Four Essays on the Economics of Oil and Gas Markets

Inauguraldissertation
zur
Erlangung des Doktorgrades
der
Wirtschafts- und Sozialwissenschaftlichen Fakultät
der
Universität zu Köln

2015

vorgelegt
von

Diplom Volkswirt Benjamin Tischler

aus

Burglengenfeld

Referent: Jun.-Prof. Van Anh Vuong

Korreferent: Prof. Dr. Jörg Breitung

Tag der Promotion: 2.Oktober 2015

Acknowledgements

First and foremost, I would like to express my gratitude to my supervisor, Assistant Professor Van Anh Vuong, Ph.D. for supporting my dissertation project. I greatly appreciated her constructive criticism and personal encouragement which strongly contributed to the success of this thesis. I would also like to thank Prof. Dr. Jörg Breitung for his helpful methodological feedback and his encouragement to bring out the best in my research projects.

I am also grateful to Prof. Dr. Marc Oliver Bettzüge for his inspiring ideas as well as for assuming the role of the chair of the examination committee for this thesis. Furthermore, I would like to thank my co-author Sebastian Nick for the productive and cordial work on our research project on the integration of gas market. Moreover, I feel indebted to Hans Manner for his great collaboration in our joint research on robust estimation and testing for market integration.

I would like to extend my gratitude to Niyaz Valitov, Meltem Kaplan and Leonie Hass who provided excellent research assistance. In addition, I would like to thank (in alphabetical order) Dimitri Dimitropoulos, Christian Growitsch, Felix Höffler, Erin Mansur, Anne Neumann and Luis Orea as well as Heike Wetzel for valuable comments and discussions on some of the essays included in this thesis.

Furthermore, I would like to express my gratitude to the Chair of Energy Economics and the Institute of Energy Economics at the University of Cologne (EWI) for providing an inspiring research environment and offering financial support for the participation in academic conferences. I am particularly grateful to Chris Schäfer and Monika Deckers for her pervasive support during my time at Chair of Energy Economics and the EWI. The same applies for the whole administrative, communication, and IT-support team at EWI.

Finally, I would like to thank my family and friends for their support and encouragement. Above all, I would like to thank Manuela for her love, inspiration and perennial support.

Contents

List of Figures

Abbreviations

ADF	Augmented Dickey-Fuller test
ARCH	Autoregressive conditional heteroscedasticity
BAND-TAR	Band threshold autoregressive model
BAND-TVECM	Band threshold vector error correction model
BH	Baker Hughes
CCEMG	Common correlated effects mean group estimator
CCEP	Common correlated effects pooled estimator
EIA	US Energy Information Administration
EQ-TAR	Equilibrium threshold autoregressive model
FD	First difference estimators
FE	Fixed effects estimator
FEVD	Forecast error variance decomposition
FGLS	Feasible generalized least squares estimator
GDP	Gross domestic product
HH	Henry Hub
HQ	Hannan-Quinn Information Criterion
IRF	Impulse Response Function
JB	Jarque-Bera residual normality tests
LAD	Least absolute deviations
LL	Log-likelihood
LM	Lagrange Multiplier
LNG	Liquefied natural gas
LOP	Law of One Price
LR	Likelihood ratio
LTC	Long term contract
MA	Moving average
MMBtu	Million British thermal units
NBP	National Balancing Point
OLS	Ordinary least squares estimator
SAR	Sum of the absolute values of the residuals
SC	Schwartz Information Criterion
SER	Sequential elimination of regressors procedure
SSR	Sum of squared residuals
TAR	Threshold autoregressive model

t-ML	Maximum Likelihood estimator based on the t-distribution
TVECM	Threshold vector error correction model
UK	United Kingdom
US	United States of America
WTI	West Texas Intermediate

Chapter 1

Introduction

1.1 Background and Motivation

Since the 1960s hydrocarbons like crude oil and natural gas have been the most important energy source to fuel the world economy. Despite high recent growth rates for renewable energies, hydrocarbons are expected to be a dominant source of energy for several decades to come. In the last 20 years not only fundamental trends, but also singular events such as the financial crisis have shaped the development of crude oil and natural gas markets as well as the corresponding academic debate. On a historic scale these last two decades have seen changes in oil and gas markets that - in terms of price and regional production changes - are only matched by the turmoils of the oil crisis of the 1970s. Hence, the objective of this thesis is to improve the understanding of some of the dynamics and trends underlying these developments.

Around the turn of the millenium the growth of wealth and resource consumption in emerging economies and the fear of shrinking oil and gas production in industrialized countries refueled the time-honored debate about the point of time when the supply of fossil energy resources will reach its historic all time high. Labeled as the "peak oil hypothesis" the discussion gained most momentum for the case of crude oil, but similar worries were also expressed for natural gas. These solicitudes were also reflected and amplified by prolonged and steep increases in global oil and gas prices which in turn raises the question of how strong hydrocarbon supply will react to these price shocks. Accordingly, Chapter 2 of this thesis addresses oil and gas supply response to oil and gas price changes in the US. More specifically, because oil and gas supply can only be kept constant or increased by drilling additional wells, the essay focuses on the response of oil and gas drilling activity to changes in oil and gas prices as well as drilling costs.

In the late year 2008 the oil and gas price boom came to a sudden end in a stark collapse during the turbulences of the financial crisis and its macroeconomic aftermath. Both oil and gas prices stopped plummeting only in spring 2009. Whereas global oil prices quickly recovered to pre-crisis levels, the US price of natural gas stayed at low levels and even started to trend slightly downward. Given the growing gas prices in Europe this led several market stakeholders to announce a disintegration of the Atlantic gas market into an US market and an European market. However, given quickly expanding transport capacities and trade volumes for liquified natural gas (LNG) other industry participants hesitate to consent with the market decoupling hypothesis. Accordingly, Chapter 3 empirically investigates the integration of Atlantic gas markets before and after the financial crisis.

Price spikes are a frequent phenomenon in natural gas and other commodity markets. This can result in a low power of empirical market integration tests that are based on threshold autoregressive models. Similarly, threshold autoregressive (TAR) model based estimates of the transaction costs of arbitrage can be subject to large small sample biases when extreme price movements are present. To address this Chapter 4 investigates the small sample properties of TAR estimators and TAR based market (co-)integration tests. To remedy the poor small sample properties of the usual OLS based TAR estimation we propose two robust estimators. Accordingly, Chapter 4 methodologically advances the econometric approaches used in Chapter 3.

In North America, the major driver underlying this putative gas market decoupling was the so called shale revolution - a term that describes a steep and unforeseen increase in hydrocarbon supply from unconventional shale formations in the US. The shale revolution started off in gas markets where the novel combination of hydraulic fracturing and horizontal drilling in formerly unfamiliar shale geology enabled sweeping production gains. However, oil drilling also quickly adopted the new approach. By tapping shale oil resources with cost-efficient technologies and pulled by high oil prices US crude oil production is now close to surpassing its historic all time high of November 1970. Hence, for the moment the advent of the US shale revolution has refuted most arguments for an impending now and forever peak in US oil or gas production. As shale resources are available in many countries around the globe, the success of the US shale revolution stimulates the desire to replicate the shale revolution in other countries. A large part of the technical side of the replication debate is about updating the technology of rig fleets such that they are capable of performing hydraulic fracturing and horizontal drilling. However, in the US industry participants also highlight the importance of accumulated experience in using these technologies in a given region. Against this background, Chapter 5 investigates the role of drilling technology and local drilling experience as drivers of the oil and gas production growth during the US shale revolution. When the accumulation of local experience plays a crucial role for shale oil and gas production, it might be difficult to replicate the shale revolution in other countries by merely updating the technology level of the old rig fleet.

1.2 Outline of the Thesis

This thesis comprises four empirical essays on the economics of crude oil and natural gas markets. The following paragraphs briefly outlines the background, research questions, the methodological framework and the findings of the four essays.

Chapter 2 “What Drives Oil and Gas Drilling? A SVAR approach to supply response” investigates supply response in US crude oil and natural gas markets by disentangling the drivers of oil and gas well drilling activity. Further, the feedback from drilling activity on oil and gas prices is addressed.

Since the turn of the millennium, the dramatic changes in crude oil and natural gas prices on the one hand and US oil and gas production on the other hand stress the question of how oil and gas supply react to price changes. Oil and gas production crucially depend on how many new oil and gas wells are drilled in order to replace the declining production of older oil and gas wells. For this reason, market stakeholders have a great interest in understanding how well drilling activity reacts to changes in oil and gas prices and drilling cost. Another reason for this interest is that drilling activity statistics are published with very little delay to real time. Thus, drilling activity data can be regarded as the earliest available, reliable indicator for medium term changes in oil and gas supply. Prior studies such as Ringlund et al. (2008) either focused on oil drilling or gas drilling activity. However, oil drilling activity is interrelated with and gas drilling activity in at least two ways. First, the costs of oil and gas drilling are interdependent because the same drilling rigs are used for oil and gas wells. Moreover, oil wells often produce a significant amount of gas (and vice versa for natural gas wells). Thus, the revenues of both oil and gas wells are dependent on oil and gas prices.

To account for these interrelations between oil and gas drilling a structural vector autoregressive (SVAR) model is employed to disentangle the drivers of oil and gas drilling activity. In addition, the SVAR model is used to investigate whether changes in drilling activity feed back on oil and gas prices. The results imply that gas drilling is both driven by expected revenues - represented by gas and oil prices - and drilling cost. In contrast, oil drilling activity is mainly driven by variables related to the expected revenues, namely oil prices and macroeconomic activity. Whereas gas prices are strongly affected by changes in drilling activity, oil prices seem to react to drilling activity only weakly in the medium term.

Chapter 3 “The Law of one Price in Global Natural Gas Markets - A Threshold Cointegration Analysis” investigates market integration between the US and the European market for natural gas before and after the finacial crisis. In the latter period declining US gas prices led to a debate about whether or not the US and the European gas markets are still interlinked by liquefied natural gas

(LNG) arbitrage. The essay is a joint work with Sebastan Nick, who equally contributed to all parts of the study as co-author.

In the last decade, trade volumes of liquefied natural gas have expanded largely and are thought to indicate a trend towards a stronger integration of natural gas markets around the world. According to the Law of one Price (LOP) we expect prices of a homogenous good such as natural gas to converge under such circumstances. In contrast, important benchmark prices of natural gas around the globe seem to have decoupled since the financial crisis in the year 2009 and since the US shale revolution gathered momentum. Against this background, we empirically investigate whether the LOP holds even if prices seem to wander far from each other. In specific, we focus on the relationship between the prices at the Henry Hub in the US and the National Balancing Point in the UK. Previous research such as Neumann (2009) mostly used linear cointegration approaches to study the subject matter. However, linear approaches ignore that there should be no arbitrage activity, i.e., no error correction when the price difference is smaller than the transaction costs of arbitrage. First, we improve on the existing literature by using a threshold autoregressive (TAR) framework that is more adequately capturing the dynamics of spatial arbitrage in gas markets by explicitly taking account of transaction costs. Second, we use our model to test for market (co-)integration and to estimate a measure of transaction costs. Third, we investigate the adjustment of UK and US gas prices using a threshold vector error correction model. Our empirical results reveal a stable long-run relationship with threshold error-correction dynamics for the period 2000-2008 and for the period 2009-2012 when gas prices seem decoupled. However, in the latter period we find evidence for substantial impediments to arbitrage that by far exceed the usual US-UK transport cost differential. Further, for the period 2000-2008 we find that price convergence occurred by adjustment in both the US and UK market. In contrast, since the year 2009 price convergence was mainly achieved by downward pressure on the high UK prices.

Chapter 4 “Robust testing of the law of one price in natural gas markets” investigates the small sample performance of TAR based estimates of the transaction costs of arbitrage and related tests for market integration when price series are contaminated with price spikes. As a remedy for the weak small sample performance of OLS based methods two robust procedures for TAR based transaction cost estimation and testing for market integration are proposed. Hans Manner co-authored the study, and contributions to all aspects of the essay were made in equal parts.

Threshold auto-regressive models can be used to test for market integration in accordance with the law of one price as well as to estimate the transaction costs of arbitrage between two connected markets. However, resource and commodity price series such as natural gas frequently exhibit temporary price spikes. Interpreted as outliers or fat tails, such extreme price movements may lead to poor sample properties of TAR estimates and related cointegration tests. Particularly, price spikes may result in a large small sample bias in TAR threshold parameter estimates, representing transaction costs, and in a low power of to market (co-)integration tests. Hence, this paper considers the robust estimation of transaction costs using TAR models and robust testing of the LOP in the presence of occasional extreme price movements. We use Monte Carlo simulation and an application to US and UK natural gas prices to demonstrate that, in the presence of fat-tails or outliers OLS based estimates of the threshold, i.e., the transaction cost parameter can be severly biased. This, in turn, leads to a lower power of TAR based cointegration tests for market integration. In order to mitigate these problems, we propose two robust TAR estimation procedures based on the least-absolute-deviations (LAD) estimator and the maximum likelihood estimator assuming t-distributed errors (t-ML). The Monte Carlo simulations show that in the presence of fat tails or outliers the robust estimators have a lower small sample bias regarding the threshold parameter than the OLS estimator. Moreover, the our robust cointegration tests for market integration have a higher power than OLS based tests when fat tails or outliers are present.

Chapter 5 “The Role of Experience for the Shale Revolution” investigates the role of local experience and drilling technology for the US shale revolution.

