \sum_{i=0}^{n^4} \rho_i \Delta HDD_{t-i} + \sum_{i=0}^{n^5} \psi_i \Delta CDD_{t-i} + \sum_{i=0}^{n^6} \phi_i \Delta SC_{t-i} + \sum_{i=0}^{n^7} \chi_i \Delta SI_{t-i} + \phi ECT_{t-1} + \epsilon_t
\]

(3-5)

where \( \phi \) is the coefficient of the error correction term, \( ECT_{t-1} \) is the lagged error correction term which is supposed to be based on the residuals from the long-run equilibrating relationship. Thus, \( ECT_{t-1} \) is the speed of adjustment for long run equilibrium each period in dependent variable when a change occurs in other variables, \( m \) is the lagged length, and \( \epsilon_t \) is the serially uncorrelated error term. The sign of \( ECT_{t-1} \) should be negative, which indicates how fast the variables converge towards equilibrium in the long run.

Following Brown et al. (1975), the cumulative sum (CUSUM) test and cumulative sum of squares (CUSUMSQ) test have applied to assure the stability of the error correction model. Stability of the model can be ensured of plots of CUSUM or CUSUMSQ measurements remain inside basic limits of the 5 percent significance level. Finally, the Impulse Response Function (IRF) is employed to investigate how coal consumption respond to shocks in other variables included in the model. Moreover, in order to ensure models are proper to estimate, we imply tests for normality and serial correlation.

### 3.3 Structural Breaks

Structural break test takes into consideration structural changes which can occur in the intercept and/or the trend of the series ((Zivot and Andrews, 1992). Thus, in view of the Perron's
portrayal of shaping the auxiliary break, Zivot and Andrews (1992) considered three models to test for a unit root. These model equations are:

Model 1: $$\Delta y_t = c + \alpha y_{t-1} + \beta t + \gamma DU_t + \sum_{j=1}^k d_j \Delta y_{t-j} + \varepsilon_t$$  \hspace{1cm} (3-6)

Model 2: $$\Delta y_t = c + \alpha y_{t-1} + \beta t + \theta DT_t + \sum_{j=1}^k d_j \Delta y_{t-j} + \varepsilon_t$$  \hspace{1cm} (3-7)

Model 3: $$\Delta y_t = c + \alpha y_{t-1} + \beta t + \gamma DU_t + \theta DT_t + \sum_{j=1}^k d_j \Delta y_{t-j} + \varepsilon_t$$  \hspace{1cm} (3-8)

where $$y$$ represents the variable of interest in detecting a structural break in equation 3-1. $$DU_t$$ is a dummy variable mean shift happening at every conceivable date of break (TB). $$DT_t$$ is relating pattern shift variable. $$DT_t = 1$$ if $$t > TB$$, else 0; $$DT_t = t - TB$$, else 0. Equation 3-6 deals with a level one shift as it takes into consideration an adjustment in the level of the trend function. Equation 3-7 presents growth shift and equation 3-8 is viewed as level-cum growth shift one in which both the level and the slope of trend function are permitted to change after the break. The null hypothesis for these models is $$\alpha = 0$$ indicating that the series contains a unit root that with a drift that bars a structural break. The alternative hypothesis $$\alpha < 0$$ infers that the arrangement is a trend stationary process with a one-time break happening at an obscure point in time. From among all conceivable break-focuses (TB), the system chooses as its decision of break-date (TB) the date which limits the one-sided testing of t-statistics. As indicated by Zivot and Andrews (1992), the significance of t-statistics is evaluated by utilizing asymptotic assessed breakpoint basic qualities.

4. Data

The empirical estimation of the research is based on monthly data for the period of January 2000 to June 2016. The dataset consists of 198 observations of eight variables: coal consumption, coal price, and price of natural gas, Gross Domestic Products (GDP), Heating Degree-Days (HDD), Cooling Degree-Days (CDD), solar energy consumption, and a seasonal
dummy variable\textsuperscript{7}. The idea is to include a seasonal dummy variable for each month which will
then change in a variable resulting from any seasonal fluctuations. Seasonality, in fact, is the
presence of variations that happen at specific regular intervals less than a year, such as weekly,
monthly, or quarterly. In this study we have monthly data for seasonality dummy variable.
Seasonality may be caused by different factors, such as economic conditions, weather, natural
disaster, and holidays. However CDD and HDD are only affected by temperatures and amount of
energy that need to cool or heat a space. The data was collected from U.S. EIA energy market
data (U.S. EIA, 2016), Federal Reserve Economic Data (FRED) and Bureau of Labor Statistics
(BLS). The monthly data of coal consumption is measured by Million British thermal unit
(MBtu)\textsuperscript{8}, coal and natural gas price are measured by U.S dollar/MBtu. In addition, the amount of
solar consumption in the U.S’ energy market as the proxy for policies supporting green energy
came from U.S. EIA. Hence, Solar Energy Consumption (SC) is measured by trillion Btu. The
dependent variable is coal consumption, where monthly consumption of coal is used. The
explanatory variables are GDP of the U.S, prices of coal and natural gas, HDD, and CDD. All
the nominal variables were converted into real variables using consumer price index (CPI). The
monthly data of coal consumption, coal prices, and natural gas prices were retrieved from U.S.
Energy Information Administration (EIA). In addition, HDD is a measurement designed to
quantify the demand for energy needed to heat a location which is how cold the temperature was
on a given day or over a period of days. A CDD reflects the amount of energy used to cool a
location (U.S. EIA, 2016). Data for HDD, and CDD were retrieved from National Oceanic and

\textsuperscript{7} The seasonal dummy variable is one for winter (December, January and February) and summer months (June, July,
and August) and zero for spring and autumn months.

\textsuperscript{8} A Btu is a unit of energy measured as the amount of heat required to raise the temperature of one pound of water
by one degree Fahrenheit. In this study, coal consumption is measured by the quantity of western coal in tons and
multiplied by its heat content per ton, and the quantity of eastern coal in tons multiplied its heat content per ton, and
then add the two together to get the total coal consumed in MBtu.
Atmospheric Administration (NOAA). Time series data on real GDP have come from Bureau of Economic Analysis (BEA). Based on shale gas revolution date in the U.S’ energy market and Zivot – Andrews test, July, 2008 was identified as the breakpoint for coal consumption data. Thus, we split data to two sub samples which are pre-shale gas period of 2000-2008 and post shale gas period of 2009-2016. Table 3-1 presents the summary statistics of the data.

As far as trends of the variables are concerned, Figure 3-3 presents the monthly GDP of the U.S from 2000 to 2016. It can be observed that GDP witnessed upward trend till May 2008, then slowed down due to financial crisis. Since March 2009, GDP continued upward trend and continued till then. As indicated from the Figure 3-4, coal consumption exhibited a falling trend throughout the study period. On the other hand, coal prices did not show volatile price behavior except for a period in 2010 and 2011 (Figure 3-5). Lastly, natural gas prices followed a range bound pattern throughout the study period, varies between $1 and $7/MBTU (Figure 3-5). Figures 3-6 and 3-7 show the trend of monthly heating degree-days and cooling degree- days. Figure 3-8 is monthly solar consumption which increased dramatically towards the end of the time period.
5. Results

The stationarity properties are checked using ADF and PP unit root test. Moreover, Zivot-Andrews structural breaks are calculated to find unit root and the structural break of the variables used, particularly coal consumption and prices of coal and natural gas which are important from the point of view of shale gas revolution. The cointegrating relationships between variables are tested using bounds test for cointegration using ARDL approach. The cointegrating equation is estimated using ARDL model to estimate long-run own price and cross price elasticity. Finally, short-run elasticity is calculated using ECM. In the present study, we have conducted the long-run and short-run linkage between coal consumption and its determinants separately from 2000 to 2008 for pre shale and 2009 to 2016 for post shale gas because of structural break\(^9\) in the U.S. ‘energy market.

Table 3-2 presents the results of ADF and PP tests for each variable considered at their levels and first difference. It is hypothesizing that all the variables have the unit root. It is evident from the ADF test statistics that coal prices, coal consumption, natural gas price, HDD, CDD, and solar energy use are non-stationary at levels. Therefore, null hypothesis of unit root could not be rejected. In PP test, except for coal consumption and HDD, all other variables have unit root and null hypothesis could not be rejected. Therefore, the first difference of each variable was conducted, where both the ADF and the PP test statistics reject the null hypothesis of unit root since the absolute values of test statistics for each variable are greater than 5% critical values. Therefore, it can be said that all the variables are integrated of order one, I(1), i.e., stationary at first difference and therefore, suitable to investigate the presence of a dynamic long-run relationship between them.

\(^9\) Based upon Zivot-Andrews test.
It is also assumed that for the two sub periods the variables are integrated at first difference and fit for long-run further analysis. The results of endogenous structural break test of Zivot -Andrews’s unit root test results show that structural breaks and non-stationary behaviour in all variables used in the study. Utilizing this procedure, the time of the structural changes for each of the variable is identified based on the most significant t statistics. The most significant structural break occurred in July 2008 for coal consumption, coal price, natural gas price and GDP. The first test in ARDL procedure is to investigate the presence of long-run relationship of equation 3. In order to run the ARDL model, we need to select the number of lags. There are several indicators such as LR, FPE, AIC, Schwarz information criterion (SBC), and Hannan-Quinn information criterion are applied to select the optimum lag. The order of the lag distribution on the dependent variables and regressors can be selected on the basis of the AIC since it chooses the maximum important lag length (Pesaran and Shin 1999).

Table 3-3 presents the results of F bound test. The calculated F-statistics of coal consumption, coal price, natural gas price, GDP, HDD, CDD, and SC are higher than the upper bound critical value of 3.79 at a 5 percent level. As this falls beyond the inconclusive region, a null hypothesis can be rejected of no cointegration relationship, indicating that a long-run cointegrating exists linking the dependent and independent variables in this model. Accordingly, in these cases, we conclude the presence of a long-run cointegrating relationship among variables. That is, the variables tend to move together dynamically over time. The same exercise was carried out for pre / post shale gas revolution periods. The F-statistics in both sub periods signify cointegrating relationships exist between the variables. However, there is no short run significant impact of coal price on coal consumption in post shale period.
After confirming cointegration, we move to estimate long-run and short-run coefficients. Long-run coefficients for all three time periods are reported in Table 3-4 where coal consumption is treated as the dependent variable. For the entire period, the explanatory variables represent determinant factors over the long-run - the 2000 to 2016 period. There exists dynamic long-run relationships between coal consumption and all independent variables based upon statistically significant coefficient estimates over the entire period model. Weather variability is accounted for by including HDD and CDD. An increase in HDD by 1% leads to a decrease in the monthly coal consumption 0.09%. Similarly, an increase in the CDD by 1% reduces coal consumption, but by only 0.03%. While these elasticities are small, the results are the opposite signs of those expected. The findings of HDD and CDD suggest that other energy sources (i.e. natural gas or solar) may be used during winter and summer time since there is negative relationship between HDD, CDD and coal consumption.

Income elasticity of coal consumption shows that the level of responsiveness of the demanded amount of coal in respect to changes in income level and economic growth. The income elasticity for the entire period is -0.41 and statistically significant, which is in line with expectations, so that an increase in GDP leads to a fall in coal consumption. The results provide evidence that there is a significant negative impact of GDP on coal demand, implying that coal is an inferior energy sources in the market. Also, this finding implies that economic growth has stimulated declining coal consumption via more efficient and environment friendly energy sources to generate electricity.

During the pre-shale period (2000-2008), long run linkages are found to exist between coal consumption and coal price, natural gas price, HDD, and CDD. The other explanatory variables do not show long run linkages. There exists statistically significant for own price and
cross price elasticity for the coal consumption. Since the primary goal is to examine linkages between coal consumption and energy prices, the coefficient estimates meet with the expectations discussed in the methodology section. These findings are in line with previous literature estimates (Hayhoe et al., 2002; Xie et al., 2016), who found that natural gas is an alternative for coal in energy market and power plants. During the post-shale gas period of 2009-2016, there are not any statistically significant long run linkages between coal consumption and the independent variables. While this result is difficult to explain, further discussion of this result will be provided in the conclusions.