“Hydraulic fracturing” and “horizontal drilling” are without doubt the core technologies behind the sweeping rise of US oil and gas production during the shale revolution. Additionally, industry representatives often stress the importance of experience for oil and gas extraction from shale formations. More specifically, local experience, i.e., the part of overall experience which cannot spill over to other regions or countries seems to play a decisive role in the shale industry. Furthermore, a strong role of local experience has implications for a replication of the shale industry in other countries. Whereas the technology level in other countries can be adjusted to US levels by upgrading the drilling rig fleet, local experience has to be accumulated by a large number of drilling operations. This in turn, would make it even more difficult for other countries to catch up on the US shale revolution.

Against this background, this essay empirically investigates the role of both drilling technology and local drilling experience for the production of shale oil and gas. To address this, we estimate oil and gas production functions that account for drilling technology and drilling experience. To take account of experience spillovers and to identify the effect of local drilling experience, we use the common correlated effects mean group estimator (CCEMG) and the pooled common correlated effects (CCEP) estimator introduced by Pesaran (2006). We find robust evidence for strong local experience effects in shale oil and gas production. Further, by disaggregating drilling and experience into different technology classes we can confirm that horizontal drilling indeed contributes most to the production of shale oil and gas. Moreover, the impact of experience gathered by using the horizontal drilling technology has the largest impact on shale oil and gas production compared with other production factors.

Chapter 2

What Drives Oil and Gas Drilling? A SVAR Approach to Supply Response

2.1 Introduction

Around the turn of the millenium the growth of wealth and resource consumption in emerging economies and the fear of shrinking oil and gas production in industrialized countries refueled the time-honored debate about the point of time when the supply of fossil energy resources will reach its historic all time high. Labeled as the "peak oil hypothesis", the discussion gained most momentum for the case of crude oil, but similar worries were also expressed for natural gas. The upsurge in the peak oil and gas debate coincided with shrinking US oil and gas production as well as with substantial increases in both global oil prices and US natural gas prices. However, in the years around 2005 gas production in the US started to grow again and US oil production followed this turnaround with a short delay. After a strong downturn during the financial crisis in the late year 2008, domestically determined US gas prices stayed on low levels. In contrast, globally determined oil prices quickly recovered almost to pre-crisis levels. On a historic scale, the movements of hydrocarbon prices during the first decade of the new millenium are only matched by the oil price movements of the 1970s oil crisis. These dramatic changes in crude oil and natural gas prices on the one hand and US oil and gas production on the other hand stress the question of how oil and gas supply react to price changes.

In the medium term, oil and gas production crucially depend on how many new oil and gas wells are drilled in order to replace the declining production of older oil and gas wells.[1] Therefore,

[1] "Medium term" refers to reactions in a variable that need between a few weeks and several months to take place after a shock in a driver of that variable has occurred.

market stakeholders who are interested in medium term supply response have a great interest in understanding how drilling activity reacts to changes in oil and gas prices and drilling cost. Another reason for this interest is that drilling activity statistics are published with very little delay to real time. Thus, drilling activity data can be regarded as the earliest available, reliable indicator for medium term changes in oil and gas supply. Accordingly, the main objective of this essay is to empirically investigate how the development of US oil and gas drilling activity is interrelated with the movement of oil and gas prices. We also address the role of drilling costs for oil and gas drilling activity. Further, because drilling activity is the most important early indicator of future supply, prices may react to changes in drilling activity. Hence, we also investigate whether changes in oil and gas drilling activity feed back on oil and gas prices. Finally, oil and gas prices as well as oil and gas supply may depend on overall economic activity in the US. Therefore, we also account for the state of the US economy in our analysis.

Up to now there are only few empirical studies on supply response in oil and gas extraction in general and drilling in particular. Studies on drilling responses start with an early work by Renshaw (1989) who employs a simple regression analysis to estimate the effect of an oil import fee on drilling activity in the USA. Iledare (1995) estimates the response of drilling measured as total footage drilled for natural gas to changes in wellhead prices and other variables. He uses pooled cross-sectional data for 1977 to 1987 for West Virginia. Ringlund et al. (2008) estimate price elasticities of oil drilling for different world regions and find that the US have a comparably high long-run price elasticity of oil drilling of 1.2 percent. Ringlund et al. (2008) use a bivariate autoregressive distributed lag model. Due to low data availability, they only include prices and drilling activity in their model and try to capture all other influences on drilling activity with deterministic regressors as well as with a stochastic time trend. Boyce and Nostbakken (2011) derive a theoretical model to explain exploratory and development drilling in the very long run. The model is tested using random and fixed effects estimators and a panel of US states with yearly observations from 1955 to 2002. A positive relationship between the number of drilled oil and gas wells and real oil and gas prices is found. The most recent paper in the field is Kellogg (2014) who uses monthly data from 1993 until 2003 and estimates the response of oil drilling in Texas to changes in the expected volatility of crude oil future prices.

Prior studies either focused on oil drilling or on gas drilling activity. However, oil drilling activity is interrelated with gas drilling activity in at least two important ways. First, the costs of oil and gas drilling are interdependent because the same drilling rigs are used for oil and gas wells. Moreover, oil wells often produce a significant amount of gas (and vice versa for natural gas wells). Thus, the revenues of both oil and gas wells are dependent on oil and gas prices.

To account for these interrelations, we use a structural vector autoregressive (SVAR) model to disentangle the drivers of oil and gas drilling activity. In addition, the SVAR model is used to investigate whether changes in drilling activity feed back on oil and gas prices. Further, the SVAR framework accounts for feedback between the variables that may lead to endogeneity problems in single equation approaches. Impulse response functions (IRF) and Forecast error variance decompositions (FEVD) are used to present and analyze the estimation results.

The results indicate that gas drilling is driven by variables that determine the revenues and by variables that determine the costs of gas drilling projects. In contrast oil drilling is mainly driven by variables that determine the revenues of oil drilling projects. Oil prices are only weakly affected by changes in oil drilling activity, whereas gas prices respond strongly and instantaneously to shocks in gas drilling activity. The remainder of this essay is organized as follows. In Section 2, the economics of the US oil and gas drilling are discussed. Section 3 presents the data. In section 4 the econometric approach is discussed, the SVAR models are specified and the identification approach is explained. In section 5 the SVAR estimates are presented and section 6 concludes.

2.2 Background

This section explains the technological and institutional setting in which oil and gas drilling activity take place and discusses the potential medium term drivers of drilling activity. Hydrocarbons such as crude oil and natural gas are extracted from fields i.e. underground geological formations that store hydrocarbons. Oil and gas production firms extract, process and sell oil and gas. Extraction is preceded by the exploration of a field by seismic surveys and scattered wildcat drilling to obtain some knowledge about the expected well output and the necessary drilling effort.The exploration phase is succeeded by the field development phase. In the development phase a great number of wells is drilled with the objective to profitably extract oil and gas for processing and sale. According to conversations with industry participants drilling processes last between a few weeks and many months - depending on geological characteristics and corresponding technological challenges.[2] Even given the geological field information available from the exploration phase, there is still some uncertainty left about the output and necessary drilling effort of development wells. Completed oil and gas wells exhibit high initial production rates which subsequently decline at decreasing rates due to the loss of geological pressure on the emptying reservoir.

Exploratory and development drilling in the US is sometimes conducted by the oil and gas production firms themselves, but more often realized by independent drilling companies.[3] Accordingly, drilling

[2] In the definition used here, the drilling process starts when the borehole is spud and ends when the well is completed.

[3] See Kellogg (2011) for details on the cooperation of drilling companies and production companies.

is often not an integral part of the activity of production companies, but there exists a leasing market for drilling rigs where suppliers are independent drilling companies and demanders are production companies. The costs of new wells, but also the total upstream costs of oil and gas production companies consist mainly of drilling cost.[4] Other upstream costs that occur after a well was drilled such as land lease, maintenance or pumping cost are very low in comparison. Whether or not a potential well with certain technological and geological characteristics is drilled is therefore mostly determined by expected drilling costs and expected revenues from the well. Hence, drilling activity should react to changes in expected revenues i.e. expected oil and gas prices as well as changes in expected drilling cost.

Expected revenues from drilling a well are determined by expected gas and oil prices over the total lifespan of the investment. Hence, the question arises how drilling investors form their expectations of future oil and gas prices. At least three possibilities seem sensible. First, future or forward prices could be used as predictors of expected spot prices in the future. However, according to Hamilton (2009) and Farzin (2001) future prices have theoretical as well as empirical drawbacks. Generally, future prices are not ideal predictors for spot prices in the future because they typically include costs of carry and risk premia. The latter could be positively correlated with the behaviour of current spot oil and gas prices as increased spot price volatility might lead to higher risk premia. Consequently, future prices may have weak predictive power especially in the volatile gas market. Second, following the theory of rational expectations (Muth 1961) current spot prices should already include all available information about future market development that is presently available. Hence, current spot prices could be used as a proxy for expected future prices. Third, the formation of price expectations could be adaptive i.e. market participants expect prices in the future to be a weighted average of current and past spot prices. Farzin (2001) argues that the adaptive version is also useful because it smoothes out short-run price volatility. The practical disadvantage of the adaptive expectations approach is that the number of time periods over which the average is calculated is arbitrary.

In this study the rational expectations approach is followed and current gas and oil spot prices are used to proxy expected future prices. Correspondingly, changes in drilling activity and resulting changes in expected oil and gas production might also have an instantaneous impact on oil and gas spot prices. In contrast to many other countries, in the US drilling statistics are published with only a few days delay to real time. Hence, drilling statistics are the earliest available, reliable indicator of future US production and supply of oil and gas. Accordingly, market participants are very vigilant about drilling statistics. Changes in drilling activity may, therefore, have an effect on buy and sell decisions in oil and gas trading which in turn influences the oil and gas price.

[4] The upstream part of the oil and gas sector is defined here as starting with the drilling of oil and gas wells and ends with the wholesale of extracted and processed oil and gas.

However, the response of US natural gas prices to changes in drilling activity can be expected to be stronger than the response of oil prices. The majority of US natural gas consumption is covered by domestic gas production. Hence, gas prices are mainly influenced by US domestic production and consumption. In contrast, the share of US domestic oil production in US oil consumption is smaller compared to natural gas. Moreover, the oil price is closely linked to global oil market developments and the condition of the overall US economy. Hence we expect the feedback from gas drilling to gas prices to be larger than the feedback from oil drilling to oil prices.

An important aspect of the revenue side of drilled wells is that they often produce both crude oil and natural gas in proportions that depend on field characteristics. Thus, registering a drilling project e.g. as a "gas drilling project" only means that the most of the revenues are expected to be generated by natural gas. But, in many gas fields oil or associated hydrocarbons such as ethane, propane, butane or natural gasoline can be found which are priced close to crude oil or natural gas prices depending on their specific characteristics. Hence, oil prices may also play a role for the revenues of declared "gas wells".[5] The same argument makes gas prices relevant for the revenues of oil wells. Therefore, in an empirical analysis of the determinants of oil or gas drilling activity both oil and gas prices have to be considered as explanatory variables.

Not only expected revenues, but also expected drilling cost are among the determinants of drilling activity. The costs of drilling projects can typically be fixed or at least quite precisely estimated at the point of time when the decision to realize a drilling project is made.[6] The costs of drilling a well are mainly determined by the geology, technology and drilling rig rental rates. Data on geological and technological determinants of drilling cost are hard to measure on an aggregate level and are not publicly available. The role of technology for aggregate drilling activity is ambiguous. On the one hand, technological progress can increase drilling activity when formerly inaccessible or unprofitable oil and gas deposits become accessible or profitable. On the other hand, technological progress can also decrease drilling activity as drilling projects can be finalized in shorter time. Both points are certainly true for unconventional drilling techniques such as hydraulic fracturing and horizontal drilling. In addition, Geology and the corresponding drilling technology change only in the long run. The objective of this essay, however, is to investigate the medium term determinants of drilling activity. Accordingly, we focus on the components of drilling costs that vary in the medium term.

[5] In some world regions that concentrate on crude oil extraction the share of gas in oil wells is negligible and vice versa for some "gas regions". However, in the US there is a substantial share of wells that deliver both gas and oil in economically significant amounts even if the wells are registered as "gas wells" or "oil wells". This phenomenon is particularly prevalent for the case of unconventional oil and gas extraction.

[6] Thus, uncertainties about drilling cost and expectation formation do not play a major role on the cost side of drilling projects.

Thus, we explicitly account only for rig scarcity and rig rental costs by using the utilization rate of the US rig fleet as a proxy for drilling cost.[7]

Exogenous factors that influence drilling activity include seasonal effects such as the spring thaw in northern US states that impedes the movement of heavy equipment and natural disasters such as the hurricane season in the year 2005 which lead to the temporary shut down of drilling activities. Other determinants of drilling activity are labor availability and tax issues. Due to data constraints these determinants have to left unaccounted for. However, it seems unlikely that the omission of these variables creates important inconsistencies in the estimates as they are arguably not strongly correlated with the other explanatory variables used in the models below.

Finally, the empirical analysis is facilitated by the assumption of competitive upstream oil and gas markets in the US. The assumption is underpinned by the high number and the not too heterogeneous size of market participants in both the rig market as wells as the wholesale markets for oil and gas.

2.3 Data

Monthly observations covering the period from October 1998 to September 2012 are examined. Data with monthly frequency are adequate to study the medium term response structure of oil and gas drilling activity. This sample period starts well after the gas and oil markets had adapted to reforms and deregulation that took place in the late 1980s. The sample is also recent enough to include observations since the rise of unconventional oil and gas extraction.