The cross price elasticities of coal consumption and natural gas, as expected, are positive at 0.29 for entire period, 0.05 for pre-shale, and 0.002 for post-shale gas. The coefficients for the entire period and pre-shale are statistically significant which suggests that coal and natural gas are substitute goods in U.S. energy markets. For the entire period, a 1 percent decrease in natural gas price leads to 0.3 percent decrease in coal consumption. Comparing this finding to the literature reveals that coal-natural gas cross elasticity during the entire period is higher than a previous estimate by the EIA (2012) at 0.17. Moreover, the own price elasticity is lower than the cross price elasticity over the entire period.

The third stage of the ARDL procedure is to estimate the Error Correction Model (ECM) which includes the coefficients of short run model and the Error Correction Term (ECT). As explained previously in equation 3-5, the ECT contains residuals of the long run model. Coefficients of ECT indicate whether or not the variables in the model have a long run association. Hence, the ECT coefficients estimate the speed of adjustment for long run equilibrium which a dependent variable returns to equilibrium after a shock in the other
variables. The sign of ECT coefficient should be negative and statistically significant which expresses adjustment speed towards long-run equilibrium in the ARDL model.

Table 3-5 presents the estimated ECM of ARDL equations of short-run\textsuperscript{10}. The ECT coefficient gives an indication of short run deviations from the long run equilibrium. As expected, the adjustment variable ECT (-1) in each of the three time periods is negative and is statistically significant at a 1 percent level. These results ensure that stable long-run relationships exist among variables in the coal consumption model. The magnitude of the error correction coefficient is -0.95 in 2000-2016, -1.18 in 2000-2008, and -0.76 in the period of 2009-2016. Therefore, it implies that disequilibria in this model is corrected by approximately 0.95 percent every month over the entire study period. During the pre-shale revolution period, coal consumption adjusts towards long-run equilibrium at much faster rate compared to the post-shale revolution period. Coal consumption, coal- natural gas prices, GDP and other explanatory variables have a tendency to converge towards an equilibrium path in the long run.

Considering the short-run or static analysis, the estimated short-run elasticities of all variables are statistically significant except CDD over the entire period. In the pre-shale gas period, the variable coefficients of coal price, the price of natural gas, GDP and solar consumption are not statistically significant. The results also show that HDD and CDD coefficients are statistically significant at 1\% level of significance for the pre-shale period. In the post-shale period, coal price, natural gas price and solar consumption have statistically significant coefficients in determining coal consumption at 1\% level of significance. There are no statistically significance for CDD, HDD, and GDP coefficients.

\textsuperscript{10} One lag report in Table 3-5 which represents to short run and static relationship.
The final stage of ARDL model is model diagnostic tests including serial correlation, stability, and normality. The results show that optimal lag length is two for all three time periods. In the ARDL model, in order to ensure models are proper to estimate, we need to examine diagnostic tests for normality, serial correlation, and stability. Table 3-6 provides these diagnostic tests for the ARDL model. The results show that there exist normality, no serial correlation, and no zero coefficients. Residual diagnostic using an F test where with a p value smaller than 0.05, we cannot reject the null hypothesis so that there is no serial correlation.

We apply the CUSUM and CUSUMSQ tests for stability of the model which proposed by Brown et al. (1975) to the residuals of the error-correction model. The graphical results of these tests are illustrated in Figure 3-9. For the most part, the plots of the CUSUM and CUSUMSQ statistics stay within the critical bounds indicating the stability of the estimated coefficients. Thus, the coal consumption function remains stable with no regard to the specific lag selection criterion. The CUSUM and CUSUMSQ tests suggest that the estimated parameters are stable over time since the plot of the two tests statistics fall within the boundaries and statically significant at 5%. Hence, it proves and emphasizes about existence of a dynamic long run linkage between dependent and explanatory variables.

Finally, the IRF analysis reveals two aspects: first the initial response to a shock in the other variables, and secondly, whether the shock impact persists or dies out in the short run (Soytas and Sari, 2006). Results of IRF for 2000 to 2016 are reported in panel A from Figure 3-10. The first graph shows that the response of coal consumption to coal consumption shocks. Coal consumption starts positive and declines sharply, reaching zero by period four. The response becomes positive again by period 6 and finally moves to negative after period 9. The response of coal consumption to a one standard deviation change in coal price is that it fluctuates
very little around zero and positive or negative. The response of coal consumption to natural gas price starts at zero and increases constantly and remains positive over all of the 10 time periods.

The response of coal consumption to GDP is that it is below zero initially and then positive briefly for periods 6 and 7 before going to negative to the end. The response of coal consumption to CDD is that it increases initially but later decreases and touches the zero line at period 8. The response becomes positive to the end. The response of coal consumption to HDD is that it is negative or zero until period 9 when it turns positive. The response of coal consumption to solar consumption is initially negative but later increases and touches the zero line at period 3. Then it peaks in period 5 before becoming negative in period 7. Later, it becomes positive and tries to move up after period 9. The response of coal consumption to seasonality variable shows that initially it is negative but it become positive at the period of 6, it turns to positive and finally downward to negative.

Results of IRF in the period of 2000-2008 are reported in the panel B from Figure 3-10. The impulse is coal consumption, the responses from natural gas price and HDD shocks have a positive movement but obvious fluctuation during period 6. The responses to coal price and GDP shocks are negative and or positive, and there are distinct effects on the period 2 and 4. The response to CDD is positive until period 7 and decreasing into negative over the final three periods. The responses to solar energy consumption and seasonal dummy changes by coal consumption show fluctuation below and above zero throughout the time periods.

The IRFs in the period 2009-2016 in panel C from Figure 3-10 show that the response of coal consumption to a one standard deviation change in coal price as initially zero or negative through period 4 and positive thereafter until the end of the time horizon. Response of coal consumption to natural gas prices starts at zero and then remains positive throughout the entire
period. The response of coal consumption to GDP changes also starts at zero and remains consistently negative throughout the entire period. The response of coal consumption to CDD changes is that it increases initially but later decreases and touches the zero line at period 7 and remains negative thereafter. The response of coal consumption to HDD changes is mostly negative or zero until period 8, after which it remains positive until the end. The response of coal consumption to solar consumption and the seasonal dummy variable both shows fluctuations between negative and positive throughout the entire period. Overall, the results provide evidence that there exists a consistency with the empirical results from ARDL model in the three periods of study.

6. Conclusions

In this study, we have estimated long-run and short-run responses of coal consumption in the U.S. energy market to changes in coal and natural gas prices, GDP, HDD, CDD, solar usage, and seasonality. Using monthly data from January 2000 to June 2016, the main purpose of these results is to reveal the dynamic linkages among the variables by using the ARDL approach to examine energy switching from coal to natural gas. These include the ARDL cointegration and error correction techniques to establish dynamic relationships among the variables. Implications from the bounds test method confirm the existence of cointegration among coal consumption, coal and natural gas prices, GDP, HDD, CDD, and solar energy consumption. Moreover, the empirical results from ECT coefficient show that a long-run relationship exists between the consumption of coal, coal price, the price of natural gas, economic growth, HDD, CDD, and solar consumption. During the entire time period, for instance, the ECT coefficient indicates that, if coal consumption drifts away from its long-run equilibrium level, it adjusts towards long-run equilibrium at 0.95 percent per month.
From the ARDL model, long-run results show that the impact on coal consumption from natural gas prices to be statistically significant and positive over 2000-2016, suggesting that these two energy inputs are substitutes. Based upon the long-run analyses, a 1% decrease in the price of natural gas will lead to decrease by 0.29% of coal consumption. Also, own price elasticity shows that a 1% increase in coal price will lead to decrease by 0.21% of coal consumption. In addition, it is observed that in the short-run and long run, coal consumption is inelastic with respect to both coal and natural gas prices. The effect of solar energy consumption and economic growth on coal demand are statically significant and negative in the entire period of study. As the income elasticity estimate shows, we can argue that coal is an inferior good. Therefore, increments in economic growth prompt a decrease in coal consumption and increases in other more efficient sources such as natural gas. Finally, the results of IRFs provide evidence that there exists a consistency with the empirical results from ARDL model.

The implications of this study for policy are as follows. First, U.S utilities are using more clean energy and substituting natural gas for coal, as this substitution is affected by both efficiency and the price of natural gas. The coal industry has been in crisis due to substitution of coal to natural gas. Coal market crisis has affected not only job market as well as local government income in the most coal producing states. Since the coal industry has been in crisis, one possible policy is to promote exporting of coal to other countries as a solution to overcome the industry’s crisis. Also, we conjecture that an important role has been played by natural gas price and effectiveness of the natural gas consumption in power plants.

Secondly, since there is no statistically significant, long run impacts from any variables on coal consumption in the post-shale gas time period, the expansion of shale gas production has substantially changed the factors that determine coal consumption. However, decreasing natural
gas prices compared to coal prices during this time period along with a statistically significant short-run impact might illustrate short term substitutions of natural gas for coal still exist. On the one hand, this substitution has resulted in important shifts in the U.S. energy market and reducing CO₂ emissions. On the other hand, this substitution implies that another proper policy could be encourage using coal for other uses (i.e. manufacturing or steel industry). These results are in line with previous studies (U.S. EIA, 2012; Nakicenovic, 2000; Hayhoe et al., 2002; Xie et al., 2016). Lastly, economic growth and maintaining the growth of the U.S economy has led to a path of consuming more low carbon energy and less coal. To maintain this energy path, energy policies such as renewable portfolio standards implemented by states around the U.S should aim at increasing the use of both natural gas and renewable energy.

These empirical results not only contribute to advancing the current literature, but also deserve certain attention from energy policy makers in the U.S. market. Since more than 90% coal has been used in power plants in the U.S, as capacities of power plants are decreasing and new capacities are being established to replace them, power firms are promoting to use more efficient and cost-effective energy sources (i.e. natural gas) to build new power plants over coal-fired plants. Given the coal issue of the energy market in the U.S., more study needs to be undertaken so as to offer a comprehensive resource for policymakers pursuing actual solutions to alleged problems surrounding this industry.
7. References


### Tables and Figures

#### Table 3-1: Descriptive statistics

<table>
<thead>
<tr>
<th>Variables</th>
<th>Coal Consumption (MBTU)</th>
<th>GDP (Million $)</th>
<th>Coal Price ($/MBTU)</th>
<th>Natural Gas Price ($/MBTU)</th>
<th>Heating degree-days (HDD)</th>
<th>Cooling degree-days (CDD)</th>
<th>Solar consumption (TBTU)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>1700044</td>
<td>14520.82</td>
<td>1.04</td>
<td>2.85</td>
<td>362.1</td>
<td>132.15</td>
<td>11.41</td>
</tr>
<tr>
<td>Std.Dev</td>
<td>150045.5</td>
<td>1169.42</td>
<td>0.53</td>
<td>1.24</td>
<td>316.61</td>
<td>150.06</td>
<td>11.37</td>
</tr>
<tr>
<td>Minimum</td>
<td>845520</td>
<td>12294.89</td>
<td>0.442</td>
<td>1.014</td>
<td>2.98</td>
<td>1</td>
<td>2.91</td>
</tr>
<tr>
<td>Maximum</td>
<td>2129983</td>
<td>16659.99</td>
<td>3.78</td>
<td>7.17</td>
<td>995.45</td>
<td>621</td>
<td>58.66</td>
</tr>
</tbody>
</table>

Note: Number of observations = 198, All variables are converted to natural logarithms

#### Table 3-2: Tests for unit root test

<table>
<thead>
<tr>
<th>Log of variables</th>
<th>Level</th>
<th>First difference</th>
<th>Zivot-Andrews</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ADF</td>
<td>PP</td>
<td>ADF</td>
</tr>
<tr>
<td>Coal consumption</td>
<td>0.479</td>
<td>-3.97**</td>
<td>-2.267**</td>
</tr>
<tr>
<td>Coal price</td>
<td>-1.871</td>
<td>-2.001</td>
<td>-13.585**</td>
</tr>
<tr>
<td>Natural gas price</td>
<td>-2.55</td>
<td>-2.795</td>
<td>-12.97**</td>
</tr>
<tr>
<td>GDP</td>
<td>-0.708</td>
<td>-0.769</td>
<td>-18.465**</td>
</tr>
<tr>
<td>HDD</td>
<td>-2.026</td>
<td>-6.080**</td>
<td>-14.88**</td>
</tr>
<tr>
<td>CDD</td>
<td>-0.207</td>
<td>-1.862</td>
<td>-11.411**</td>
</tr>
<tr>
<td>SC</td>
<td>-2.76</td>
<td>-0.98</td>
<td>-13.56**</td>
</tr>
</tbody>
</table>

Notes: *p<0.10, **p<0.05, ***p<0.01. For Zivot-Andrews test; the critical values at 1%, 5% and 10% level are -4.93, -4.42 and -4.11 respectively
Table 3-3: Results of F bound test

<table>
<thead>
<tr>
<th>Study period</th>
<th>F Statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total period (2000-2016)</td>
<td>22.012</td>
</tr>
<tr>
<td>Pre shale gas revolution (2000-2008)</td>
<td>9.635</td>
</tr>
<tr>
<td>Post shale gas revolution (2008-2016)</td>
<td>6.52</td>
</tr>
</tbody>
</table>

Note: Critical value of bounds for all time periods are 2.62 bottom and 3.79 top.

Table 3-4: Estimated long run coefficients

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>34.89***</td>
<td>17.07***</td>
<td>28.45***</td>
</tr>
<tr>
<td></td>
<td>(3.25)</td>
<td>(2.52)</td>
<td>(9.50)</td>
</tr>
<tr>
<td>Coal price (cp)</td>
<td>-0.21***</td>
<td>-0.05**</td>
<td>-0.01</td>
</tr>
<tr>
<td></td>
<td>(0.019)</td>
<td>(0.03)</td>
<td>(0.03)</td>
</tr>
<tr>
<td>Natural gas price (gp)</td>
<td>0.29***</td>
<td>0.05***</td>
<td>0.002</td>
</tr>
<tr>
<td></td>
<td>(0.03)</td>
<td>(0.01)</td>
<td>(0.02)</td>
</tr>
<tr>
<td>GDP</td>
<td>-0.41***</td>
<td>0.08</td>
<td>0.015</td>
</tr>
<tr>
<td></td>
<td>(0.09)</td>
<td>(0.07)</td>
<td>(0.02)</td>
</tr>
<tr>
<td>HDD</td>
<td>-0.09***</td>
<td>-0.04*</td>
<td>-0.02</td>
</tr>
<tr>
<td></td>
<td>(0.01)</td>
<td>(0.03)</td>
<td>(0.02)</td>
</tr>
<tr>
<td>CDD</td>
<td>-0.03***</td>
<td>-0.07**</td>
<td>-0.01</td>
</tr>
<tr>
<td></td>
<td>(0.009)</td>
<td>(0.02)</td>
<td>(0.02)</td>
</tr>
<tr>
<td>SC</td>
<td>-0.02***</td>
<td>0.02</td>
<td>0.06</td>
</tr>
<tr>
<td></td>
<td>(0.02)</td>
<td>(0.05)</td>
<td>(0.08)</td>
</tr>
<tr>
<td>SIDV</td>
<td>0.06***</td>
<td>0.01</td>
<td>-0.02</td>
</tr>
<tr>
<td></td>
<td>(0.01)</td>
<td>(0.02)</td>
<td>(0.04)</td>
</tr>
<tr>
<td>Adjusted R²</td>
<td>0.72</td>
<td>0.80</td>
<td>0.62</td>
</tr>
<tr>
<td>F Statistics</td>
<td>23.62</td>
<td>20.33</td>
<td>7.99</td>
</tr>
<tr>
<td></td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
</tr>
<tr>
<td>Durbin-Watson</td>
<td>2.07</td>
<td>1.84</td>
<td>2.14</td>
</tr>
</tbody>
</table>

Note: All variables are measured as natural logs and *p<0.10, **p<0.05, ***p<0.01. Standard errors are in parenthesis.
Table 3-5: Estimated short run coefficients

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>0.003</td>
<td>-0.001</td>
<td>-9.13</td>
</tr>
<tr>
<td></td>
<td>(0.004)</td>
<td>(0.005)</td>
<td>(15.94)</td>
</tr>
<tr>
<td>Δcp(-1)</td>
<td>-0.13*</td>
<td>-0.05</td>
<td>-0.71***</td>
</tr>
<tr>
<td></td>
<td>(0.07)</td>
<td>(0.06)</td>
<td>(0.013)</td>
</tr>
<tr>
<td>Δgp(-1)</td>
<td>0.12***</td>
<td>0.03</td>
<td>0.223***</td>
</tr>
<tr>
<td></td>
<td>(0.04)</td>
<td>(0.03)</td>
<td>(0.012)</td>
</tr>
<tr>
<td>ΔGDP(-1)</td>
<td>-1.85*</td>
<td>-0.19</td>
<td>-2.52</td>
</tr>
<tr>
<td></td>
<td>(0.83)</td>
<td>(0.67)</td>
<td>(1.63)</td>
</tr>
<tr>
<td>ΔHDD(-1)</td>
<td>-0.06***</td>
<td>-0.06***</td>
<td>-0.04</td>
</tr>
<tr>
<td></td>
<td>(0.008)</td>
<td>(0.01)</td>
<td>(0.02)</td>
</tr>
<tr>
<td>ΔCDD(-1)</td>
<td>-0.005</td>
<td>-0.06***</td>
<td>-0.002</td>
</tr>
<tr>
<td></td>
<td>(0.008)</td>
<td>(0.02)</td>
<td>(0.01)</td>
</tr>
<tr>
<td>ΔC(-1)</td>
<td>0.69***</td>
<td>1.17***</td>
<td>0.49***</td>
</tr>
<tr>
<td></td>
<td>(0.13)</td>
<td>(0.31)</td>
<td>(0.18)</td>
</tr>
<tr>
<td>ΔSC(-1)</td>
<td>-0.28***</td>
<td>0.23</td>
<td>-0.39***</td>
</tr>
<tr>
<td></td>
<td>(0.06)</td>
<td>(0.14)</td>
<td>(0.12)</td>
</tr>
<tr>
<td>ECT(-1)</td>
<td>-0.95***</td>
<td>-1.18***</td>
<td>-0.76***</td>
</tr>
<tr>
<td></td>
<td>(0.05)</td>
<td>(0.20)</td>
<td>(0.19)</td>
</tr>
<tr>
<td>Adjusted R²</td>
<td>0.62</td>
<td>0.67</td>
<td>0.68</td>
</tr>
<tr>
<td>F Statistics</td>
<td>19.53</td>
<td>12.20</td>
<td>10.95</td>
</tr>
<tr>
<td></td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
</tr>
<tr>
<td>Durbin-Watson</td>
<td>2.31</td>
<td>2.38</td>
<td>2.25</td>
</tr>
</tbody>
</table>

Observations: 195, 102, 96

Note: All variables are measured as natural logs and *p<0.10, **p<0.05, ***p<0.01. Standard errors are in parenthesis.

Table 3-6: Diagnostic tests statistics

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>F Statistics (Wald test)</td>
<td>18.51</td>
<td>19.38</td>
<td>2.91</td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.00)</td>
</tr>
<tr>
<td>Durbin-Watson</td>
<td>2.07</td>
<td>1.84</td>
<td>2.14</td>
</tr>
<tr>
<td>Serial Correlation</td>
<td>4.76</td>
<td>1.21</td>
<td>1.93</td>
</tr>
<tr>
<td></td>
<td>(0.11)</td>
<td>(0.29)</td>
<td>(0.16)</td>
</tr>
<tr>
<td>Normality</td>
<td>0.49</td>
<td>0.54</td>
<td>0.23</td>
</tr>
<tr>
<td></td>
<td>(0.75)</td>
<td>(0.31)</td>
<td>(0.17)</td>
</tr>
<tr>
<td>Stability (CUSUM test)</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Observations: 198, 108, 90

Note: Wald test represents for coefficient diagnostics. J-B test null is normality. B-G test null is no serial correlation up to the selected lag. Parentheses are p-value.
Figure 3-3: Economic Growth (US$)

Figure 3-4: Coal consumption (MBTU)

Figure 3-5: Coal – natural gas prices ($/MBTU)

Figure 3-6: Heating degree-days
Figure 3-7: Cooling degree-days

Figure 3-8: Solar consumption (TBTU)

Figure 3-9: Plot of the cumulative sum of recursive residuals (CUSUM)
Figure 3-10: Impulse response function (combined)

Panel A: 2000 to 2016 Data\textsuperscript{11}

Panel B: 2000 to 2008 Data

\textsuperscript{11} For all three panels, LCC is on the y axis and variable shocks are on the x axis.
Panel C: 2009 to 2016 Data

Response of LCC to LCC

Response of LCC to LCP

Response of LCC to LGP

Response of LCC to LGDP

Response of LCC to LCDD

Response of LCC to LHDD

Response of LCC to LSC

Response of LCC to SIDV

Response to Cholesky One S.D. Innovations ± 2 S.E.
1. Introduction

Urbanization refers to the gradual increase in the proportion of people living in urban areas and the increasing of human populations into discrete areas which leading to transformation of land for residential, commercial, industrial, and transportation purposes (Henderson, 2003). In fact, urban activities and energy use in urban areas have become the principal sources of CO₂ emissions. Moreover, the urbanization progress has caused increased natural resource use, fossil fuels consumption, and more impacts on the environment and ecosystem. As shown in Figure 4-1, the United Nations Convention on Climate Change indicates that 75% of world population will be living in cities and suburb areas by the year 2030 (UN, 2014). Also, this report makes clear that CO₂ emissions are decreasing in industrialized countries and increasing in the rest or world in recent decades (Figure 4-2).

Figure 4-1: Urban and rural populations of the world, 1950–2050

Figure 4-2: CO₂ emissions from 1750 and projected to 2050
Due to concerns about global climate change, the general public has an interest in policies that advocate limiting the rise of CO\textsubscript{2} emissions. Studies of urban environmental transition emphasize that urbanization can have both negative and positive effects on the environment making it difficult to determine the net effect a priori (Cole and Neumayer, 2004; Liddle and Lung, 2010). Examining the linkages among CO\textsubscript{2} emissions, urbanization and other factors in the U.S. is significant as the U.S. is the second largest aggregate emitter in the globe in 2016, based upon its share of global energy-related CO\textsubscript{2} emissions (International Energy Agency, [IEA], 2016). The availability of state level data from U.S. EIA and the Carbon Dioxide Information Analysis Center [CDIAC] is the basic information to study CO\textsubscript{2} emissions, urbanization, energy consumption, and socio-economic variables in the U.S. This research investigates how urbanization and other variables reflecting energy use impact carbon emissions in the U.S.