Oil and gas drilling activity data are from the US Energy Information Administration (EIA) and originally from the Baker Hughes‘s North American Rotary Rig Count which is a monthly census of the number of onshore rigs actively drilling for oil or natural gas in the United States.[8] The rig count is a very comprehensive measure of exploration and field development activity as it only excludes insignificant rigs such as very small truck mounted rigs, cable tool rigs and rigs that can operate without a permit. The real natural gas wellhead price for the US is used as the price of natural gas. Nominal wellhead prices were obtained from the EIA database and were deflated using the

[7] Drilling rigs are usually used for both oil and gas drilling. Thus, the approximation of the medium term variation of oil and gas drilling cost by rig utilization rates is reasonable. In the estimation below, we alternatively use the amount of oil drilling activity as an approximation of the costs of gas drilling (and vice versa). The quality of this approximation approach depends on further assumptions, e.g. about the size of drilling rig fleet and its reaction to rig scarcity. Hence, this second approximation approach will only serve as a robustness check, when the drivers of oil and gas drilling activity are investigated.

[8] “A rig is considered active from the moment the well is “spudded” until it reaches target depth [...]. Rigs that are in transit from one location to another, rigging up or being used in non-drilling activities [...] are not counted as active”. Drilling data are available at http://www.eia.gov/dnav/ng/ng˙enr˙drill˙s1˙m.htm.

monthly US implicit GDP deflator.[9] Wellhead prices are calculated by dividing the total reported value of all sold gas at the wellhead by the total quantity produced as reported by the appropriate US agencies. Therefore, wellhead prices include all costs prior to shipment from the lease including gathering and compression costs, severance and similar charges, but exclude pipeline transmission costs which can be significant at least when the pipeline system reaches its capacity limit. Wellhead prices can be considered as spot prices, but adjusted for transmission costs associated with current pipeline capacity limitations. Therefore, the wellhead price is more closely related to the revenues of gas extraction companies than, say, the Henry Hub spot price. Still, wellhead prices closely follow the spot prices at the most liquid trading places such as the Henry Hub. However, wellhead prices experience softer spikes than are typically observed at the Henry Hub which is another reason for preferring wellhead prices when it comes to econometric model estimation.

The real free on board prices of the benchmark crude oil West Texas Intermediate (WTI) as traded at the Cushing-Oklahoma spot market are used as oil prices. The nominal WTI oil prices were obtained from the US Energy Information Agency (EIA) database and were deflated using the monthly US implicit GDP deflator[10]. To proxy the costs of onshore rigs monthly rig utilization rates are obtained from the Guiberson-AESC Well Service Rig Count.[11] Albeit the Guiberson-AESC rig count actually measures onshore oil rig utilization rates, it closely resembles rig utilization rates because as outlined above almost all rigs can be used for both oil and gas drilling. Rig utilization rates are measured as the percentage share of rigs that are actively drilling compared to the total number of available rigs in the market. A rig is considered as active if, on average, it is crewed and worked every day during the month. The monthly, seasonally adjusted industrial production index of the US Federal Reserve Database represents the situation of the overall US economic situation. It is used as driver of oil and gas demand and therefore serves as an additional proxy for expected oil and gas well revenues.[12] The industrial production index measures the real production output of manufacturing, mining, and utilities and is standardized such that the average index values for the months of the year 2007 are equal to one hundred.

The Figures A.1, A.2, A.3 and A.4 in Appendix A show the plots of all time series used. Table A.1 and A.2 in the appendix provide detailed definitions, sources and descriptive statistics on the data and variables.

[9] The monthly values for the implicit GDP deflator were obtained by linear interpolation of quarterly deflator values to a monthly frequency. The wellhead prices are available at http://www.eia.gov/dnav/ng/ng˙pri˙sum˙dcu˙nus˙m.htm and the gdp deflator is available at http://research.stlouisfed.org/fred2/series/GDPDEF/

[10] WTI spot price are available at http://www.eia.gov/dnav/pet/pet˙pri˙spt˙s1˙m.htm

[11] Available at http://www.c-a-m.com/Forms/Product.aspx?prodID=cdc209c4-79a3-47e5-99c2-fdeda6d4aad6

[12] The US industrial production index is available at http://research.stlouisfed.org/fred2/series/INDPRO/

2.4 Structural VAR Models for Oil and Gas drilling

In this section the empirical approach to estimate the drivers of oil and gas drilling activity is presented and corresponding econometric models are specified. Our approach also addresses how oil and gas drilling activity feeds back on oil and gas prices. The empirical analysis is based on the estimation of a system of dynamic simultaneous equations in the form of a structural vector autoregressive model (SVAR). The SVAR is estimated in levels and identified by short run restrictions on the instantaneous relationships between the variables.

Table 2.1: Set of variables in the gas model, the oil model and the joint model

Model	Gas drilling	Oil drilling	Gas price	Oil price	Rig utilization	Macroeconomy
Gas model	x	-	x	x	x	x
Oil model	-	x	x	x	x	x
Joint model	x	x	x	x	-	x

The "x" represents endogenous variables that are included in the respective model given in the left column. A "-" means that a variable is not included in the model.

In order to provide robust results for the response of oil and gas drilling activity three models are specified. The first model focuses on the explanation of gas drilling activity (*gas model*), the second model explains oil drilling activity (*oil model*) and the third model aims at explaining both oil and gas drilling activity jointly (*joint model*). Moreover, the joint model is used to investigate how shocks in drilling activity affect oil and gas prices. Each model contains a different subset of the variables as shown in Table 2.1. Gas drilling and oil drilling is the number of onshore rigs actively drilling for oil and natural gas, respectively. Gas prices are real US natural gas wellhead prices and oil prices are real Western Texas Intermediate spot prices proxying expected future oil and gas prices. The rig utilization rate is from the Guiberson-AESC data set and approximates drilling cost. The macroeconomy variable is the industrial activity index of the Federal Reserve Economic Data base and represents the overall economic situation in the US.

The gas and oil model include gas and oil drilling activity, respectively, as the dependent variables of main interest. All three models proxy the expected revenues of gas and oil drilling, respectively, by the oil and gas price as well as by the macroeconomy variable. In the gas model and the oil model drilling costs are approximated by the rig utilitzation rate. In the joint model gas drilling activity is used as a proxy for the cost of oil drilling and vice versa for the cost of gas drilling. With this latter approach drilling costs can only be approximated roughly. Therefore, the joint model is only used as a robustness check for the results on the drivers of oil and gas drilling obtained from estimating the gas model and the oil model, respectively.

The joint model is primarily used to estimate how oil and gas drilling activity affect oil and gas prices. As oil and gas wells often produce both oil and gas, both oil and gas drilling are included in the joint model.[13] Therefore, it is important to not only include variables representing oil and gas supply, but also variables representing oil and gas demand. Accordingly, the joint model contains oil and gas drilling as a measure of (expected future) supply of oil and gas. The macroeconomy variable as well as deterministic terms capturing seasonalities (explained below) serve as a rough measure of (expected future) oil and gas demand.

As several demand and supply side variables are missing in the joint model, it clearly cannot fully explain oil and gas price formation. Other supply side variables include net storage withdrawals as well as pipeline and overseas imports. The demand side is mainly determined by demand from electricity generation, transport, heating and industrial activity. However, these omitted demand and supply side variables are arguably not strongly correlated with oil or gas drilling activity in the short and medium term once macroeconomic activity and seasonality are accounted for in the joint model. Thus, the corresponding bias in the parameter estimates regarding price responses to shocks in drilling activity is arguably small.

The variables were tested for unit roots and cointegration.[14] The unit root tests provide evidence that the variables are integrated of order one. Only the tests on the utilization rate are ambiguous. Depending on the test specification utilization rates are sometimes positively sometimes negatively tested for unit roots. The full set and relevant subsets of the variables were tested for cointegration. The tests were inconclusive about whether or not the variables are cointegrated. Gospodinov et al. (2013) discuss the advantages of alternative estimation approaches when there is uncertainty about the exact nature of (co-)integration. They conclude that unrestricted SVAR models that are estimated in levels and identified by short run restrictions are the most robust specification when tests are inconclusive about the magnitude of the largest roots and cointegration between the variables. Moreover, any estimated cointegration relationships established with the variables at hand would lack a clear theoretical interpretability. Therefore a level SVAR model with restrictions on the instantaneous relationships between the variables is used as the basic specification.

The reduced form of all three models is specified with the help of maximum lag order selection as well as subset selection procedures and evaluated with diagnostic tests. Thereafter, a structural model is identified and estimated using restrictions on the instantaneous relationship between the

[13] In contrast, the gas model and the oil model only contain gas drilling or oil drilling, respectively. Hence, these two models cannot take account of the fact that oil drilling activity may increase gas supply which in turn may affect gas prices. The same argument holds for gas drilling activity, its effect on oil supply and, eventually, oil prices.

[14] The tests proposed by Elliott et al. (1996) and Kwiatkowski et al. (1992) were used to test for unit roots. The test proposed by Johansen (1988) and Johansen and Juselius (1990) was used to determine the cointegration rank. Different lag orders and deterministic specifications were tested. The test results are left out here for brevity.

variables. All three models are based on a reduced form VAR model of order P that can be written as

$$y_t = \sum_{i=1}^{P} B_i y_{t-i} + \Phi D_t + e_t \tag{2.1}$$

where y_t is a N-dimensional vector of the endogenous variables presented in Table 2.1, the B_i are $N \times N$ coefficient matrices for the lags of y_t. Φ is a $N \times K$ matrix where K is the number of deterministic regressors. Further, e_t is a N-dimensional vector of reduced form errors which have a zero expected value $E(e_t) = 0$, have no autocorrelation $E(e_t e_s') = 0$ for all $s \neq t$ but are allowed to be contemporaneously correlated $E(e_t e_t') = \Sigma_e$, that is, $e_t \sim (0, \Sigma_e)$.

The capricious paths of drilling activity, utilization rates and natural gas prices (see Figures A.1, A.4 and A.2 in Appendix A) create problems with the normality of the residuals which leads to less precise estimates. Given our moderate sample size, influential outliers may distort the parameter estimates. To mitigate these problems, we include event dummy variables where there there are extreme residuals and where there is economic reason to do so. Hence, D_t is a vector of deterministic variables that includes impulse dummies for major events such as hurricanes and the downturn of the US economy during the financial crisis. Additionally, D_t contains a time trend and seasonal dummies. The time trend accounts for general technological progress. The seasonal dummies capture seasonal variations in oil and gas drilling as well as seasonal variations in oil and gas demand represented by oil prices, gas prices and the macroeconomy variable. To account for the trend shift after the economic crisis that is apparent in the drilling activity time series a trend shift variable is included that starts in May 2009.[15] This point of time roughly corresponds with the time when the US shale revolution starts to gain momentum. Hence, by including a trend shift variable we take account of technological change associated with the US shale revolution. The trend shift dummy is put where the strongest downturns of the financial crisis have ended and where visual inspection of the oil drilling data strongly suggests a trend shift. Table A.3 in Appendix A gives a detailed overview over the deterministic regressors in the models.

First, for each of the three models the maximum lag order of the VAR system is determined using the Schwartz (SC) and Hannan-Quinn (HQ) information criteria.[16] The SC and the HQ criteria point to a maximum lag length of two or three. A maximum lag length of three is choosen as specification tests detect no residual autocorrelation with this maximum lag order.[17] Second, the large sets of

[15] The trend shift variable is zero for all time periods before May 2009 and increases by one each month afterwards.
[16] Lütkepohl (2005) argues that the other alternatives for information criteria - namely the Akaike information criterion and the Final Prediction Error criterion - asymptotically overestimate the true VAR order, whereas the SC and HQ criteria offer consistent estimates of the true order.
[17] Residual autocorrelation can be detected when a maximum lag length of only two is used.

regressors in the VAR models reduce the degrees of freedom for the estimation and would lead to too imprecise estimates when the full set of regressors was used. Therefore, the sequential elimination of regressors (SER) algorithm proposed by Brüggemann and Lütkepohl (2000) is employed to further decrease the number of parameters in each of the three models by imposing subset restrictions on the reduced form VAR. The SER sequentially eliminates the lags of the explanatory variables that lead to the largest reduction of the selected information criterion until no further reduction is possible for the system of VAR equations. The remaining subset of regressors is deemed to include the most relevant regressors in statistical terms. In addition, Breitung et al. (2004) highlight that the asymptotics of SVAR estimates and the related impulse response functions are strongly improved when variables with a coefficient that is indeed zero are restricted to zero in the estimation. Third, the reduced forms of the three models are estimated with the SER subset restrictions in place using the feasible generalized least squares estimator (FGLS) outlined in Lütkepohl (2005).[18]

Diagnostic tests for residual autocorrelation and residual normality as well as tests for ARCH-effects are employed to evaluate the consistency and the efficiency of the estimates. Model checking results for all three models are shown in Table 2.2 (tests for residual autocorrelation) and in Table 2.3 (tests for residual normality and ARCH-effects).