Yet, at the same time, a growing body of literature now implies that urbanization holds the key factor to sustainability since, as a multidimensional socio-economic and environmental process, it is highly carbon-efficient at the local level in the developed countries. Hence, urbanization growth reduces environmental costs of infrastructure per capita, facilitates alternative transit use, increases energy efficiency, encourages distributed use of solar energy, and stimulates environmental awareness - all in ways that could lead to decrease the CO\textsubscript{2} emissions. Some researchers conclude that more urbanized communities create higher CO\textsubscript{2} emissions in the globe (Poumanyvong and Kaneko, 2010; Hossain, 2011; Kindle and Shawhan, 2011; Zhang and Lin, 2012; Leitao and Shahbaz, 2013; Liddle, 2014; Liddle, 2014). Other scholars working at the local and state levels have also argued that urbanization leads to less CO\textsubscript{2} emissions (Newman and Kenworthy, 1999, Andrews, 2008; Gonzalez, 2009, Bettencourt et al.,
2010; Glaeser, 2011). Together, this literature argues that urbanization’s impact on CO$_2$ emissions has advantages locally but drawbacks globally. Estimates of CO$_2$ emissions represent energy-related which is based upon state-level energy use and the consumption of fossil fuels (U.S. EIA, 2016). The crucial reason to utilize energy-related estimate is due to high cost to monitor wide variety of static and dynamic sources of CO$_2$ emissions (Auffhammer and Steinhauser, 2007).$^{12}$

Furthermore, a controversial discourse has recently advanced about the subject of whether CO$_2$ emissions measurement must be based upon space-related consumption or production method (Davis et al. 2011; Peters et al. 2012; Steininger et al. 2014; Liu, 2015; Fan et al. 2016; Fernandez-Amador et al. 2017). Production-based accounting (PBA) of CO$_2$ emissions for fossil fuel consumption estimates the greenhouse gas emissions from all the coal, crude oil, and gas consumed in a state by power plants, residential, and industrial production of goods and services.

However, PBA has multiple drawbacks. First, it excludes CO$_2$ emissions derived from international trade and transportation. Since such CO$_2$ emissions do not occur within a specific state, allocation of these emissions to specific territories is complicated. Second, industries with energy-intensive in states with CO$_2$ emission regulations and taxes may move into other states with fewer regulations and lower costs of energy. However, the goods and services produced in the less regulated states could then be transported to the more regulated states. Hence, reducing CO$_2$ emissions in one state might be directly correlated to rising CO$_2$ emissions in the other states due to CO$_2$ emissions leakage. Thus, one state's production can also be driven by consuming in other states.

$^{12}$ The top five non-energy sources of climate change are: Deforestation, Methane from livestock, Soil carbon, Landfills and wastewater, and Permafrost.
A consumption-based accounting (CBA) is an approach that considers these issues. It subtracts from states all CO$_2$ emissions that are incorporated in exported goods and services, including CO$_2$ emissions of transportation, and the contained CO2 emissions in the inventories of the importing states (Aldy, 2005; Auffhammer and Carson, 2008; Peters et al., 2011; Fan et al., 2016). Hence, with respect to PBA, low emission states may be less clean in the CBA approach as high CO$_2$ emissions states could be producing goods that are transported for consumption in low CO$_2$ emission states. Econometric analyses of PBA CO$_2$ emissions can find that CO$_2$ emissions are primarily driven by city size, population, GDP, and the energy consumption of an economy.

More importantly, it can be argued that not all energy used to produce goods and services in one state are consumed in that same state. For example, West Virginian generates more than 95 percent of its electricity from coal-fired power plants while exporting a large portion of electricity to other neighboring states (IER, 2013). This creates a likely mechanism for the spatial spillover effects of CO$_2$ emissions. More explicitly, it is argued that there exists a spatial dependence among the forces of energy-related emissions and key economic factors which cross state levels. Moreover, it is hypothesized that transportation of goods and services inter-state and across the country has spillover impacts in state-level. In particular, recent studies address the spatial correlation between the factors of emissions interstate. Examples include U.S. EIA (2016) and Burnett et al. (2013). Thus, since economic activities and urban area distances matter, consideration of spatial spillovers in energy use is a chief motivation in this study.

To address spatial spillovers, we employ several methods that complement both PBA and CBA approaches to investigate the contradiction of urbanization and CO$_2$ emissions. These contradictions appear not just from a variety of scales of analysis (i.e. state versus national), but
also from limitations in how previous studies have theorized and estimated the factors of CO₂ emissions and urbanization. To address these shortcomings, we synthesize insights from non-spatial and spatial econometric methods to develop a framework for understanding how urbanization and energy factors affect CO₂ emissions across 48 continental states and District of Columbia. Hence, spatial linkages are examined by determining spatial panel data models that control for spatial impacts over time and space (LeSage and Pace, 2009; Aldy, 2005). Indeed, estimated parameters might be biased when spatial correlation is not taken into account (Kalenkoski and Lacombe 2006).

Using spatial econometric models is an appropriate approach to study the effects of urbanization, per-capita gross state product (GSP), energy use, energy prices, and coal consumption on CO₂ emissions at the state level from 2000 to 2015. Additionally, the spatial reliance signifies that an energy policy implemented in one state could have spillover impacts on neighboring states (Kindle and Shawhan, 2011). The estimate of such spillovers is significant to determine the direct and indirect impacts of state-level policies adopted in the U.S. which affect the level of CO₂ emissions. Hence, this study attempts to control for spatial correlation in estimating the impact of urbanization and other driving forces on CO₂ emissions.

This research makes contributions to the literature through its use of a spatial panel data approach. More specifically, we have taken into consideration urbanization and energy-related factors on CO₂ emissions when examining state-level panel data for 2000-2015. This time period captures recent developments in state-level economic growth and energy use policies, like passage of renewable portfolio standards by states. Looking ahead, our findings suggest that a statistically significant spatial dependence among the key factors and CO₂ emissions at the state-level in the U.S.
The remainder of this paper is organized as follows. Section 2 surveys the relevant literature on urbanization, energy consumption, and CO₂ emissions. Section 3 describes theory related to driving forces of CO₂ emissions. Section 4 is about empirical models. The next section describes data. Section 6 presents the results and reports empirical findings in the paper. Finally, section 7 concludes this study.

2. Literature Review

Urbanization is often discussed in the context of urban economics and it is a demographic indicator that increases with population density and transforms the organization of human behavior, affecting energy consumption and CO₂ emission patterns (Jorgensen et al., 2014). Previous research has applied various time series and panel data models to examine the relationships between economic activities, urbanization level, and CO₂ emissions (Martínez-Zarzoso et al., 2007; Parshall et al., 2009; Poumanyvong and Kaneko, 2010; Hossain, 2011; Kindle and Shawhan, 2011; Zhang and Lin, 2012; Leitao and Shahbaz, 2013; Liddle, 2014; Sadorsky, 2014; Xu and Boqiang, 2015). These previous studies provide the extensive literature reviews on urbanization, the factors of energy, environmental pollution and CO₂ emissions.

Rising CO₂ emissions are both an environmental and public health concern. CO₂ emissions and environmental pollution have grown due to industrialization, fossil fuel consumption, and urbanization that have become main major environmental problems not only for developed economies but severely in the developing countries. Consequently, CO₂ emissions have affected climate via global warming and impacted individuals’ health condition in terms of disease and life expectancy (Galor and Weil, 2000; Schafer and Van Basshuysen, 2013). It is evident that global CO₂ emissions from fossil fuel consumption and from urban development do play a key part in the expansion of global warming as well as ozone exhaustion (Canadell et al.,
2007). Previous research mostly has applied non-spatial econometric models and have been criticized because of omitted distance variables bias (LeSage and Pace, 2009; Elhorst, and Vega, 2013). However, recent studies have investigated the spatial dependencies between carbon dioxide emissions, energy consumption, and economic growth in a spatial framework (Chuai, et al., 2012; Yu, 2012; Burnett et al., 2013).

The issue of how urbanization and energy use affect CO₂ emissions has been studied an empirical question. Existing literature shows that the principal forces driving U.S. CO₂ emissions come from urbanization, transportation, coal-fired power plants, and energy use (Aldy, 2005; Carson, 2010; Burnett et al., 2013). These key factors have had positive impacts on CO₂ emissions, but their effects have been cautiously decreasing over the past decade (U.S. EIA, 2016). Some studies have revealed that there is a positive relationship between urbanization and CO₂ emissions (York, 2003; Sadorsky, 2014). Burnett et al. (2013) used a long-run panel data set in fixed- effect and spatial models to examine the relationship between U.S. state-level CO₂ emissions and economic activity. They show that economic distance plays a crucial role in intra-state and inter-state emissions; and there exist statistically significant, positive economic spillovers and negative price spillovers to state-level emissions in the United States. As stated by Chuai et al. (2012) and Yu (2012) the relationships between energy consumption, urbanization and CO₂ emissions have examined even though the studies are not clear about how urbanization affects the environmental quality of neighboring regions. The magnitudes and signs of the impacts are not comparable because of various data sets, assumptions, and estimation methodologies (Sadorsky, 2014). In one of the earliest studies, Parikh and Shukla (1995) use a data set of 83 developed and developing countries for the year 1986 to investigate the effect of
urbanization level on energy consumption and emissions. Their empirical results show that urbanization has a positive and significant effect on CO₂ emissions.

In addition, there are also some studies in the related literature. For instance, Martinez-Zarzoso and Maruotti (2011) study the relationship between urbanization and CO₂ emissions using the Stochastic Impacts Regression and Population, Affluence and Technology (STIRPAT) model in a panel of 88 developing countries over the period 1975 to 2003. The empirical results show an inverted-U shaped linkage between urbanization and CO₂ emissions. This paper relates to the previous studies by using a semi-parametric mixture model that allows for unknown distributional shapes and endogenously classifies countries into homogeneous groups. The results show that urbanization's effect on CO₂ emissions are different among the country groups. Hossain (2011) studies the dynamic causal relationships between CO₂ emissions, energy use, economic growth, trade openness, and urbanization level for the panel of newly industrialized countries using the time series data for the period 1971–2007. The Granger (1988) causality test results supported that there is no evidence of long-run causal relationships between unidirectional short-run causal, economic growth, and trade openness to CO₂ emissions. In addition, Sadorsky (2014) applies developed panel regression approaches that allow for heterogeneous slope coefficients and cross-section dependence to model the effect that urbanization has on CO₂ emissions for a panel of emerging economies. Zhang and Lin (2012) investigate urbanization and CO₂ emissions in China. They show that the urbanization affect more energy use across regions in China. The researchers investigate the impact of urbanization growth on CO₂ emissions and focus on regional differences. They use a stochastic IPAT- impact of human activity on the environment- model which is STRIPAT approach using provincial panel data from 1995 to 2010. The findings reveal that urbanization raises energy use and CO₂
emissions gradually in provinces of China. Moreover, the results show that the impacts of urbanization on energy use vary across regions and decrease constantly from the western region to the central and eastern regions in China.

Furthermore, Poumanyvong and Shinji (2010) examine linkage between urbanization, energy consumption, and CO$_2$ emissions with highlighting on income levels. They also used the STIRPAT approach employing panel data for 99 countries from 1975 to 2005. It is clear that the relationship between CO$_2$ emissions and urbanization is complex and it is not only a linear link between variables. Xu and Boqiang (2015) investigate the effects of industrialization and urbanization on CO$_2$ emissions in 30 provinces of China using nonparametric additive econometric methods and regional panel data from 1990 to 2011. This study is a comprehensive research not only about urbanization but also consideration of industrialization factors for environmental impacts. The findings indicate that there exists an inverted U-shaped linkage between industrialization and CO$_2$ emissions in the three provinces in China. In fact, urbanization caused an inverted U-shaped model with CO$_2$ emissions in the eastern region, and a positive U-shaped trend in the central provinces. Hence, one of the most important aspects of this study is using a dynamic approach to determine the impact of urbanization and energy factors on CO$_2$ emissions in China. However, this study did not use panel data model features such as fixed or random effects.

The issue of environmental emissions has also become the subject matter of scholarly research that address main drivers of CO$_2$ emissions. Some recent studies indicate that the consumption of energy is as a basic determinant of CO$_2$ emissions through estimating an emissions-energy-output (EEO) model (Ang, 2007). However, Itkonen (2012) criticizes the EEO model approach by examining the impact of income on CO$_2$ emissions when energy use is a
positive linear function of income. Itkonen (2012) argues that income variable included in the EEO model can only explain the carbon intensity\textsuperscript{13} of energy consumption and not the level of CO\textsubscript{2} emissions. Moreover, Jafarullah and King (2017) investigate a model, using the EEO framework to explore determinants of CO\textsubscript{2} emissions. They extend Itkonen's (2012) critique of the EEO model in several ways. First, they state that Itkonen’s remarks regarding estimation bias are critical to the functional form of the linkage between energy use and income. Second, they highlight the conceptual dependence between energy use and CO\textsubscript{2} emissions data from Itkonen's (2012) model. This, in turn, implies that the coefficient on the energy consumption variable should equal positive one when all related factors of CO\textsubscript{2} emissions are included in the model. As the estimates found for this coefficient often diverge markedly from unity, this suggests that most of empirical EEO models are misspecified in some way. Finally, they conclude that under a plausible nonlinear model to Itkonen’s (2012) model, the consumption of energy and income variables are fundamental factors in the model.