Table 2.2: Tests for residual autocorrelation for the gas, the oil and the joint model

Test	$Q_{gas,16}$	$LM_{gas,5}$	$Q_{oil,16}$	$LM_{oil,5}$	$Q_{joint,16}$	$LM_{joint,5}$
Test statistic	392.5	136.1	376.5	144.3	362.2	132.3
Appr. distrib.	$\chi^2(374)$	$\chi^2(125)$	$\chi^2(376)$	$\chi^2(125)$	$\chi^2(376)$	$\chi^2(125)$
p-value	0.24	0.23	0.48	0.11	0.68	0.30

Q_{16} and LM_5 represent the Portmanteau and Breusch-Godfrey LM test for autocorrelation as described in Lütkepohl (2004) for lag orders of $h = 16$ and $h = 5$ for the gas, the oil and the joint model, respectively. In both cases the null hypothesis of no residual autocorrelation is tested against the alternative that at least in one residual series there is autocorrelation up to the specified lag order. The Brensch-Godfrey LM test has most power when testing for low orders of residual autocorrelation, whereas the Portmanteau test is preferable for large lag orders.

Most importantly, there is no evidence for autocorrelation in the residuals in all three models. There is some evidence for ARCH effects in some of the residuals in the oil and the joint model. Despite event dummy variables included there is still non-normality in some residuals of the oil and the joint model. Particularly, the residuals of the drilling and the utilization rate equations show ARCH effects and residual non-normality. The latter might be due to the fact that the sudden movements in drilling activity that produce extraordinarily high residuals cannot be fully captured by the event dummy variables. However, the inclusion of more event dummies is not considered as sensible as this would consume more degrees of freedom in the estimation and there is potentially too much arbitrariness involved when dummies are included for purely statistical reasons and without a strong

[18] The reduced form estimates are not shown as they are of no interest in themselves due to their lack of economic interpretability.

Table 2.3: Tests for residual normality and ARCH-effects

Model	JB Normality	Gas drilling	Oil drilling	Gas price	Oil price	Utilization rate	Macro-economy
Gas model	Test statistic	21.8	-	3.79	6.74	59.9	2.04
	p-value	0.00	-	0.15	0.03	0.00	0.36
Oil model	Test statistic	-	535.9	3.86	5.32	106.8	0.50
	p-value	-	0.00	0.14	0.06	0.00	0.77
Joint model	Test statistic	5.14	150.2	6.16	3.20	-	0.31
	p-value	0.07	0.00	0.04	0.20	-	0.85
Model	**ARCH-LM**	**Gas drilling**	**Oil drilling**	**Gas price**	**Oil price**	**Utilization rate**	**Macro-economy**
Gas model	Test statistic	14.6	-	15.8	15.1	13.8	11.8
	p-value	0.26	-	0.19	0.23	0.31	0.45
Oil model	Test statistic	-	75.7	16.2	14.4	19.3	6.19
	p-value	-	0.00	0.17	0.27	0.08	0.90
Joint model	Test statistic	12.4	78.3	16.4	14.6	-	13.8
	p-value	0.40	0.00	0.17	0.26	-	0.31

Results for univariate Jarque-Bera (JB) residual normality tests and univariate ARCH-LM tests as described in Lütkepohl (2004) are presented. The JB tests the null hypothesis of residual normality for one of the VAR equations. The ARCH-LM test has the null hypothesis of no ARCH effects in any of the residual series up to the specified lag length. Here, ARCH effects are tested up to a lag order of twelve. The rows of the panel show the test statistic and the p-value for the gas model, the oil model and the joint model for the JB test in the upper part of the panel and for the ARCH-LM test in the lower part. The columns two to seven show the results for the residuals of the specific VAR equations.

economic argument. To summarize, the tests show that the model is not fully satisfactory in every regard. But despite some deficits in efficiency due to non-normality and ARCH effects the estimates can be still regarded as appropriate for the purpose of this study. As linear dependencies are analyzed in the form of impulse response functions and variance decompositions, the reduced form estimates can nevertheless be considered as an adequate basis for the subsequent estimation of the structural from of the VAR model.

The estimation of a reduced form model is already a statistically valid description of the data generating process. But a reduced form model can only be a starting point as it generally has no economic interpretation.[19] To allow for an economic interpretation the corresponding structural form of the model has to be identified and estimated. In the structural representation instantaneous relations between some of the endogenous variables are allowed, but a sufficient amount of such relation has to be restricted based on economic intuition. That is, contemporaneous values of some endogenous variables are allowed as explanatory variables in some equations. The structural representation corresponding to equation 2.1 is given by

$$B_0 y_t = \sum_{i=1}^{P} B_i^* y_{t-i} + \Phi^* D_t + \epsilon_t \tag{2.2}$$

[19] The reduced form VAR model particularly lacks economic interpretability when economic intuition or theory suggest instantaneous relations between the endogenous variables and when there is correlation between the reduced form errors.

where y_t is defined as in the reduced form versions of the three models. B_0 is an invertible $N \times N$ coefficient matrix representing the instantaneous relations between the endogenous variables.[20] The B_i^* are $N \times N$ matrices of structural coefficients where $B_i^* = B_0 B_i$. Φ^* is a $N \times K$ matrix of structural coefficients for determinstic regressors where $\Phi^* = B_0 \Phi$ and ϵ_t is a N-dimensional vector of structural errors where $\epsilon_t = B_0 e_t$. A number of restrictions has to be imposed on the structural model to make the number of estimated structural parameters equal to the number of reduced form parameters. Ignoring deterministic terms in the reduced form model $(N^2)P + 1/2[N^2 - N]$ parameters are estimated. Again ignoring the determinstic terms and under the standard assumptions that structural errors are white noise and independently distributed and that the variances of the structural errors are normalized to one in the structural form $(N^2 - N) + N^2 P$ parameters have to be estimated.[21] Thus the difference in the number of parameters is $1/2(N^2 - N)$. This is just equal to the number of additional restrictions that need to be imposed to identify the structural model.

In our case these restrictions are imposed as zero restrictions on the matrix of instantaneous relations B_0. That means, using this identification scheme assumes that shocks may affect a subset of variables directly within the current time period, whereas another subset of variables reacts only after a time lag. An advantage of zero restrictions is that researchers can often find clear arguments for (not) imposing them based on economic intuition or theory. With arguably more responsive variables such as oil and gas prices on the one hand and presumably more laggard variables such as utilization rates, drilling and the macroeconomy variable on the other hand there should be enough pronounced differences in the (instantaneous) reaction patterns to make this identification scheme valid. Accordingly, with $N = 5$ endogenous variables for identification at least ten zero-restrictions have to be imposed on B_0.[22]

The identification schemes for the gas model, the oil model and the joint model are presented in the Tables 2.4, 2.5 and 2.6. Table 2.4 shows how the matrix of instantaneous impacts is restricted for the gas model which is mainly designed to explain the drivers of gas drilling activity. Drilling activity is allowed to instantaneously react to changes in all other variables. Even though large movements in drilling activity can only be expected in the medium term, i.e., with a lag of several weeks or months, drilling activity may still react very quickly to changes in its drivers. This can be due to idle drilling rigs that are already located in the large oil and gas regions in Texas and Louisiana and can be made operational in less than a month. This point is backed by Ringlund et al. (2008) who find an instantaneous reaction of drilling activity to price changes on a monthly basis. In contrast, prices

[20] In its normalized version the diagonal of B_0 contains only ones. The matrix is normalized such that each row represents an equation where the variable with coefficient one is the dependent variable.

[21] That is, the variance-covariance matrix of the structural errors is an identity matrix.

[22] For just identification exactly $1/2(N^2 - N) = 1/2(5^2 - 5) = 10$ restrictions have to be imposed. For over-identification more than ten restrictions have to be imposed. When more restrictions can be imposed than are necessary for just identification over-identification tests offer the possibility to test for the validity of the set of restrictions. This allows for potential statistical support for a certain choice of zero restrictions.

Table 2.4: Matrix B_0 with identifying zero restrictions for the gas model

Dependent variable	Gas drilling	Gas price	Oil price	Utilization rate	Macroeconomy
Gas drilling	*	*	*	*	*
Gas price	*	*	*	0	*
Oil price	0	0	*	0	*
Util. rate	*	0	0	*	0
Macroeconomy	0	0	0	0	*

Each row represents the coefficients of the instantaneous relationships between the endogenous variables of an equation in the system of VAR equations with the left column showing the dependent variable in that equation. An asterisk * means that the respective instantaneous relationship is estimated without restrictions. A zero means that a zero restriction is imposed on this instantaneous relationship.

are generally thought to have a fast reaction when new information about market fundamentals is available. Thus, gas prices are allowed to react instantaneously to gas drilling because drilling is perceived by market stakeholders as one of the most important and the most quickly available indicators of changes in future domestic US gas supply. Oil is a substitute for gas on the demand side in heating, the chemical industry and in electricity generation, hence, the gas price is allowed to instantaneously react to oil price changes. Similar to oil prices shocks to the US economy might also have a direct effect on gas prices because gas demand at least from electricity generation is strongly linked to overall output of an economy. Only the utilization rate is not allowed to have an instantaneous impact on the gas price. This is because almost any movement in the utilization rate that could have an instantaneous impact on the gas price is already captured by allowing for an instantaneous reaction of the gas price to shocks in gas drilling activity.

Oil prices are determined mainly by developments in the global market for crude oil of which the US economy is an important part. The impact of oil coming from gas wells and equally the impact of gas drilling activity on oil prices should be negligible in the very short run. The same is true for utilization rates proxying drilling cost which are most likely too remotely related to oil prices on a monthly basis. Hence, only the US macroeconomy is allow to influence oil prices instantaneously. According to Ramberg and Parsons (2012) oil and gas prices are cointegrated, but whereas gas prices react to deviations from the long run relationship oil prices do not adjust to deviations. That is oil prices can be considered to "pull" the gas prices in the very short run, but not the other way around. Due to this weak exogeneity property of oil prices gas prices are restricted to have no instantaneous impact on oil prices. Utilization rates naturally should react instantaneously only to drilling activity as the share of utilized rigs clearly changes with the number of rigs drilling for natural gas. However, the size of the rig fleet certainly cannot react to any influences in the short run.

Macroeconomic activity is assumed to feature no instantaneous reaction to any of the variables which is supported by the fact that the energy sector constitutes only a small part of the overall

US economy. Any change in gas or oil prices - let alone rig utilization rates or gas drilling activity - should trickle down only slowly into the overall economic situation. The number of restrictions is therefore eleven.[23] With eleven restrictions the model is overidentified by one restriction which offers the advantage that the validity of the identification scheme is testable. The LR Test as proposed in Breitung et al. (2004) for the validity of the overidentifying restriction uses a χ^2 distribution with one degree of freedom and yields a test statistic of 0.0846 and a p-value of 0.77. Thus, the overidentification test does not reject the overidentifying restriction which can be seen as some additional statistical support for the identification scheme used here.

Table 2.5: Matrix B_0 with identifying zero restrictions for the oil model

Dependent variable	Oil drilling	Gas price	Oil price	Utilization rate	Macroeconomy
Oil drilling	*	*	*	*	*
Gas price	0	*	*	*	*
Oil price	*	0	*	0	*
Util. rate	*	0	0	*	0
Macroeconomy	0	0	0	0	*

See notes for Table 2.4.

Table 2.5 shows the restrictions on the instantaneous relationships for the oil model. In the oil model drilling activity for crude oil is the variable of main interest. So, it is allowed to instantaneously react to all the other modeled variables. The gas price is allowed to instantaneously react to all variables, but oil drilling activity. This restriction is arguably valid as gas coming from oil wells is typically of minor importance for the overall US gas supply in the sample period. Further, as geological surveying is costly the amount of gas coming from oil fields and, thus, oil drilling is often not explored with equal precision as expected oil quantities. This uncertainty in how much additional oil drilling is contributing to gas supply and, eventually, gas prices might also decrease the instantaneous response of gas prices to oil drilling. Hence, it seems acceptable to restrict the instantaneous impact of changes in oil drilling activity on gas prices to zero. Oil price responses are restricted in the presented way for similar reasons as in the gas model. However, as the oil sector is the major focus of the oil model oil prices are allowed to react instantaneously to oil drilling activity. As in the gas model utilization rates again only react to drilling activity and macroeconomic activity is considered exogenous within a one month time horizon.

Table 2.6 shows the restrictions on the instantaneous relationships for the joint model. The joint model aims at incorporating the responses of both gas and oil drilling activity in one model and

[23] Alternative just- and overidentified models were also estimated. The results in the most reasonable specifications were qualitatively the same. However, there was a small number of reasonable identification schemes where the estimation procedure - the scoring algorithm explained below - did not converge. The overidentifying restrictions used here are considered as the most credible restrictions among the models in which the algorithm converged.

Table 2.6: Matrix B_0 with identifying zero restrictions for the joint model

Dependent variable	Gas drilling	Oil drilling	Gas price	Oil price	Macroeconomy
Gas drilling	*	*	*	0	*
Oil drilling	*	*	0	*	*
Gas price	*	0	*	*	*
Oil price	0	0	0	*	*
Macroeconomy	0	0	0	0	*

See notes for Table 2.4.

to provide a framework for investigating the feedback of oil and gas drilling activity on oil and gas prices. In the joint model oil and gas drilling are allowed to instantaneously react to all variables except the price of the respective other type of hydrocarbon. For instance, oil drilling is not allowed to immediately react to changes in gas prices. Gas prices are allowed to react to all variables except oil drilling. Oil prices are only instantaneously affected by macroeconomic activity. This slight coarsening of the set of restrictions on the contemporaneous effects is necessary to have enough restrictions for just identification. Particularly, the need to restrict the instantaneous reaction of oil prices to oil drilling activity is not desirable as impedes detecting instantaneous effect of oil drilling shocks on oil prices. However, lagged feedback can of course still be detected. On the other hand, this might not be too strong restriction for the overall picture as in the short run the oil price is probably more determined by other variables than domestic drilling.[24] Overall, the restrictions are credible and represent the most important economic relationships.

Finally, using the outlined identification schemes and the estimates of the reduced form covariance matrices the structural versions of the three VAR models are estimated with the scoring algorithm proposed by Amisano and Giannini (1997) and described in Breitung et al. (2004).[25] The parameter estimates of the structural VAR estimations are not presented here because the impulse response functions and forecast error variance decompositions employed in the next section offer a more handy way to show and interpret the results of the estimation.

2.5 Results

2.5.1 Impulse Response Functions and Forecast Error Variance Decompositions

To get an understanding of the system dynamics the estimation results are presented in two standard forms of innovation accounting: impulse response functions (IRF) and forecast error variance

[24] In the short run the oil price is arguably less dependent on supply and demand fundamentals than in the medium and longer term.

[25] In general the structural parameters are non-linear functions of the reduced form parameters. That renders conventional estimation such as standard maximum likelihood difficult as non-linear optimization methods are needed for computation. The scoring algorithm uses an iterative approach that albeit sometimes having convergence problems in practical applications mitigates the problems associated with non-linearity.

decompositions (FEVD). FEVD give insights into the size or economic relevance of the response of a variable to shocks in other variables. IRF are used to evaluate the sign, statistical significance and time structure of a response. First, IRF and FEVD are presented formally. IRF show how a one-off standard deviation structural shock in one variable propagates itself in the system. The IRF are based on the estimates of the structural VAR parameters of the three models: To obtain the IRF the structural VAR representation has to be inverted to get the structural vector moving average (MA) representation of the model also known as the Wold representation. Ignoring determinstic regressors as they have no direct impact on IRF analysis gives the MA representation

$$y_t = (B_0 - B_1^* L^1 - ... - B_P^* L^P)^{-1} \epsilon_t = \sum_{j=1}^{\infty} \Xi_j \epsilon_{t-j} \tag{2.3}$$

where L is the lag operator, the Ξ_j are $N \times N$ coefficient matrices that can be recursively obtained from the estimated coefficient matrices of the structural VAR model. The response of $y_{k,t+j}$ to a one standard deviation impulse in $y_{m,t}$ - the shock $\epsilon_{m,t}$ - is given by the (k,m)th elements of the matrices Ξ_j when viewed as a function of j. Thus, the elements of Ξ_j give the impulse responses to the structural shocks ϵ_t.[26] Confidence intervals for the IRF are estimated with the bootstrap version proposed by Hall (1992) and described in Breitung et al. (2004). The inference with bootstrapped confidence intervals is preferred here to Monte Carlo integration methods or delta method based inference. The reason is that the standard Monte Carlo approach relies on normal likelihood for the residuals which is not the case given the results from the normality tests for the reduced form model. The delta method heavily relies on asymptotics and has particularly poor small-sample properties in the presence of integrated variables and should therefore not be used. The bootstrap combined with the subset restrictions obtained before by the SER procedure is probably the best to deal with small sample problems and non-normal error distributions. For some discussion of IRF bootstrapping see Breitung et al. (2004). In all models, 5000 bootstrap replications were used to obtain reliable IRF confidence intervals.

Further, forecast error variance decomposition (FEVD) is used to gain a better understanding of the relative importance of the drivers of drilling activity and prices. In terms of the structural residuals the h-step forecast error is

$$y_{T+h} - y_{T+h|T} = \Xi_0 \epsilon_{T+h} + \Xi_1 \epsilon_{T+h-1} + ... + \Xi_{h-1} \epsilon_{T+1} \tag{2.4}$$

where the Ξ_i are obtained from the MA representation. The structural errors are assumed to have unit variance and are not contemporaneously correlated i.e. $\sigma_\epsilon = I_N$. The structural errors are also

[26] Recall that there was no conclusive evidence of (co-)integration of the variables in the models. Therefore, some shocks may show permanent effects in the IRF plots, that is, some impulse responses may not die out as j increases.

assumed to be uncorrelated over time. Accordingly, the forecast error variance of the nth element of y_{T+h} is

$$\sigma_k^2(h) = \sigma_{j=0}^{h-1}(\xi_{n1,j}^2 + ... + \xi_{nN,j}^2) = \sigma_{j=1}^{N}(\xi_{nj,0}^2 + ... + \xi_{nj,h-1}^2) \tag{2.5}$$

where $\xi_{km,j}$ denotes the (k, m)th component of Ξ_j. $(\xi_{nj,0}^2 + ... + \xi_{nj,h-1}^2)$ represents the contribution of the j shock to the h-step forecast error variance. The latter quantity is divided by $\sigma_k^2(h)$ to obtain the relative contributions of the shocks to a variable which are then reported for various forecast horizons.

In the following the IRF and FEVD of the estimated structural models are presented and discussed with respect to the two research questions. First, the IRF and FEVD are used to identify the statistical significance and relative strength of the drivers of gas and oil drilling activity. The analysis focuses on the gas and the oil model. The results of the joint model are reported as a robustness check. Second, the IRF and FEVD of gas and oil prices from the joint model are presented to investigate how oil and gas prices respond to changes in oil and gas drilling activity.

2.5.2 The Drivers of Oil and Gas Drilling Activity

Figure 2.1 shows the impulse responses of gas drilling activity to shocks in the other variables as estimated with the gas model. Table 2.7 shows the corresponding FEVD. The IRF show that, gas price shocks have a positive impact on gas drilling activity that becomes statistically different from zero only after about 5 months. The drilling response to a gas price shock reaches its maximum after one year and then slowly tapers off. Gas drilling activity shows a similar positive albeit somewhat shorter lived response to shocks in the oil price. These response patterns correspond to the reasonable duration of a planing and preparation phase between the investment decision to drill and the time when the well is spud. A reaction time of only a few weeks is possible where idle rigs are available or when rigs can be moved within the same area, the land is already leased by the production company and where drilling crews are readily available. Exploration projects probably take much longer from the price signal until the well is spud. Also, when new regions are developed the drilling response likely takes longer as a part of the rig fleet has to be moved and extraction infrastructure has to be built before large scale development drilling is feasible.

Shocks in the utilization rates that proxy drilling cost lead to an immediate negative response in gas drilling activity. A macroeconomy shock leads to no statistically significant response of gas

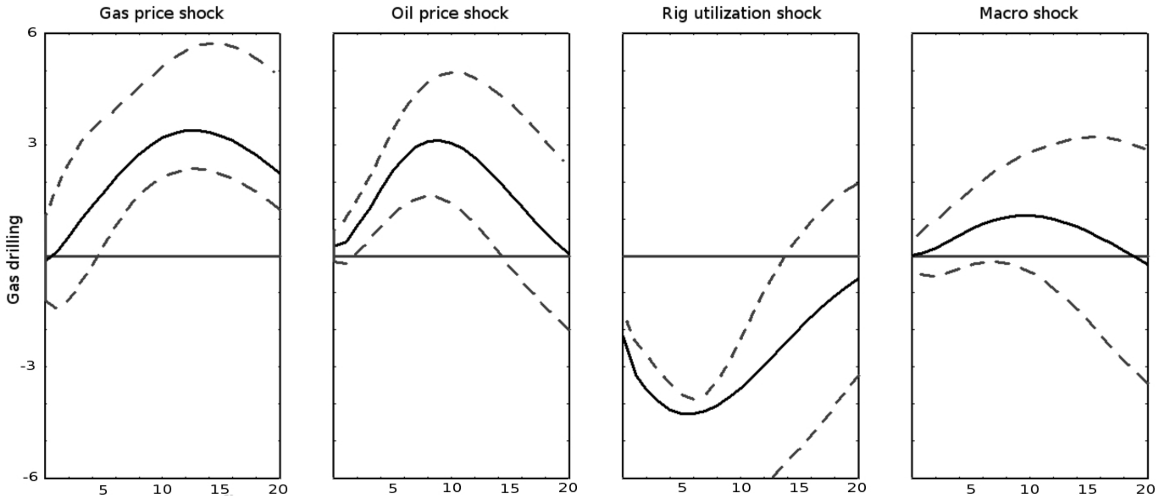


Figure 2.1: Responses of gas drilling activity in the gas model

drilling. This might be due to the fact that short term demand increases for natural gas triggered by temporary macroeconomic shocks might be mostly captured by storage withdrawals and gas imports. The FEVD show that gas prices and utilization rates play a crucial role for gas drilling activity as they jointly explain about 62 per cent of the error variance of gas drilling activity after 20 months. Oil prices have a somewhat smaller influence on gas drilling activity with up to 13 per cent after 12 months. This could be due to the fact that oil and natural gas liquids (that are priced similar to crude oil) that are extracted from gas wells account for some, but not a decisive part of the revenues of gas drilling projects. In correspondence with the IRF the macroeconomy has only a minor influence on gas drilling activity.

Table 2.7: Forecast error variance decomposition for gas drilling activity in the gas model

Forecast horizon	Gas drilling	Gas price	Oil price	Rig utilization	Macro economy
1	1.00	0.00	0.00	0.00	0.00
2	1.00	0.00	0.00	0.00	0.00
4	0.95	0.03	0.01	0.00	0.00
8	0.68	0.12	0.09	0.10	0.02
12	0.41	0.18	0.13	0.25	0.03
20	0.25	0.24	0.11	0.38	0.02
40	0.23	0.24	0.10	0.36	0.07

The columns show the estimated shares of the variance of the natural gas drilling activity accounted for by the respective structural errors of each variable at different time horizons.

As a robustness check Figure 2.2 and Table 2.8 show the IRF and FEVD for gas drilling activity in the joint model. As far as the models are comparable the results are qualitatively the same. However, in the joint model gas drilling shows an immediate reaction to a shock in the gas and oil price. The FEVD confirms that gas drilling is dominated by the influence of gas prices and to a smaller extent also by oil prices. The joint model also indicates that a shock to oil drilling activity

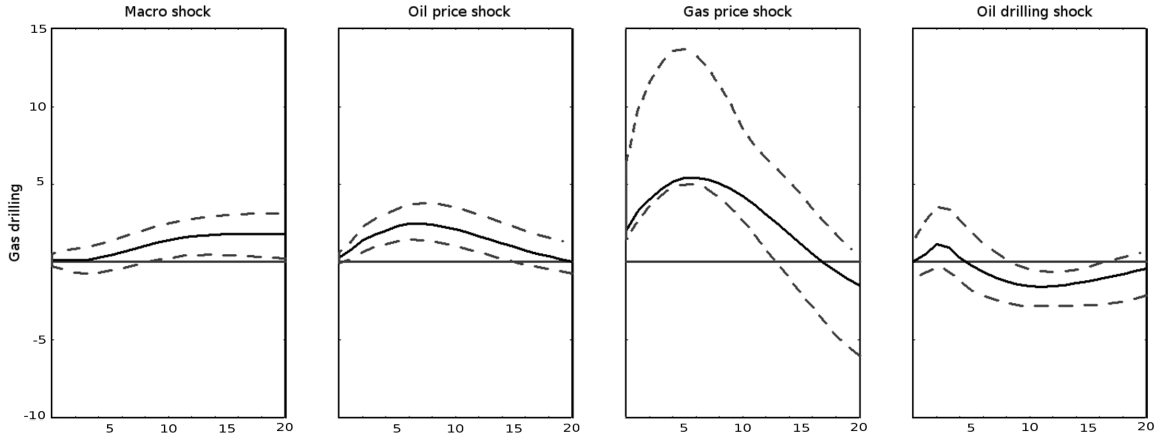


Figure 2.2: Responses of gas drilling activity in the joint model

increases the scarcity of drilling rigs and thereby significantly decreases gas drilling activity after several months. However, according to the FEVD oil drilling explains only four per cent of the error variance of gas drilling after 40 months. Thus, oil drilling is a statistically significant, but economically unimportant factor for gas drilling activity. Relatively low utilization rates of the US rig fleet throughout most of the sample period could explain this relatively weak negative tie between oil and gas drilling. Another explanation could be a rig fleet that reacts quite elastic to rig scarcity.

Table 2.8: Forecast error variance decomposition for gas drilling activity in the joint model

Forecast horizon	Gas drilling	Oil drilling	Gas price	Oil price	Macroeconomy
1	0.37	0.00	0.62	0.01	0.00
2	0.28	0.01	0.68	0.03	0.00
4	0.14	0.03	0.74	0.08	0.00
8	0.06	0.02	0.79	0.13	0.01
12	0.13	0.04	0.69	0.13	0.02
20	0.29	0.05	0.50	0.10	0.06
40	0.27	0.04	0.47	0.08	0.14

The columns show the estimated shares of the variance of the gas drilling activity accounted for by the respective structural errors of each variable at different time horizons.

To sum up, gas drilling is determined by both the revenue side and the cost side of drilling projects. On the revenue side gas prices play the major role, but to a lesser degree also oil prices are a factor to be considered. Once, gas and oil prices are controlled for the overall US economic situation has almost no impact on gas drilling.