Majority of previous studies assume that inter-jurisdiction local regions to be independent throughout cross-sections so that spatial interactive impacts are disregarded. Ultimately, local region’s attributes could rely on its neighbors, for example, urban areas in one state utilize electricity generated in another state. Thus, disregarding spatial dependency may lead to a spurious model or estimate biased parameters in OLS approach (Anselin, 2002; Lesage and Pace, 2009). Burnett et al. (2013) used a spatial econometric method to determine US energy emissions using state-level data. They incorporated spatial forces driving energy pollution in the U.S. Although the authors utilize spatial models, SDM model and spatial dependence have not investigated by various data generating process. Therefore, it is adequate to examine the impacts of urbanization and energy factors on environment with a variety of econometric and spatial

\textsuperscript{13} The ratio of CO\textsubscript{2} emissions to energy consumption.
econometric approaches. This study explores to contribute the literature by adding new data and more variables by using both non-spatial and spatial models.

3. Theoretical Model

Urban areas are the center of energy consumption and the consumption of fossil fuels in transportation, electricity and industrial goods, all of which generate CO$_2$ emissions. However, many power plants and industrial firms are located in rural regions and they burn fossil fuels - emitting a high amount of CO$_2$ and pollute entire environment. It has been theoretically shown that an increase in energy use occurs in urban areas related to higher economic activities and income levels (Canadell et al., 2007). Moreover, Itkonen (2012) assumes that energy consumption and income are linearly associated. Itkonen (2012) modifies the EEO model by examining the effect of income on CO$_2$ emissions when the consumption of energy has a positive linear relationship with income. Also, Jaforullah and King (2017) model CO$_2$ emissions, using the EEO approach to identify determinants of CO$_2$ emissions. The EEO model, although, indicates that energy consumption is a function of income, energy use and income are two critical exogenous variables of CO$_2$ emissions. However, based upon their model, they concluded that the consumption of energy was not an independent determinant of CO$_2$ emissions (Jaforullah and King, 2017).

We argue that the energy consumption can be used as an independent variable to explain CO$_2$ emissions. First, most of the empirical studies on the relationships between energy use, GDP, and CO$_2$ emissions have included energy consumption as a main independent variable (Ang, 2007; Soytas, et al., 2007; Ang, 2008; Martinez-Zarzoso, 2008; Zhang and Cheng, 2009; Halicioglu, 2009; Soytas and Sari, 2009; Poumanyvong and Shinji, 2010; Acaravci and Ozturk, 2010; Hossain, 2011; Wang, et al., 2011; Chuai et al., 2012; Yu, 2012; Gonzalez and Martinez,
2012; Saboori and Sulaiman, 2013; Sadorsky, 2014). This literature has focused on the linkages between CO₂ emissions, income and energy consumption by utilizing energy use variables as explanatory variables. Second, a single study by Jaforullah and King (2017) may not be generalizable to ignore energy consumption effect on CO₂ emissions. Lastly, the spatial spillover effects of both income and energy consumption on CO₂ emissions in other states will be explored in this research, hence, both variables are included in the model below.

The idea of spatial spillovers among economic activities is related to the concept of economic distance, which indicates that the closer two locations are to one another in a geographic distance, the more possibility that their economies are interconnect (Conley and Ligon, 2002). Spatial linkages suggest that policies adopted in one region would impact the whether or not policies are implemented in neighboring regions. Thus, for example, a U.S. state is more likely to adopt a law and or policy if its neighboring states have already done so (Mooney, 2001). In fact, geographical location has been identified as a critical factor of cross-region economic growth due to indicators like the diffusion of technology (Keller, 2004). We could argue that CO₂ emissions may decrease with technological development, then the diffusion of technology would likely enhance conditions of neighboring environment.

Furthermore, it is considered that spatial panel data models have divided in two categories. One is a non-dynamic, which has utilized in the context of forecasting recently (Elhorst, 2009; Baltagi et al., 2012). Another spatial panel is dynamic which controls for either time-invariant heterogeneity across geographical areas or spatial autocorrelation between areas (Anselin, 2001; Anselin et al., 2008). Since the main empirical aim of this study is to precisely model the main drivers of CO₂ emissions in the U.S’ state level, we formulate the dynamic spatial panel data models. Given the theoretical concepts cited literature above, we can formulate
an equation as follows:

\[ y_i = (X_i, G_i, RE_i) \]  \hspace{1cm} (4-1)

The dependent variable \( y_i \) represents per capita CO\(_2\) emissions in state \( i \), \( X_i \) is an \( n \times k \) matrix of \( k \) state-level characteristics such as, urbanization rate, GSP, the consumption of energy, coal consumption, energy prices, and RPS in state \( i \) and time period \( t \). \( G_i \) represents the geographical characteristics (i.e. longitude and latitude) of the sample and \( RE_i \) includes a dummy variable to describe a renewable policy proxy for state level.

The linkages between energy use, urbanization, and CO\(_2\) emissions will be examined while controlling for possible spatial impacts in the panel data. To account for this insight a term for spatial linkage is taking into account association between states, arguing that there exist possible spatial linkages between state-level economic activities and energy use, which in turn produce CO\(_2\) emissions (Auffhammer and Carson, 2008). Spatial linkages can be constructed in various ways, depending on the relationship between the dependent variable and the explanatory variables. When specifying interaction between spatial units, the model may contain a spatially lagged dependent variable or a spatial autoregressive process for the error term, including the spatial autoregressive model (SAR), spatial lag model (SLM), or a spatial error model (SEM) (Anselin et al., 2008). The spatial Durbin model (SDM) contains a spatially lagged dependent variable and spatially lagged independent variables (LeSage and Pace, 2009). All spatial models have a weight matrix (W), which quantifies the spillover between regions. Elhorst (2014) names the weight matrix as a tool to explain the spatial linkages of the geographical units in the sample. There are variety of units of measurement for spatial dependency such as neighbors, distance, and links (Getis, 2007). In this study, a contiguity weight matrix is utilized for each spatial model. This matrix specifies the value of one between adjacent states and zero between states.
with no common border.

The main factors which lead to spillover impacts across states allow us to include spatial lags. First, there exist more than 3200 fossil fuel–fired power plants within all states in the U.S. (EIA, 2016) and these power plants are likely to spillover emissions to neighboring states. The second factor giving rise to spatial lags is urbanization and its impact on the own state and neighboring state. The third factor is inter-state of goods transportation which lead to consider spillover effects of pollution across states for energy consumption. These would suggest that state-level of urbanization and energy use would lead to state spillover impacts of CO₂ emissions by spatial lags. The spatial dependence method provides the theoretical basis for the literature on interstate level impacts of CO₂ emissions. Anselin (2002) states two important motivations for considering spatial impacts in regression models from a theory-driven as well as from a data-driven perspective. A theory-driven framework follows from the formal specification of spatial interaction, which are interacting agents and social interaction in an econometric model, where an explicit interest in the spatial interaction of a particular variable conquers and a theoretical model creates the model specification.

If we have two regions \( i \) and \( j \) which are spatially correlated and supposed error terms have normality, then:

\[
\begin{align*}
\nu_i &= \alpha_i \nu_j + x_i \beta + \epsilon_i \\
\nu_j &= \alpha_j \nu_i + x_j \beta + \epsilon_j \\
\epsilon_i &\sim N(0, \sigma^2 x^2) \quad i = 1 \\
\epsilon_j &\sim N(0, \sigma^2 x^2) \quad j = 2
\end{align*}
\]

Where the dependent variable in neighbor i influences the dependent variable in neighbor j and vice versa. In non-spatial models each observation has a mean of \( x_i \beta \) and a random component \( \epsilon_i \) where the observation i represents a location or point in space at one location and
is considered to be independent of observations in other locations. In other words, statistically independent observations imply that $E(\varepsilon_i \varepsilon_j) = E(\varepsilon_i)E(\varepsilon_i) = 0$. This assumption of independence greatly simplifies models. However, a spatial approach implies that each observation corresponds to a location or region (LeSage, 2009).

4. Empirical Model

In this study, we examine the linkages between energy factors, urbanization, and CO$_2$ emissions while controlling for possible spatial impacts in the panel data. In order to test and estimate the proper spatial model, this study selects an ordinary panel model, spatial lag model (SAR), spatial error model (SEM) and the spatial Durbin model (SDM). In addition, the SDM is utilized to determine direct and indirect effects which contains a spatially lagged dependent variable and spatially lagged independent variables (LeSage and Pace, 2009). According to the variables that described, we can have a main empirical model to estimate which is given by the following:

$$\ln(c_{it}) = \beta_0 + \beta_1 \ln(u_{rt}) + \beta_2 \ln(\varepsilon_{ct}) + \beta_3 \ln(\varepsilon_{ct}) + \beta_4 \ln(g_{it}) + \beta_5 \ln(t_{it}) + \beta_6 r_{ps_{it}} + \mu_i + \eta_t + \varepsilon_{it}$$

$i = 1 ... N; t = 1 ... T$  \hspace{1cm} (4-3)

Where $c_{it}$ represents to per capita CO$_2$ emissions, $u_{rt}$ is urbanization rates, $\varepsilon_{ct}$ represents to total energy consumption, $\varepsilon_{ct}$ represents the consumption of coal, $g_{it}$ is per capita Gross State Product (GSP), $t_{it}$ denotes energy prices and $r_{ps_{it}}$ is dummy variable for renewable energy policy which is renewable portfolio standard (RPS) for state $i$ over time $t$. $\beta_1, \beta_2, \beta_3, \beta_4, \beta_5, \beta_6$ are coefficients of explanatory variables. Moreover, all variables are converted to natural
logarithms in order to interpret the coefficients as elasticities.\(^{14}\) The parameter \(\mu_i\) denotes the entity effect (or heterogeneity) for each state and \(\eta_t\) denotes a common time fixed effect. The entity effect as fixed meaning that it is assumed that this variable is correlated with the explanatory variables and approximately fixed over time for each state within the sample. If the empirical model estimate with no controlling for the entity effect, then estimation might result in omitted variable bias while the fixed effect is correlated with the explanatory variable. The entity effect can be interpreted as characteristics within states that do not change over time such as unobservable geographic characteristics such as lake and rivers. The time period impacts control for time-specific shocks that influence all states within a given period of time; e.g., federal policies on renewable energy production tax credits that impact \(\text{CO}_2\) emissions throughout the U.S.

Following the previous discussion, there are three main types of spatial panel specifications to explore spatial relationships between variables. Consistent with the insight of a simple spatial method, the Spatial Autoregressive Regression (SAR) model is:

\[
Y_{it} = \rho \sum_{j=1}^{N} W_{ij} Y_{jt} + X_{it} \beta + \mu_i + \eta_t + \varepsilon_{it}
\]

\[i = 1, 2, \ldots, N\]

\[t = 1, 2, \ldots, T\]

\[\varepsilon \sim N(0, \sigma^2 I_n)\]

where \(y_{it}\) is a \((n \times 1)\) vector of \(\text{CO}_2\) emissions, \(\sum_{j=1}^{N} W_{ij} y_{jt}\) is the pre-specified \(n \times n\) matrix of spatial interaction impacts, \(\rho\) is a spatial autocorrelation coefficient, \(X\) is a matrix of explanatory variables including urbanization rate, energy prices, per capita GSP, RPS, the consumption of coal, and energy use, \(\beta\) is a \(1 \times k\) vector of parameters to be estimated, and \(\varepsilon\) is a vector of errors.