Figure 2.3 shows the IRF of oil drilling activity to various variables as estimated with the oil model. Table 2.9 shows the corresponding FEVD. The IRF show that oil drilling responds positively to oil price shocks. The response in oil drilling becomes significant after three months, reaches its

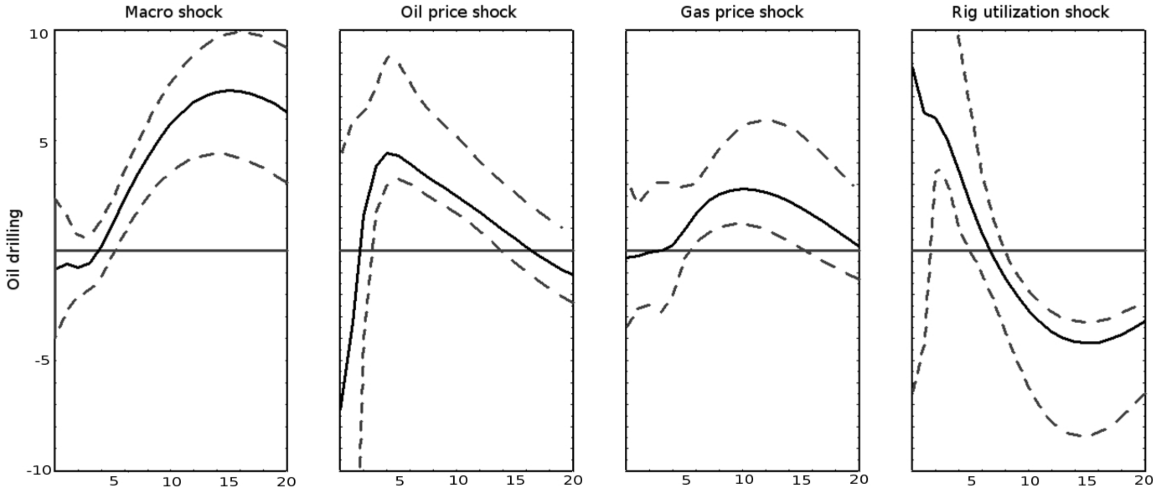


Figure 2.3: Responses of oil drilling activity in the oil model

maximum after five months and then tapers off to become insignificant after slightly more than a year. The response of oil drilling to gas price shocks is somewhat shorter lived. It becomes significant only after half a year and turns insignificant less than a year later. Oil drilling seems to react somewhat awkward to a utilization rate shock as it first reacts positively and only after several months reacts significantly in the expected negative way. This may be due to the fact that utilization rates are not a perfect proxy for drilling costs. Particularly in the very short run an increase in the utilization rate can be caused by an increase in oil drilling activity. This might be captured only imperfectly by the equation explaining the utilization rate and, in turn, might cause the IRF to show a positive response of oil drilling in the short run. A shock in the macroeconomy leads to a strong positive reaction of oil drilling activity, that becomes significant after about five months and reaches its maximum after 15 months.

Table 2.9: Forecast error variance decomposition for oil drilling activity in the oil model

Forecast horizon	Oil drilling	Gas price	Oil price	Rig utilization	Macroeconomy
1	0.74	0.00	0.11	0.15	0.00
2	0.76	0.00	0.09	0.15	0.00
4	0.69	0.00	0.10	0.21	0.00
8	0.61	0.01	0.16	0.20	0.02
12	0.50	0.03	0.15	0.19	0.12
20	0.34	0.04	0.11	0.20	0.31
40	0.28	0.05	0.11	0.19	0.36

The columns show the estimated shares of the variance of the oil drilling activity accounted for by the respective structural errors of each variable at different time horizons.

The FEVD reveal that oil drilling activity is mainly driven by macroeconomic activity. Notably, oil prices play a less important role for oil drilling activity than gas prices for gas drilling activity. This can be explained by the results below concerning the drivers of the gas and the oil price (see Table 2.12 and Table 2.11). The gas price seems to be much less influenced by macroeconomic activity than the oil price. The impact of the gas price on drilling is relatively robust to the inclusion of macroeconomic activity in the model whereas the influence of the oil price vanishes once the macroeconomy - presumably the main driver of the oil price - is included. With a share of up to 21 per cent of the error variance of oil drilling after four months the rig utilization rate also plays an important role for oil drilling. But, compared to its impact on gas drilling the rig utilization rates play a less important role for oil drilling activity. This can be explained by the relatively high profit margins given the high oil prices in the sample period. The gas price has only negligible influence on oil drilling activity.

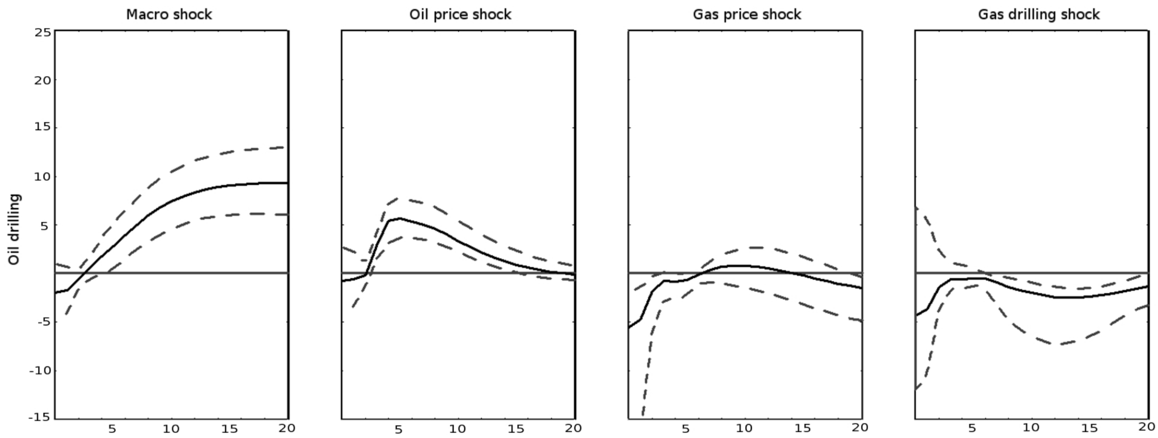


Figure 2.4: Responses of oil drilling activity in the joint model

Again, as a robustness check the IRF (Figure 2.4) and FEVD (Table 2.10) for oil drilling activity in the joint model are presented. The results for the responses to oil price and macroeconomy shocks are roughly the same. Gas price shocks seem to have an immediate negative impact on oil drilling that, however, stays significant only for two months. Shocks in gas drilling activity lead to a significant and negative response in oil drilling activity after about six months. This parallels the finding above that a shock in oil drilling leads to a decrease in gas drilling (see Figure 2.2) and supports the presumption for a relatively inelastic supply side in the drilling rig market. The FEVD confirm that oil drilling is predominantly driven by macroeconomic activity. The influence of all other variables is comparatively small.

In summary, oil drilling and correspondingly the expected profits of such projects are mainly influenced by expectations about the macroeconomy. Even if the scarcity of drilling rigs has a statistically significant effect on oil drilling, this cost related variable does not seem to be a too important determinant of expected profitability of drilling projects in the oil sector.

Table 2.10: Forecast error variance decomposition for oil drilling activity in the joint model

Forecast horizon	Gas drilling	Oil drilling	Gas price	Oil price	Macroeconomy
1	0.04	0.89	0.06	0.00	0.01
2	0.04	0.89	0.06	0.00	0.01
4	0.04	0.88	0.06	0.01	0.01
8	0.03	0.74	0.05	0.12	0.06
12	0.04	0.60	0.04	0.13	0.19
20	0.04	0.40	0.03	0.09	0.44
40	0.03	0.21	0.02	0.05	0.69

The columns show the estimated shares of the variance of the oil drilling activity accounted for by the respective structural errors of each variable at different time horizons.

2.5.3 How Oil and Gas Prices Respond to Shocks in Oil and Gas Drilling

To address the second research question of whether prices respond to drilling activity shocks the IRF and FEVD of gas and oil prices in the joint model are examined. Figure 2.5 shows the impulse

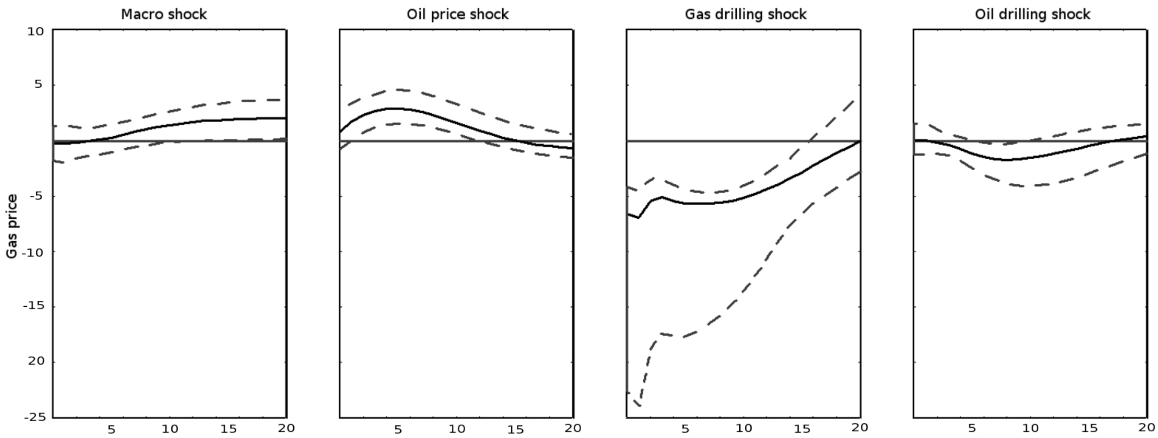


Figure 2.5: Responses of gas price in the joint model

responses for gas prices in the joint model. In accordance with economic intuition gas prices react positively to oil price and macroeconomic shocks. The response to the oil price shock becomes significant after one month and turns insignificant again after about eleven months. This reflects that oil and gas are imperfect substitutes for hydrocarbon demanders. The response to a macroeconomic shock is merely significant. A positive shock to the overall US economy seems to only effect gas prices in the long run. Gas prices instantaneously react significantly and negatively to a gas drilling shock. This strong, immediate reaction of gas prices underpins the claim that drilling statistics are a particularly important signal for US gas market stakeholders about future gas supply. In contrast, oil drilling shocks lead to a negative, but only merely significant response of gas prices. Hence, the gas from oil wells has a comparatively small effect on gas supply and gas prices, accordingly. The variance decomposition for gas prices in the joint model are shown in Table 2.11. The FEVD show that the error variance of gas prices is explained to a large part by drilling activity both in the short

Table 2.11: Forecast error variance decomposition for natural gas prices in the joint model

Forecast horizon	Gas drilling	Oil drilling	Gas price	Oil price	Macroeconomy
1	0.48	0.00	0.51	0.01	0.00
2	0.48	0.00	0.50	0.02	0.00
4	0.47	0.00	0.48	0.05	0.00
8	0.52	0.01	0.37	0.09	0.00
12	0.57	0.03	0.30	0.09	0.01
20	0.53	0.02	0.32	0.08	0.04
40	0.46	0.02	0.33	0.06	0.13
70	0.40	0.02	0.29	0.05	0.24

The columns show the estimated shares of the variance of the natural gas prices accounted for by the respective structural errors of each variable at different time horizons.

and the long run. Interestingly, the oil price only has a comparatively small influence on the gas price. Macroeconomic activity plays an increasingly important role in the long run explaining 24 per cent of the error variance after 70 months. Oil drilling activity plays only a negligible role for the trajectory of gas prices over time.

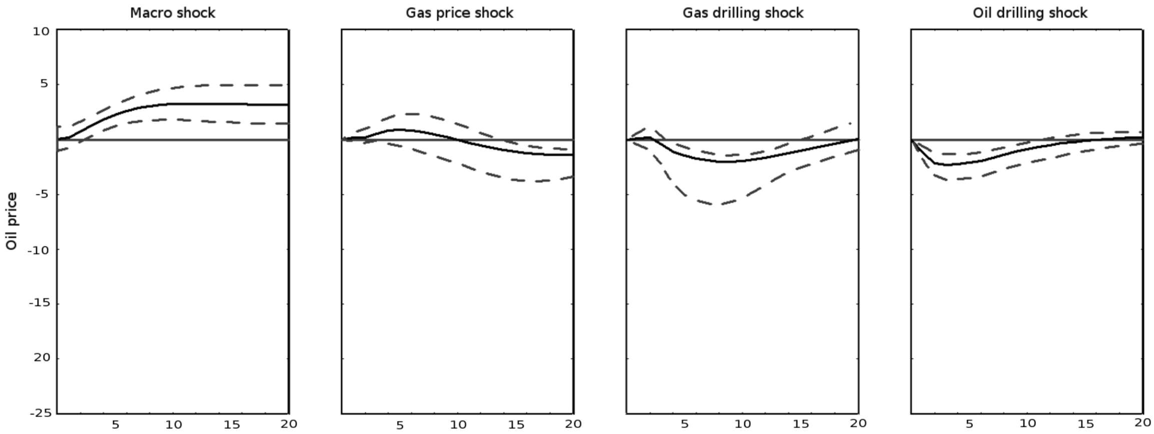


Figure 2.6: Responses of the oil price in the joint model

Figure 2.6 shows the impulse responses for oil prices in the joint model. As expected oil prices are significantly positively related to macroeconomic shocks. This effect becomes significant after two months and is persistent possibly implying that macroeconomic activity is a likely candidate for the stochastic trend driving the system of variables. The response to gas price shock stays insignificant for many time periods and turns merely negatively, significant after more than a year. Oil prices react negatively significant both to oil and gas drilling shocks. The response to an oil drilling shock takes place after one month whereas the response to a gas drilling shock takes about three months to become statistically significant. The responses turn insignificant again after about one year after the respective response became significant. The variance decomposition for oil prices in the joint model is shown in Table 2.12. The FEVD reveal that oil prices are predominantly driven by macroeconomic activity in the long run explaining 75 per cent of the error variance of oil prices after 70 months.

In the long term all other variables play a negligible role. In the medium term oil and gas drilling gain some importance explaining nine and twelve per cent of the forecast error variance after one year. This mirrors the fact that in contrast to domestic gas production domestic US oil production is relatively small compared to overall US oil consumption. Consequently, oil drilling has a smaller effect on the US oil price than gas drilling has on the gas price. Further, gas prices play almost no role for oil prices neither in the long run nor in the short run.

Table 2.12: Forecast error variance decomposition for oil prices in the joint model

Forecast horizon	Gas drilling	Oil drilling	Gas price	Oil price	Macroeconomy
1	0.00	0.00	0.00	1.00	0.00
2	0.00	0.02	0.00	0.98	0.00
4	0.00	0.09	0.00	0.89	0.02
8	0.05	0.13	0.01	0.69	0.12
12	0.09	0.12	0.01	0.54	0.24
20	0.09	0.09	0.04	0.40	0.39
40	0.07	0.05	0.04	0.25	0.59
70	0.04	0.03	0.03	0.16	0.75

The columns show the estimated shares of the variance of the oil prices accounted for by the respective structural errors of each variable at different time horizons.

In summary, there is evidence for a strong, instantaneous response of gas prices to shocks in gas drilling activity. The response of gas prices to changes in oil drilling are statistically significant but much weaker in quantitative terms. In the longer run oil prices are almost exclusively driven by the overall economic situation of the US, but in the medium term oil prices also depend on oil and gas drilling shocks. In line with Ramberg and Parsons (2012), we find that oil prices play a more important role for the evolution of gas prices than vice versa.

2.6 Conclusions

Gas and oil drilling statistics are arguably the most important predictors of future gas and oil production. The drivers of US gas and oil drilling activity were examined using a structural VAR model. This framework has the advantage that it explicitly accounts for the interrelation between oil and gas drilling activity. First, the SVAR takes account of the interrelations between the costs of oil and gas drilling. Second, it accounts for the fact that oil and gas wells often produce both oil and gas. Therefore, the revenues of oil and gas drilling projects depend on both oil and gas prices.

Gas drilling and the corresponding expected profitability of such investments is both strongly determined by variables determining the revenue side - namely, gas prices and oil prices- and by drilling costs. In contrast, oil drilling activity is predominantly driven by variables related to the expected revenues of oil wells, namely overall US economic activity and the oil price. Drilling costs turn out to be less important for oil drilling investments than for gas drilling projects. These findings shed light on the relative importance of the drivers of drilling activity and give us insights into the time structure of the drilling responses to changes in its main drivers.

Further, the response of oil and gas prices to shocks in oil and gas drilling variables was examined. Gas prices show a strong, immediate and negative response to shocks in gas drilling activity. Oil prices show a significant negative reaction to increased oil and gas drilling activity in the medium term. However, in the longer run oil prices are clearly dominated by the overall economic activity in the US. These results confirm that market participants at least in the gas sector are very vigilant about drilling statistics as one of the earliest, reliable sources of information on future oil and gas supply.

From a methodological perspective, some evidence for simultaneous relations between gas drilling activity and gas prices was found. Further, significant interrelations between oil and gas drilling as well as oil and gas prices were found. This highlights the need to take account of endogeneities when estimating oil and gas supply responses and casts doubt on reduced form single equation approaches found parts of the literature. Further research could include a more comprehensive and detailed modeling of the drilling rig market. Another issue could be to capture technological change explicitly by using additional variables or by using smooth structural transition functions.

Chapter 3

The Law of One Price in Global Natural Gas Markets - A Threshold Cointegration Analysis

3.1 Introduction

In the last decade, trade volumes of liquefied natural gas (LNG) have increased strongly and are thought to promote the integration of global markets for natural gas.[1] However, important benchmark prices of natural gas around the globe have diverged since 2009. This observation is frequently referred to as the "decoupling" of gas markets. Since then, the US Henry Hub (HH) spot price started its descent while prices at the UK National Balancing Point (NBP) increased, resulting in high price spreads between the US and UK natural gas markets (see Figure 3.1). The law of one price (LOP) theory states that prices will adjust towards each other due to arbitrage activity once the price spread between two markets is larger than the transaction costs of spatial arbitrage. Thus, the question arises whether the seemingly decoupled benchmark natural gas prices such as the HH spot price and the NBP spot price are still pulled together by forces of arbitrage even if persistent price spreads can be observed between these markets.

[1] Accordingly, since the year 2000 liquefaction capacities for LNG exports and regasification capacities for LNG imports have more than tripled, LNG shipping capacities have more than doubled.

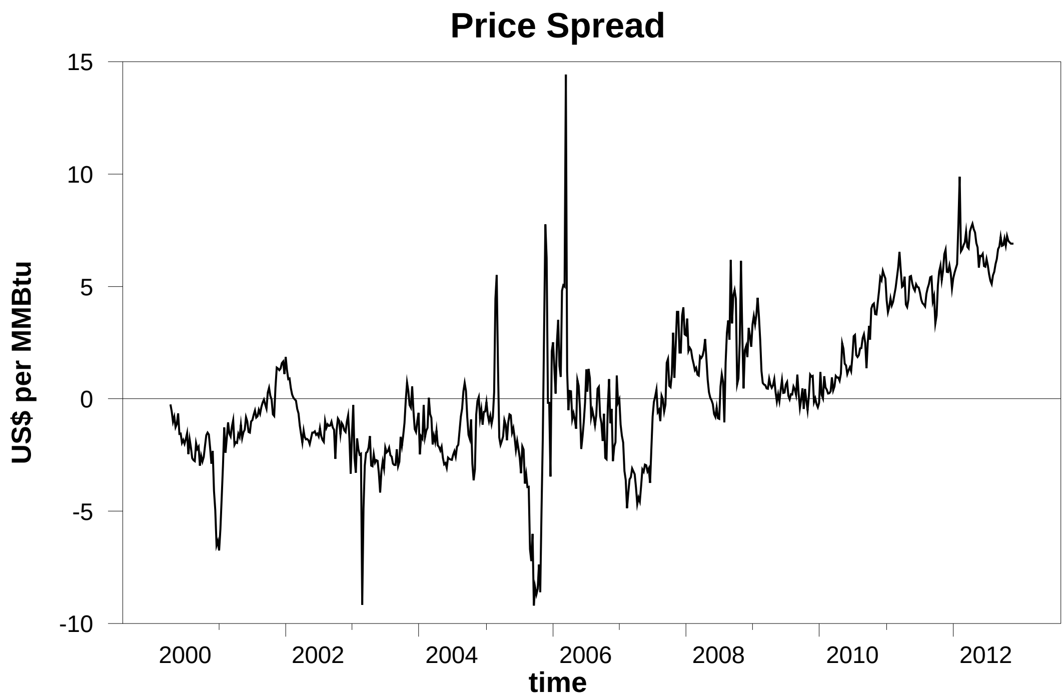


Figure 3.1: Price Difference between UK National Balancing Point and US Henry Hub

Up to now, the relationship between international gas prices has been mostly analyzed using linear cointegration analysis as done by Neumann (2009), Kao and Wan (2009) or Brown and Yücel (2009). Using US and UK gas price data up to the year 2008, all these studies find linear cointegration and therefore conclude that there is a certain degree of integration between the US and UK gas markets that may be attributed to spatial arbitrage via LNG volumes.[2] Most recently, Li et al. (2014) use data from 1997 to 2011 and employ a test for market convergence that is not related to the notion of cointegration. A subsequent Kalman filter analysis of price convergence is based on a linear relationship between gas prices. The authors find several converging gas prices around the globe. Though, no convergence between the US gas market and other gas markets can be found.

However, linear (cointegration) models may not capture arbitrage dynamics in a spatial setting adequately since they do not account for transaction costs - including transportation costs - that can impede arbitrage activity. The inability to consider transaction costs may distort the validity of the cointegration analysis and hence the corresponding economic interpretation, i.e., the degree of market integration. To overcome this methodological caveat, we use a threshold cointegration approach to test for price convergence. We use a threshold cointegration framework since this

[2] An exception is an early paper by Siliverstovs et al. (2005) who find no cointegration between the US and the European gas market from the early 1990s to 2004.

enables the explicit consideration of transaction costs and thus fits well to the specifics of the LOP in a spatial setting.

The focus of this study is on the relationship between the natural gas spot prices in the US and the UK. First, we use both linear and threshold cointegration tests to test for price convergence between the US and the UK natural gas price. In other words, we test whether we can find empirical evidence of both markets being integrated by arbitrage activity in accordance with the LOP. The threshold cointegration test is based on the estimation of a threshold autoregressive (TAR) model. The threshold parameter in the TAR model can be interpreted as transaction costs and other impediments to arbitrage activity such as transport cost as well as capacity and contractual constraints. Hence, by estimating the TAR threshold parameter we also obtain a measure for the transaction costs of arbitrage between the US and the UK market. Second, in case we find evidence of threshold cointegration, we use a threshold vector error correction model (TVECM) in order to investigate the dynamics that restore the long-run equilibrium. The TVECM thereby provides additional insights regarding the speed and direction of adjustment of the individual price series during the arbitrage process. Hence, we obtain an indication as to the degree of transatlantic gas market integration and as to how quickly arbitrage is pushing price convergence.

We conduct the steps described above for the period 2000 to 2012 and for the two sub-samples 2000 to 2008 and 2009 to 2012. The former sub-sample roughly corresponds to the samples used in most of the previous research and thus enables the comparison with studies using linear cointegration approaches. The latter sub-sample corresponds to the period when the UK and US gas prices are putatively decoupled. Linear and threshold cointegration tests both imply that natural gas markets are integrated in the full sample period 2000 to 2012. In the 2000 to 2008 sub-sample, there is empirical evidence of both linear and threshold cointegration of the UK and US price series. In the second sub-sample the linear cointegration test rejects the hypothesis of cointegration. However, using the threshold cointegration test that is more consistent with the LOP, we find empirical evidence for cointegration and thus market integration also in the second sub-sample.

The threshold estimates of our empirical approach provide an indication of the price differential that is necessary to trigger spatial arbitrage. While we find empirical evidence of arbitrage activity starting already at rather low price spreads in the first sub-sample, our results suggest that arbitrage is only triggered in the case of very high price spreads in the second sub-sample. Moreover, a comparison of our threshold estimates with LNG transport cost data reveals low non-transport cost impediments to arbitrage in the 2000 to 2008 sub-sample, but high non-transport cost impediments in the second sub-sample. The TVECM results suggest that between 2000 and 2008 both markets adjusted to restore the long-run equilibrium. In contrast, between 2009 and 2012, price convergence was exclusively achieved by movement in the UK price of natural gas.

The contribution of our study is threefold. First, we improve the econometric framework for studying price convergence and spatial arbitrage in global natural gas markets by specifying an econometric model that is more consistent with LOP theory than previous approaches. Second, our empirical approach reveals that, at least for the 2009 to 2012 sub-sample, linear and threshold cointegration frameworks deliver different results regarding price convergence and thus arbitrage activity. Third, our empirical results provide evidence of the US gas price and the UK gas price being still pulled together by arbitrage as implied in the LOP, although both prices seem to have decoupled since 2009.

The remainder of this paper is organized as follows: In Section 3.2, we show how the LOP can be represented by an econometric model explicitly accounting for transaction costs. Section 3.3 introduces the econometric procedures necessary to test for price convergence and arbitrage activity. We also present and discuss the empirical results in this Section. Section 3.4 concludes.

3.2 Theoretical Framework

3.2.1 The Law of One Price and its Application to Global Natural Gas Markets

Our study is grounded on the LOP, stating that the price for the same good should be equal in different markets. The LOP theory is based on the assumption of a homogenous good that can be resold to different markets without restrictions. In such a setting, the LOP is expected to hold since any price divergence triggers arbitrage and thus is only of transitory nature. The arbitrage conditions for traders participating in two markets i and j subject to transaction costs τ, can be stated as

$$\begin{aligned} P_1 > &\quad P_2 + \tau \\ P_1 < &\quad P_2 - \tau. \end{aligned} \tag{3.1}$$

Thus, arbitrage activity may only be triggered if the implied gross profit of the trade covers transaction costs. In a frictionless world without transaction costs, τ equals zero and thus drops out of the equation.

In our study, we empirically investigate a special case of regional price arbitrage in commodity markets, focusing on the natural gas prices in the United Kingdom (UK) and the United States (US). One has to keep in mind that direct price arbitrage, i.e., direct exports of gas from the UK to the US or vice versa, cannot the driving force of any price convergence between the two markets. In order to export natural gas it has to be liquefied in specialized liquefaction facilities. Neither in the US nor in the UK there are any significant liquefaction facilities. However, in both countries there are large regasification capacities that allow importing natural gas in the form of LNG and by transforming imported LNG back to gas form.

In fact, transatlantic convergence in gas prices may only be the result of two special forms of arbitrage. First, arbitrage can be carried out by re-exporting not yet regasified LNG received from long-term contracts (LTCs). Second, arbitrage can be carried out via the LNG spot market by a third party, namely the trader of LNG rerouting its shipments according to current market conditions.