\(^{14}\) Since coal consumption (cc) has zero values for several states over the 2000 to 2015 time period, we add 1.0 to all values of this variable so that with transformation, coal consumption is \(\text{Ln} (cc+1)\).
(Anselin, 2002).

The Spatial Error Model (SEM) mode is:

\[ Y_{it} = X_{it} \beta_{it} + \mu_{it} + \eta_{it} + \phi_{it} \]
\[ \phi_{it} = \lambda \sum_{j=1}^{N} W_{ij} \phi_{jt} + \epsilon_{it} \]
\[ i = 1,2,...,N \]
\[ t = 1,2,...,T \]
\[ \epsilon \sim N(0, \delta^2 I_n) \] (4-5)

where \( \lambda \) is a spatial parameter similar to \( \rho \) in SAR model and all other notations are as previously. The estimate results in the spatial models (SAR and SEM) show that the spatial coefficients (lambda and rho) are statistically significant with contiguity based matrix, justifying the use of spatial econometric panel data models. This specification says that the error for CO\textsubscript{2} emissions depends on the average of the errors in neighboring observations and its idiosyncratic component, implying that the unobserved errors \( u \) and \( \epsilon \) are entirely predictable from neighboring error \( W \epsilon \).

When the endogenous variable can be predicted as a function of spatially lagged of exogenous variables, the appropriate model to utilize is a Spatial Durbin model (SDM). The SDM is given by:

\[ Y_{it} = \rho \sum_{j=1}^{N} W_{ij} Y_{jt} + X_{it} \beta_{it} + \sum_{j=1}^{N} W_{ij} X_{jt} \gamma + \mu_{it} + \eta_{it} + \epsilon_{it} \]
\[ i = 1,2,...,N \]
\[ t = 1,2,...,T \]
\[ \epsilon \sim N(0, \delta^2 I_n) \] (4-6)

where parameter \( \gamma \) is a \((k \times 1)\) vector of spatial autocorrelation coefficients on the exogenous variables and \( \rho \) denotes a scalar spatial autocorrelation coefficient of endogenous variable. The SDM can be utilized to control if the model can be simplified to a SLM or a SEM since the
models nest dependence in both the disturbances and the dependent variable (LeSage and Pace, 2009). In order to specify non-spatial panel models against the spatial models (i.e. SAR and SEM), we utilize several Lagrange Multiplier (LM) tests. These tests examine whether the spatial models approach offer a proper specification. Moreover, we explore the joint significance of state fixed effects ($\mu_i$) and time period fixed effects ($\eta_t$) (Elhorst, 2012). The null hypothesis tests are:

$$H_0:\mu_1 = \mu_2 = \cdots = \mu_N = 0 \quad (4-7)$$

$$H_0: \eta_1 = \eta_2 = \cdots = \eta_T = 0$$

The LR tests are employed to investigate those null hypotheses. If the $p<0.05$, then we can reject the null hypothesis of joint insignificance (Elhorst, 2012).

If we fail to reject the spatial model, then the next step will be to explore whether the SDM model can be simplified to the SAR or SEM model. The hypothesis tests are:

$$H_0: \gamma = 0 \quad (4-8)$$

$$H_0: \gamma + \rho\beta = 0$$

where $H_0: \gamma = 0$ verifies whether the SDM can be simplified to the SAR, and $H_0: \gamma + \rho\beta = 0$ proves whether it can be simplified to the SEM (Elhorst, 2009; LeSage and Pace, 2009). All tests follow a $\chi^2$ distribution. If we reject both hypotheses, suggesting that the SDM is the best fit to the spatial panel data. Inversely, if we cannot reject the hypotheses, it suggest that the SAR and SEM provide the best fit for the panel data.

5. Data
This paper includes a panel dataset of 48 states and District of Columbia in the U.S. over the period 2000-2015. CO$_2$ emissions are estimated based upon a linear function of fossil fuels consumption where the volume of CO$_2$ emissions are calculated by the share of fossil fuels consumption (Department of Energy [DOE], 2016). The data of CO$_2$ emissions per fuel source, energy consumption, coal consumption, and energy prices are obtained from the U.S Energy Information Administration (2016) and the Carbon Dioxide Information Analysis Center (CDIAC). Urbanization rate data comes from United States Census Bureau (USCB) and United States Department of Housing and Urban Development (HUD). All data are obtained over the time period from 2000 to 2015.

The dependent variable is per capita CO$_2$ which is a measure of the total CO$_2$ emissions by state divided by the state population. Consistent with the insight of CO$_2$ emissions as a simple formulation of how the energy emissions are estimated through emissions-energy-output (EEO) model (Itkonen, 2012), the following equation is used:

$$CO_{2,t} = \alpha_{\text{coal}}E_{t}^{\text{coal}} + \alpha_{\text{oil}}E_{t}^{\text{oil}} + \alpha_{\text{gas}}E_{t}^{\text{gas}} + E_{t}^{\text{cem}} + S_t$$  

(4-9)

where $\alpha_x$ are the related thermal conversion factors and or coefficients of CO$_2$ emissions. Based upon equation 4-9, CO$_2$ emissions are calculated for the final energy use of four main sources of energy for each state: coal, crude oil, natural gas, and cement manufacturing. It is assumed that all carbon in the fuels are completely converted into CO$_2$ emissions. Due to lack of data for cement manufacturing, it is not calculated in this study. Finally, annual state population data are obtained from the U.S. Census Bureau (2016). Therefore, carbon dioxide emissions estimated based upon observation of state-level in the U.S (U.S. EIA, 2016).

Independent variables include urbanization rate, energy use, coal consumption, per capita GSP, energy prices, and RPS. Urbanization rate represents to the percentage of urban areas in
each state in the sample. Energy consumption represents per capita total energy use in each state per million Btu. The consumption of coal is determined by the amount of coal used by electric power plants per billion Btu. Per unit of energy, coal contributes more CO$_2$ emissions than any other fossil fuel, so coal consumption as part of energy use is considered as a separate explanatory variable. The GSP variable was converted to real values with a base year of 2009 by using the BEA's implicit price deflator.

It is hypothesized that per capita GSP represents to economic growth and affect CO$_2$ emissions (Markandya, 2006). An increase in economic growth or GDP require more energy sources and inputs, implying a higher use of natural resources. CO$_2$ emissions have measured by some methods and one of them related to this study is an EEO model. EEO models indicate the output-energy linkage and the output-pollution association (Itkonen, 2012). Economic growth, therefore, leads to a higher amount of waste and environmental degradation (Yandle et al., 2002; Jaforullah and King, 2017; Ahmad, et al., 2017).

Conversely, economic growth might have positive effect on the quality of the environment, generating more renewable energy and improving green technologies. Hence, it is can be argued that there is a negative correlation between GDP and emissions. Following the Environmental Kuznets Curve (EKC), initial increases in GDP would lead to environmental degradation but at some unspecified level, increases in GDP reduce environmental degradation and lead to reductions in emissions. Accordingly, the EKC indicates that there is an inverted U-shape linkage between economic growth and environmental degradation. At initial stage, economic growth increases CO$_2$ emissions while after the specific level of development, this relationship turns to inverse (Coondoo and Dinda, 2002).
Energy prices are state-level annual average prices of energy sources including coal, natural gas, and oil. The energy price index will add as another important variable that may influence CO₂ emissions. In the context of energy, there exists an inverse linkage between the price of fossil fuel and CO₂ emissions (Hammoudeh, et al., 2014). Higher prices of energy (i.e. coal and crude oil) lead higher reduction in CO₂ emissions over the long term and vice versa (DOE, 2017). It is predicted that the energy prices will be inversely correlated to CO₂ emissions since these emissions are measured based upon the use of energy sources. Energy prices were converted from nominal to real values by using the Bureau of Economic Analysis (BEA)'s implicit price deflator with a base year of 2009.

A Renewable Portfolio Standard (RPS) is another explanatory variable that reflects state level commitment to renewable energy. RPS is a state level policy instrument that aims to stimulate the supply of renewable energy in the U.S. electricity markets. Therefore, the RPS variable acts as an indicator variable for whether a state has already passed the law to encourage renewable energy. This variable takes a value of one starting in the year that a state has passed the RPS law and zero otherwise. This variable does not reflect the date when the RPS standard is implemented, as for many states, this standard comes into effect after 2015. A total of 29 states have passed and adopted RPS mandates as a method to develop diversification of their electricity markets, as well as promote reductions in air emissions. A RPS requires utility companies to provide a percentage of their electricity to be generated from renewable energy sources (Wiser et al., 2008). The variable characterizing the presence of an RPS was constructed using the

---

Database of State Incentives for Renewables and Efficiency (DSIRE). Descriptive statistics for all the variables are presented in Table 4-1.

6. Results and Discussion

This study examines the influential factors on state level CO$_2$ emissions. We have adopted spatial dynamic econometric models to avoid regression bias from omitting spatial correlation. Therefore, the principal concentration of this study is to examine how spatial relationships affect the factors of state-level CO$_2$ emissions. Accordingly, to investigate evidence of the non-spatial and spatial properties of the underlying variables, we examined both conventional and spatial models.

Prior to presentation of the empirical results for the spatial models, model misspecification tests are conducted. In order to explore the joint significance of the fixed effects, we conduct likelihood ratio (LR) tests. Table 4-2 provides the results of post diagnostics. Based upon the LR tests for non-spatial models for fixed effects, we can add the both the state fixed effects and time period fixed effects to the model. Moreover, regarding selection of panel data specifications between fixed effect and random effect, we need to conduct a Hausman test (Hausman and Taylor, 1981). The Hausman test has a null hypothesis of the preferred model is random effects versus the alternative of fixed effects (Green, 2008). It basically tests whether unique errors ($u_i$) are correlated with the regressors. The null hypothesis is they are not and if the significance level of this $\chi^2$ is < 0.05, we use fixed effects. The Hausman test results show that the calculated $\chi^2$ is 32.86 with corresponding p value 0.000, which denote that the estimated parameters are biased and inconsistent under the specification of random effects.

The estimation results for the non-spatial panel data models are shown in Table 4-3 and assuming there is no spatial impact ($\rho=\lambda=0$). Columns (1) through (4) represent the estimation

16 The data are publicly available at www.dsireusa.org.
results of pooled OLS, state fixed effects (SFE) only, time fixed effects only (TFE), and both state and time fixed effects, respectively. Examining Column (4) results, only coal consumption, per capita GSP, and RPS have statistically significant impacts on CO₂ emissions. Coal consumption has a positive impact while per capita GSP and RPS have negative impacts on CO₂ emissions.

Next, before we present estimation results from the three spatial models, we investigate which spatial model has the best fit for the panel data in this study. Therefore, the Wald and LR tests are applied to test the hypothesis whether the SDM model is the better spatial panel model compared to the SAR or SEM models. The Wald statistic's null hypothesis is that the exclusion restrictions (spatial autocorrelation parameters) hold (i.e., the model with spatial autocorrelation parameter(s) is arguably the proper specification) compared to the alternative hypothesis that the exclusion restrictions do not hold. LR tests the difference between the log-likelihood values of the unrestricted model versus the restricted model. If the difference between the two log-likelihood values is statistically significant, then the model with the spatial autocorrelation term(s) is arguably the proper specification.¹⁷ The results reported in Table 4-2 show that the SDM model is the best fit based upon the data. Hence, both the SAR and SEM models are rejected in favor of the SDM model. Estimated spatial models show that states with higher fossil fuel and energy consumption for cars, power plants, residential and industrial firms have more CO₂ emissions.

The estimated results for the three spatial models are reported in Table 4-4. The results show that at a 1% statistical significance level, CO₂ emissions is an decreasing function of

---

explanatory variables for urbanization rate, energy prices, RPS, and per capita GSP. The findings reported in Table 4-4 also reveal that at a 1% statistical significance level, CO$_2$ emissions are increasing functions of energy use and coal consumption. The $\rho_{adm}$ coefficient in the SDM model is statistically significant with contiguity based weight matrix, justifying the use of spatial econometric models by indicating the presence of a spatial autoregressive impact. This coefficient shows that an increase in CO$_2$ emissions in neighboring states leads to an increase approximately about 0.2 times emissions in the other states. The estimates of the $\rho$ parameter in the SDM specification is positive and statistically significant. This implies that if we ignore spatial dependence, we will be incorrectly interpreting the estimated effects.

The SDM model results indicates that increased urbanization, energy prices, and economic growth along with passage of a RPS will lead to decreased state-level CO$_2$ emissions; while increasing the coal consumption and energy use will lead to increased CO$_2$ emissions. Hence, not surprisingly, reducing fossil fuels in economy is an effective choice for decreasing CO$_2$ emissions. The size of the significant positive coefficient of the coal consumption indicates that encouraging the replacement of coal with renewable or natural gas energy plays a critical role in decreasing CO$_2$ emissions.

The direct and indirect impacts are important aspects of the spatial econometric models. The direct impact determines the effect of changing an explanatory variable on the dependent variable of a spatial entity. Moreover, the indirect impact is a function of how changes in independent variables in state $i$ affect state $j$. The indirect effect (neighboring) examines the impact of changing an explanatory variable in an entity on the dependent variable of all other entities (Elhorst, 2010).
Table 4-5 reports the estimation results for direct and indirect effects from the SDM model. Among the direct effects, each independent variable has a statistically significant impact at a 1% level, except for coal consumption which is significant at the 5% level. Moreover, the indirect effects of most explanatory variables also have significant impacts at the 1% level with exceptions of urbanization rate and the consumption of coal (both of which are significant at a 5% level) and GSP which does not have statistically significant impact. These results show that urbanization rate is significant in direct, indirect, and total effects on CO₂ emissions.

The negative direct effect and positive indirect effect of urbanization rate suggest that a growth in urbanization will reduce own state CO₂ while at the same time increasing CO₂ emissions of neighboring states. These results explicitly imply that a one percent increase in urbanization is linked with a 0.3 percent decrease of the per capita CO₂ emission in own state and an increase of 0.012 percent of per capita CO₂ emissions in neighboring states.

Both the direct and indirect effects of energy consumption along with consumption of coal are positive and statistically significant, implying that an increase in either leads to an increase of both own state and neighboring state CO₂ emissions. The statistically significant effect of per capita GSP is negative within own state emission, but the indirect effect is not statistically significant. This result suggests that if the own state per capita GSP increases, it will lead to reduce the CO₂ emissions in own state but not in neighboring states.

As shown in Table 4-5, the negative direct and indirect effects of energy prices suggests that if energy price increases, it will decrease CO₂ emissions in own state and neighboring states. Finally, both the direct effect and indirect effect of RPS are negative, which implies that if a RPS is implemented in one state, it will lead to reductions in own state CO₂ emissions as well as
emissions in neighboring states – indicative of interstate electricity transport and renewable energy certificates for electricity generation that can be traded across state lines.

Furthermore, the total effect estimates reported in Table 4-5 measure the sum of the direct and indirect effects. The total effects of all independent variables have statistically significant effects at a 1% level, except for energy consumption and coal consumption which are significant at the 5% level. The total impact of coal consumption and energy consumption are positive, suggesting that higher energy and coal consumption levels lead to higher CO₂ emission levels. In particular, coal consumption dramatically increases CO₂ emissions with a 1% increase leading to an own state increase of 1.31% in emissions and a total effect of almost double at 2.54%.

However, energy prices, GSP per capita, and RPS have the expected negative total effects which suggest that these three factors lead to decrease CO₂ emissions. Moreover, from these estimates we can argue that taking into account the negative direct effect of urbanization along with the positive indirect effect leads to a total negative effect. The intuition here is that urbanization advances reductions in CO₂ emissions, so growing cities across the U.S has a total effect of decreasing CO₂ emissions.

The impacts of factors and independent variables from Table 4-5 are essentially consistent with the previous literature and theoretical expectations offered in literature review and the theoretical model (Newman and Kenworthy, 1999; Cole and Neumayer, 2004; Andrews, 2008; Gonzalez, 2009; Jaforullah and King, 2017). More specifically, our results are in line with findings of Liddle and Lung (2010), Poumanyvong, and Kaneko (2010), Hossain (2011), Chuai et al. (2012), Burnett et al. (2013), and Xu and Boqiang (2015) since they show evidence favorable to the effects of urbanization, GDP and the energy factors (use, prices, and policies) on
CO₂ emissions. However, our study is in contrast to some of these studies since our findings show that urbanization has negative direct effect, while at the same time positive indirect effects on neighboring state level CO₂ emissions in the U.S.

7. Conclusions

This paper utilizes spatial econometric models to identify the impacts of urbanization rate, energy factors, GSP, and RPS on state level CO₂ emissions. Given the important role of urbanization and energy use in the U.S. economy, examining of how these influential factors affect CO₂ emissions is critical. As such, this study utilized a spatial panel data model approach to consider spatial linkages among U.S. state level CO₂ emissions, urbanization, energy consumption, energy prices, GSP, and RPS. The results from the SDM model show that all factors have statistically significant spillover effects on CO₂ emissions in the neighboring states with the exception of the variable measuring per capita income (GSP).

The empirical results suggest that urbanization rate, energy prices, RPS, and per-capita GSP are inversely related to CO₂ emissions in their own states, which implies that urban development, economic growth, and the implementing a renewable portfolio standard energy policy all reduce CO₂ emissions at the own state level. The SDM results also show that while more urbanization reduces own state CO₂ emissions, it increases CO₂ emissions of neighboring states by about 1/3 the amount of the own state reduction. One possible reason for this positive indirect effect of urbanization stems from interstate transportation of goods and services. As urban areas grow, there is increased importation of goods and products from other states, which create energy-related CO₂ emissions from production and transportation of these goods. Thus, fuels are used in one state to produce and transport goods consumed in another state so that CO₂ emissions are attributed to the state from which the products are transported. Also, the impact of
urbanization growth on CO₂ emissions has been affected by construction of well-designed cities and reduce CO₂ emissions by lower building energy-use per person and higher energy efficiency (Busch and Kennan, 2013). These results indicate the importance of developing federal and state policy to manage CO₂ emissions in the cities. Moreover, policies that decrease CO₂ emissions would be those that promote development of urban areas and implementation of renewable energy policies.

In addition, these findings include energy use and coal consumption have statistically significant positive effects on both own state and neighboring states CO₂ emissions. These results imply that higher coal consumption and energy use lead to higher CO₂ emissions on a regional basis not just in the state where these increases occur. Based upon total effects, the coal consumption has the highest impact on CO₂ emissions which suggests that a proper policy is to use more efficient energy sources, such as building natural gas power plants over coal-fired plants. Other factors are responsible for different CO₂ emissions paths followed by different states possibly due to urban infrastructure, solar and other renewable energy sources in cities, green technologies, and the development of new socio-economic systems in the U.S’ urban areas.

These results contribute to the energy and environmental economics literature by quantifying opposite direct and indirect effects of urbanization on state level CO₂ emissions along with the negative effects of income along with renewable portfolio standards (RPS) on state level emissions. The income effects provide additional support for the EKC on CO₂ emissions such that income growth reduces per capita CO₂ emissions at an own state level. For RPS, however, the indirect effects of CO₂ emission reductions are as large as the own state direct effects, indicating that cross border impacts are important. Both of these results are important
due to the prominence of state and local government actions in climate change mitigation now that the federal government has withdrawn from the Paris Climate Accords.

This study suffers from several limitations. First, the problem of measurement error of CO₂ emissions which is consistent with the rest of literature since CO₂ emission is based on the measurement of emissions from the burning of fossil fuels and not real ambient CO₂ emissions. Second, there are other factors that contribute to CO₂ emissions that include land use conversions from natural areas or farm to urban which are necessary to create urbanization. These factors are not accounted for in this analysis. Third, the spatial panel data process could suffer from problem of endogeneities within the independent variables and issue of theoretical framework (Partridge, et al., 2012).

Finally, this study only provides a primary interpretation of the effect of urbanization from a state level, hence, it excludes more explicit examining about the city size, the green technologies, a substitution from fossil fuel to renewable energy, and structural changes among various economic sectors in the U.S urban areas. Future work could be more directly investigate the roles of economic activities, sectors of economy in urban areas, renewable energy sources, and other excluded factors in explaining the effect of urbanization on environment and economic development. We look forward to future study along these lines.

8. References


LeSage, J., and Pace, K., 2009. Introduction to Spatial Econometrics. Taylor & Francis Group, LLC.


Tables

Table 4-1: Descriptive statistics

<table>
<thead>
<tr>
<th>Variables</th>
<th>Mean</th>
<th>Std.dev</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>CO₂ Emission Per Capita (C)</td>
<td>23.47</td>
<td>19.02</td>
<td>4.2</td>
<td>23.47</td>
</tr>
<tr>
<td>Urbanization growth (ur)</td>
<td>73.18</td>
<td>14.9</td>
<td>38.2</td>
<td>73.18</td>
</tr>
<tr>
<td>Energy consumption (ec)</td>
<td>31.1</td>
<td>30.25</td>
<td>244.4</td>
<td>31.1</td>
</tr>
<tr>
<td>Coal consumption (cc)</td>
<td>383270</td>
<td>366848</td>
<td>0</td>
<td>383270</td>
</tr>
<tr>
<td>GSP per capita (gsppc)</td>
<td>38121</td>
<td>8622</td>
<td>13288</td>
<td>38121</td>
</tr>
<tr>
<td>Energy prices(ep)</td>
<td>31.99</td>
<td>5.53</td>
<td>23.78</td>
<td>31.99</td>
</tr>
<tr>
<td>RPS</td>
<td>0.50</td>
<td>0.50</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Observations</td>
<td>784</td>
<td>784</td>
<td>784</td>
<td>784</td>
</tr>
</tbody>
</table>

Note: CO₂ emissions is per capita, urbanization rate is percentage of population in cities in each state, energy consumption is million Btu, coal use is the amount of coal for each state(Thousand shortens), per capita GSP is US dollar, and energy price is price index.

Table 4-2: Post diagnostic tests

<table>
<thead>
<tr>
<th>Tests</th>
<th>Non-spatial and Spatial</th>
<th>Joint significant for FE</th>
<th>Spatial Specification</th>
</tr>
</thead>
<tbody>
<tr>
<td>Likelihood ratio (LR):</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>State Fixed Effect (SFE)</td>
<td>128.35 (0.000)</td>
<td></td>
<td>-</td>
</tr>
<tr>
<td>Time Fixed Effect (TFE)</td>
<td>42.67 (0.000)</td>
<td></td>
<td>-</td>
</tr>
<tr>
<td>Spatial Lag</td>
<td>-</td>
<td>84.92 (0.000)</td>
<td></td>
</tr>
<tr>
<td>Spatial Error</td>
<td>-</td>
<td>72.54 (0.000)</td>
<td></td>
</tr>
<tr>
<td>Hausman</td>
<td>-</td>
<td>32.86 (0.000)</td>
<td>-</td>
</tr>
<tr>
<td>Wald :</td>
<td>Spatial Lag</td>
<td>-</td>
<td>70.45 (0.000)</td>
</tr>
<tr>
<td></td>
<td>Spatial Error</td>
<td>-</td>
<td>76.08 (0.000)</td>
</tr>
</tbody>
</table>

Note: All tests follow $\chi^2$ distribution with K degree of freedom. The numbers are the value of $\chi^2$ and all p-values are in parenthesis.
**Table 4-3: Estimation results of non-spatial, panel data models**

<table>
<thead>
<tr>
<th>Dependent variable: CO₂ emissions per capita</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>-5.89**</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>(Constant)</td>
<td>(-1.06)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Urbanization rate (ur)</td>
<td>-0.021**</td>
<td>0.023***</td>
<td>0.04</td>
<td>0.59</td>
</tr>
<tr>
<td>(0.013)</td>
<td>(0.004)</td>
<td>(0.58)</td>
<td>(0.61)</td>
<td></td>
</tr>
<tr>
<td>Energy consumption (ec)</td>
<td>1.039***</td>
<td>0.10</td>
<td>-0.003</td>
<td>0.023</td>
</tr>
<tr>
<td>(0.03)</td>
<td>(0.40)</td>
<td>(0.056)</td>
<td>(0.19)</td>
<td></td>
</tr>
<tr>
<td>Coal consumption (cc)</td>
<td>1.30***</td>
<td>1.04***</td>
<td>1.07***</td>
<td>1.01***</td>
</tr>
<tr>
<td>(0.03)</td>
<td>(0.02)</td>
<td>(0.07)</td>
<td>(0.03)</td>
<td></td>
</tr>
<tr>
<td>GSP per capita (gsppc)</td>
<td>-0.05**</td>
<td>-0.44***</td>
<td>-0.26***</td>
<td>-0.78***</td>
</tr>
<tr>
<td>(0.03)</td>
<td>(0.03)</td>
<td>(0.05)</td>
<td>(0.006)</td>
<td></td>
</tr>
<tr>
<td>Energy prices(ep)</td>
<td>-0.34**</td>
<td>-0.25***</td>
<td>-0.41**</td>
<td>-0.65</td>
</tr>
<tr>
<td>(0.09)</td>
<td>(0.07)</td>
<td>(0.10)</td>
<td>(2.94)</td>
<td></td>
</tr>
<tr>
<td>RPS</td>
<td>-0.026***</td>
<td>-0.017***</td>
<td>-0.15***</td>
<td>-0.31***</td>
</tr>
<tr>
<td>(0.009)</td>
<td>(0.007)</td>
<td>(0.004)</td>
<td>(0.001)</td>
<td></td>
</tr>
<tr>
<td>LM spatial lag</td>
<td>67.24***</td>
<td>20.17***</td>
<td>12.83***</td>
<td>8.94***</td>
</tr>
<tr>
<td>LM spatial error</td>
<td>27.34***</td>
<td>13.69***</td>
<td>7.01***</td>
<td>6.08***</td>
</tr>
<tr>
<td>σ²</td>
<td>0.12</td>
<td>0.03</td>
<td>0.06</td>
<td>0.04</td>
</tr>
<tr>
<td>R²</td>
<td>0.58</td>
<td>0.52</td>
<td>0.67</td>
<td>0.36</td>
</tr>
<tr>
<td>Observations</td>
<td>784</td>
<td>784</td>
<td>784</td>
<td>784</td>
</tr>
</tbody>
</table>

Note: The following models presented are: Column (1) pooled OLS (no fixed or time-period effects), Column (2), fixed effects only (individual fixed effects), Column (3) time-period effects only, and Column (4) both individual fixed effects and time-period effects, respectively.

All variables are measured as natural logs and *p<0.10, **p<0.05, ***p<0.01.

Standard errors are in parenthesis.
Table 4-4: Estimation results of spatial- panel data models

<table>
<thead>
<tr>
<th>Dependent variable: CO₂ emissions per capita</th>
<th>(SAR)</th>
<th>(SEM)</th>
<th>(SDM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Urbanization rate (ur)</td>
<td>-0.28*** (0.05)</td>
<td>-0.34*** (0.04)</td>
<td>-0.47*** (0.08)</td>
</tr>
<tr>
<td>Energy consumption (ec)</td>
<td>2.10*** (0.38)</td>
<td>0.55*** (0.07)</td>
<td>1.32*** (0.09)</td>
</tr>
<tr>
<td>Coal consumption (cc)</td>
<td>1.56*** (0.04)</td>
<td>1.34*** (0.05)</td>
<td>1.40*** (0.05)</td>
</tr>
<tr>
<td>GSP per capita (gsppc)</td>
<td>-0.22*** (0.07)</td>
<td>-0.83*** (0.25)</td>
<td>-0.41*** (0.06)</td>
</tr>
<tr>
<td>Energy prices (ep)</td>
<td>-0.12*** (0.01)</td>
<td>-0.35*** (0.11)</td>
<td>-0.62*** (0.03)</td>
</tr>
<tr>
<td>RPS</td>
<td>-0.18*** (0.007)</td>
<td>0.21*** (0.009)</td>
<td>-0.25*** (0.005)</td>
</tr>
<tr>
<td>ρ</td>
<td>0.22*** (0.02)</td>
<td>NA</td>
<td>0.20*** (0.015)</td>
</tr>
<tr>
<td>ξ</td>
<td>NA</td>
<td>0.56*** (0.04)</td>
<td>NA</td>
</tr>
<tr>
<td>w* Urbanization rate (ur)</td>
<td>NA</td>
<td>NA</td>
<td>0.043*** (0.004)</td>
</tr>
<tr>
<td>w* Energy consumption (ec)</td>
<td>NA</td>
<td>NA</td>
<td>0.55** (0.06)</td>
</tr>
<tr>
<td>w* Coal consumption (cc)</td>
<td>NA</td>
<td>NA</td>
<td>1.001*** (0.34)</td>
</tr>
<tr>
<td>w* GSP per capita (gsppc)</td>
<td>NA</td>
<td>NA</td>
<td>-0.64*** (0.02)</td>
</tr>
<tr>
<td>w* Energy prices (ep)</td>
<td>NA</td>
<td>NA</td>
<td>0.52 (0.93)</td>
</tr>
<tr>
<td>w* RPS</td>
<td>NA</td>
<td>NA</td>
<td>0.07*** (0.005)</td>
</tr>
<tr>
<td>σ²</td>
<td>0.07</td>
<td>0.06</td>
<td>0.04</td>
</tr>
<tr>
<td>R²</td>
<td>0.72</td>
<td>0.69</td>
<td>0.75</td>
</tr>
<tr>
<td>Observations</td>
<td><strong>784</strong></td>
<td><strong>784</strong></td>
<td><strong>784</strong></td>
</tr>
</tbody>
</table>

Note: All variables are estimated as natural logs and *p<0.10, **p<0.05, ***p<0.01. Standard errors are in parenthesis.
Table 4-5: Estimation results of direct, indirect and total effect of the SDM

<table>
<thead>
<tr>
<th>Dependent variable: CO₂</th>
<th>Direct effect</th>
<th>Indirect effect</th>
<th>Total effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Urbanization rate (ur)</td>
<td>-0.30***</td>
<td>0.012**</td>
<td>-0.288**</td>
</tr>
<tr>
<td></td>
<td>(0.03)</td>
<td>(0.005)</td>
<td>(0.09)</td>
</tr>
<tr>
<td>Energy consumption (ec)</td>
<td>0.53***</td>
<td>0.49***</td>
<td>1.02**</td>
</tr>
<tr>
<td></td>
<td>(0.05)</td>
<td>(0.07)</td>
<td>(0.15)</td>
</tr>
<tr>
<td>Coal consumption (cc)</td>
<td>1.31**</td>
<td>1.23**</td>
<td>2.54**</td>
</tr>
<tr>
<td></td>
<td>(0.60)</td>
<td>(0.37)</td>
<td>(1.20)</td>
</tr>
<tr>
<td>GSP per capita (gsppc)</td>
<td>-0.57***</td>
<td>-0.74</td>
<td>-1.31***</td>
</tr>
<tr>
<td></td>
<td>(0.10)</td>
<td>(1.12)</td>
<td>(0.11)</td>
</tr>
<tr>
<td>Energy prices(ep)</td>
<td>-0.28***</td>
<td>-0.46***</td>
<td>-0.74***</td>
</tr>
<tr>
<td></td>
<td>(0.047)</td>
<td>(0.020)</td>
<td>(0.08)</td>
</tr>
<tr>
<td>RPS</td>
<td>-0.014****</td>
<td>-0.015***</td>
<td>-0.029***</td>
</tr>
<tr>
<td></td>
<td>(0.006)</td>
<td>(0.0047)</td>
<td>(0.004)</td>
</tr>
</tbody>
</table>

Observations 784 784 784

Note: All variables are estimated as natural logs and *p<0.10, **p<0.05, ***p<0.01. Standard errors are in parenthesis.
Chapter 5: Conclusions

This dissertation has examined three aspects of energy economics: (1) the dynamic linkage between crude oil prices and stock market capitalization; (2) the relationship between coal consumption, the coal price, natural gas prices, weather, and GDP; and (3) spatial spillover among energy factors, urbanization and CO₂ emissions. Econometric methods utilized in this research included VAR, VECM, ARDL, and spatial econometric models. The three essays explore the influential factors of energy prices and energy factors in the ME economies, the U.S. coal market, and how the non-spatial times series and spatial panel data models perform in estimating linkages between variables. We have provided a nuanced view of three methods of econometric models that are important to the international energy markets and the U.S economy. This dissertation will be useful to anyone interested in global and national energy policies along with understanding how changes in energy factors have led to considerable changes in each of these markets. Similarly, anyone interested in competitive energy markets and the Middle East (ME) economies could benefit from this research.

The first essay (Chapter 2) investigates the extent of how changes in crude oil prices (OP) impact market capitalization (MC) in ME stock markets. This provides a more nuanced view of the crude oil prices and stock market capitalization. The results suggest that MC has reacted to the changes of OP generated by global economic crisis and shale revolution. Moreover, dependency on income from crude oil, OP increases have benefits for most nine oil-importing and oil-exporting countries in the ME.

The second essay (Chapter 3) explores the coal industry, paying particular attention to the relationships between coal consumption and coal plus natural gas prices. The results suggest that cross price elasticity of coal consumption with natural gas prices is 0.29 and own price elasticity
is -0.21 while previous elasticities were estimated at 0.17 and -0.11 in the U.S. (U.S. EIA, 2012).

In addition, a significant negative impact of GDP on coal demand, implying that coal is an
inferior energy source in the market and economic growth prompts decreases in coal consumption.

The spillover effects of energy factors on CO\textsubscript{2} emissions state level is the topic of the
third essay (Chapter 4). This essay examines the influential factors of energy-related CO\textsubscript{2}
emission levels within states in the U.S. We argue that there is spatial dependence among these
energy drivers and other main factors which cross state lines. Spatial dependence implies that
policies adopted within one state will impact emissions in neighboring states. Specifically, we
estimate a model of state level CO\textsubscript{2} emissions based upon per-capita GSP, energy prices,
urbanization, energy consumption, coal consumption, and renewable portfolio standards (RPS)
from 2000-2015. The findings suggest that energy use and coal consumption have statistically
significant positive effects on both own state and neighboring states CO\textsubscript{2} emissions. Also, CO\textsubscript{2}
emissions are negatively affected by energy prices, per-capita GSP, urbanization growth, and
RPS.

The main goal of this dissertation is to provide a basic reference for researchers and
policy makers to make proper decisions and set CO\textsubscript{2} emission reduction targets and policy. The
ME economies rely on oil income and their stock market values are connected to crude oil
prices. Thus, the ME economies need to diversify the economies to non- oil income sources. In
the coal industry, natural gas is substitute for coal and the U.S energy market should focus on
natural gas power plants since natural gas is more efficient and releases less emissions than coal.
Perhaps, exporting coal is an approach to improving the market for coal. Finally, urbanization
growth contributes less emissions for own state. Therefore, cities development should be
promoting for reducing emission level. Energy use and Coal consumption accounted for the highest total effects on CO₂ emissions in the U.S, and a combination of policies that address substitute renewable energy over fossil fuels and the transition to more fuel efficiency standards is necessary.

In this dissertation, we only considered one explanatory variable as the impact factor in the first essay because of the data limitation for political instability and macroeconomic activities in the ME. However, a lot of factors such as political and macroeconomic variables will indeed affect market capitalization in ME countries. Future studies for the first essay should be include more explanatory variables in the model. Moreover, the second and third essay need to include more exogenous variables including wind energy, city size, transportation, and population density. So future studies should consider the non-spatial and spatial panel data models by incorporating more explanatory variables in the models.