In the global natural gas market, the majority of LNG shipments is based on long-term contracts. These trade flows represent constant deliveries regardless of current gas market prices or gas demand and supply balances. Hence, quantities traded via LTCs cannot contribute to arbitrage directly. However, LNG that is imported via LTCs can still contribute to arbitrage activity indirectly by re-export. Typically, LNG import terminals unload LNG and store it in liquid form before the LNG is eventually regasified and fed into the gas pipeline network. Stored LNG that is still in liquid form can be loaded back onto LNG carrier ships which in turn can be re-exported to other markets. Assuming equal regasification cost, this type of arbitrage will occur when the gas price at the terminal where the liquid LNG is stored is lower than the gas price in a foreign market minus the transaction cost of shipping the LNG to the foreign market. In 2009 when the putative decoupling of the US and the UK started, we also observe a steep increase in LNG re-exports from the US making the US the worldwide largest LNG re-exporting country in the year 2010 (see IGU 2014). Hence, the re-export of LNG from LTCs may constitute an important driver of any price convergence in the "decoupling" period.

Moreover, there is also a growing spot market for LNG where LNG volumes are traded on a short-term basis accounting for current regional gas prices and transaction costs. Within this spot market, the exporter of LNG is expected to serve the market where the greatest revenue (adjusted for transaction costs) can be obtained from selling the regasified LNG volumes. As a consequence, changes in regional supply and demand balances may represent an incentive for the LNG exporter to divert its spot volumes to other destinations. Thus, the rerouting of LNG spot market deliveries may constitute an effective element of spatial arbitrage in the natural gas market.

The arbitrage condition for third parties can be stated in terms of the prices in the potential destination regions and the respective transportation costs. For the sake of simplicity, we assume an exporter having the opportunity to deliver spot LNG volumes to two markets 1 and 2. In addition, market 1 is assumed to be the more remote market for the exporter, resulting in increased transportation costs when this market is served instead of market 2. As long as the regional price differential $P_1 - P_2$ does not exceed the difference in transaction costs, $\Delta TC_{1,2}$, all spot volumes are shipped to market 2. In contrast, market 1 is exclusively served when the transaction cost differential is covered by the greater revenues that can be generated by selling the gas to market 1. Equation (3.2) states the indifference condition for the arbitrageur with regard to the potential destinations:

$$\Delta TC_{1,2} = P_1 - P_2 \tag{3.2}$$

with

$$\Delta TC_{1,2} = TC_1 - TC_2 \tag{3.3}$$

where TC_1 and TC_2 denote the transaction costs for the exporter to market 1 and market 2, respectively. The situation when switching from market 1 to market 2 is profitable for the arbitrage player is illustrated in Figure 3.2.

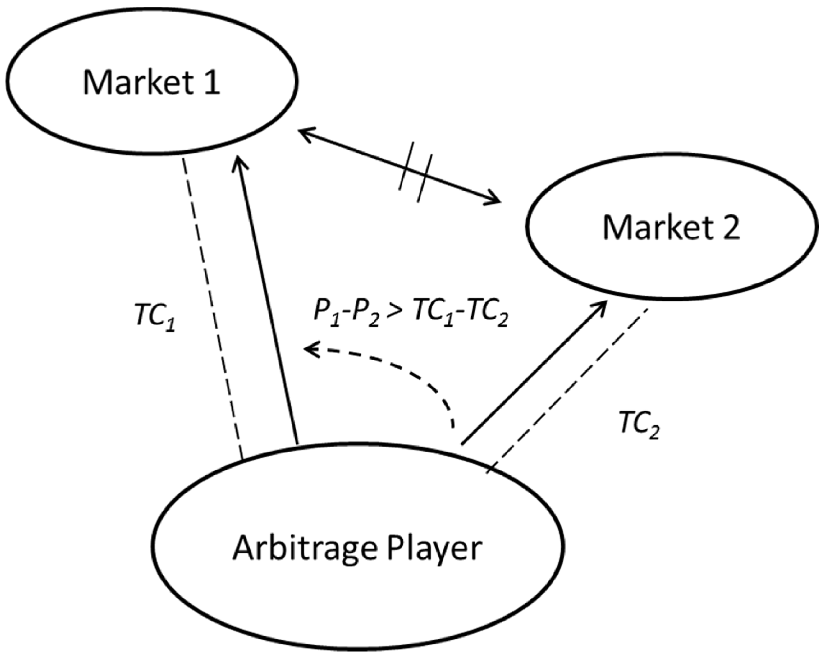


Figure 3.2: Regional Price Arbitrage in the Global Natural Gas Market

Even if LNG imports in some years of the last decade only constitute a minor share of overall US or UK gas consumption, LNG may still have an important impact on natural gas prices in the US and the UK. LNG related gas import is often more expensive than pipeline imports or domestic production. THerefore, LNG often is the marginal supply in UK or US gas markets and, thus, has an important impact on gas prices.

3.2.2 Econometric Approach

Disregarding transaction costs, the LOP requires that any deviation from the LOP is corrected by instantaneous arbitrage activity. Based on this assumption, deviations of the prices in the two markets under consideration cannot be persistent. From an econometric perspective, this requires the stationary behavior of the price spread $s_t = P_1 - P_2$. Since the LOP calls for the unity of prices in the long run equilibrium, the price spread reverts to an expected value of zero. The time series behavior of the price spread can be represented by the following autoregressive process:

$$s_t = (\varrho + 1)(s_{t-1}) + \epsilon_t \tag{3.4}$$

or equivalently

$$\Delta s_t = \varrho(s_{t-1}) + \epsilon_t \tag{3.5}$$

with $-1 < \varrho < 0$ to ensure stationarity of the process and ϵ_t as an error term with zero mean and finite variance.

The existence of transaction costs causes a "band of no arbitrage" between both markets as stated in equation (3.1). Within this band, the price spread s_t does not cover the transaction costs of the arbitrage activity, τ. In contrast, if the absolute value of s_t exceeds τ, arbitrage is profitable resulting in an adjustment of the prices towards a price spread of zero. However, the price spread never actually becomes zero as the adjustment towards the equilibrium stops when the price spread is equal to transaction costs and arbitrage is not profitable anymore. From this point on, the two prices move independently until a shock drives the prices apart and makes arbitrage profitable again. Thus, in the presence of transaction costs, the short-run behavior of s_t may be more adequately captured by a threshold autoregressive model (TAR) than by an ordinary autoregressive process:

$$\Delta s_t = \begin{cases} \gamma_1(s_{t-1} - \tau) + \epsilon_t, & \text{if } s_{t-1} > \tau \\ \epsilon_t, & \text{if } |s_{t-1}| \leq \tau \\ \gamma_3(s_{t-1} + \tau) + \epsilon_t, & \text{if } s_{t-1} < -\tau \end{cases} \tag{3.6}$$

where γ_1 and γ_3 are the adjustment coefficients and ϵ_t is an error term with zero mean and finite variance. The representation in equation (3.6) shows that within the band of no arbitrage, the price spread follows a random walk without drift, that is, it does not exhibit any mean reverting behavior. In contrast, in the outer regimes, the price spread follows a stationary process reverting to the threshold τ. Balke and Fomby (1997) who first introduced this type of TAR-model call the corresponding stochastic process a band threshold autoregressive (BAND-TAR) process. In our empirical application, τ represents the time-invariant transaction costs of arbitrage.[3] However, we allow for different threshold values in our sub-samples. Note, that our definition of transaction cost is rather broad. Accordingly, transaction cost include transport cost as well as any more intangible cost of arbitrage such as the cost of information procurement, as well as the implict cost of contractual and technical constraints. Since there are many countries and firms acting as suppliers and demanders in the global LNG market, we do not see the necessity to account for non-competitive behaviour which could also lead to price differentials.

In general, the model specified in equation (3.6) accounts for asymmetric dynamics in the adjustment process. This means that the adjustment behavior may be different depending on whether the price spread is positive or negative. The specific case of symmetric adjustment, i.e., identical adjustment behavior regardless of the direction of the equilibrium deviation, is nested in equation (3.6) by the coefficient restriction $\gamma_1 = \gamma_3$.

Accounting for transaction costs, the threshold vector error correction model (TVECM) seems thus well suited for the empirical investigation of arbitrage between two markets. We thus specify the following TVECM to investigate arbitrage dynamics between the two markets:

$$\Delta P_{t,1} = \begin{cases} \gamma_1^{(1)}(s_{t-1} - \tau) + \epsilon_{1t}, & \text{if } s_{t-1} > \tau \\ \epsilon_{1t}, & \text{if } |s_{t-1}| \leq \tau \\ \gamma_1^{(3)}(s_{t-1} + \tau) + \epsilon_{1t}, & \text{if } s_{t-1} < -\tau \end{cases} \tag{3.7}$$

$$\Delta P_{t,2} = \begin{cases} \gamma_2^{(1)}(s_{t-1} - \tau) + \epsilon_{2t}, & \text{if } s_{t-1} > \tau \\ \epsilon_{2t}, & \text{if } |s_{t-1}| \leq \tau \\ \gamma_2^{(3)}(s_{t-1} + \tau) + \epsilon_{2t}, & \text{if } s_{t-1} < -\tau \end{cases} , \tag{3.8}$$

where the threshold τ equals the τ from the a univariate TAR model and ϵ_{1t} and ϵ_{2t} are error terms with zero mean and finite variance. In equation (3.7) and equation (3.8), prices stop to

[3] We assume time-invariant transaction cost, since we are not aware of significant changes in LNG shipping costs over the sample period.

adjust towards the long-run equilibrium when the price spread drops below the transaction cost threshold. The TVECM accounts for asymmetric error correction, meaning that the adjustment of each individual price is allowed to differ with respect to the sign of the price spread. However, the specifications in equation (3.7) and equation (3.8) also nest the symmetric case via the coefficient restrictions $\gamma_1^{(1)} = \gamma_1^{(3)}$ and $\gamma_2^{(1)} = \gamma_2^{(3)}$. With this restriction in place, we assume that prices in both markets adjust with the same speed irrespective of whether the price spread is positive or negative.

3.3 Empirical Application

This section comprises four subsections and presents the econometric estimation and testing procedures as well as the corresponding empirical results. The first subsection presents the data and variables. In the second subsection, we test for market (co-)integration in the presence of threshold effects as implied by the LOP. In the third subsection, we estimate threshold vector error correction models to investigate how the US and the UK prices adjust in order to restore equilibrium. Finally, we interpret the estimation results from an economic perspective and draw conclusions with regard to the degree of transatlantic gas market integration.

3.3.1 Data

We use weekly price data for the period April 2000 to November 2012. The Henry Hub (HH) spot price, measured in US-dollar per million British thermal units (MMBtu) is used as the US price of natural gas. The National Balancing Point (NBP) spot price which is measured in pound sterling per kilowatt-hour is used as the UK price. For the sake of comparability, the UK price is converted to US dollars per MMBtu using appropriate physical conversion factors and the weekly exchange rate published by the Bank of England.

In order to investigate whether or not there is still price convergence after the "decoupling of gas prices", we perform all our empirical approaches for the full sample and for the sub-samples 2000 to 2008 and a 2009 to 2012. The starting date of the latter sub-sample coincides roughly with the so called "decoupling" of the US and UK gas market and allows for an adequate number of observations in the second sub-sample.

In contrast to other studies of spatial arbitrage such as Lo and Zivot (2001), we do not use the logarithm of the natural gas prices but the price levels themselves. Using the logarithms is a typical transformation done in order to get the data more consistent with the assumptions needed for efficient estimates. However, taking the logarithms of both prices effectively assumes an isoelastic relationship between the variables. In our view, an isolelastic relationship is not in complete accordance with the LOP, most likely not even locally in the range of values of our data.[4] Thus, we argue that only a specification using actual price levels is in line with LOP theory. LOP theory implies that in the long run, prices should be equal and thus tied together by a linear relationship.

3.3.2 Threshold Estimation and Testing for (Threshold) Cointegration

First and foremost, it has to be determined whether UK and US natural gas prices share a long-run equilibrium relationship. Without the two time series being cointegrated, any further analysis of adjustment processes of individual price series would be in vain. Only if UK and US prices are cointegrated, error correction, price convergence and hence market integration will occur. For the purpose of comparison, we use linear as well as threshold cointegration tests. The analysis involves several sequential steps. First, unit root tests are used to determine whether both price series are integrated of the same order as this is the prerequisite for cointegration. Table 3.1 shows the results from augmented Dickey-Fuller (ADF) unit root tests for the levels and the first differenced versions of the natural gas prices. The results imply that both of the prices series are integrated of order one.

Table 3.1: Unit Root Tests for US and UK Natural Gas Spot Prices

	UK price	Δ UK price	US price	Δ US price
ADF	-1.101	-23.61 ***	-1.12	-23.44 ***

Results for unit root tests with a null hypothesis of a unit root and a maximum lag length of $T^{1/3} = 9$ The Schwarz Criterion is used in order to select the lag lengths for the ADF unit root test. A rejection of the null hypothesis of a unit root at the 1, 5 and 10 percent significance level is denoted by ***, ** and *, respectively.

As a second step cointegration tests are used to determine whether the prices are tied together by a long-run relationship over the sample periods. Initially, linear cointegration tests are used to make our results comparable with previous studies and to provide a reference for the threshold cointegration tests. When testing for cointegration, we use the structural economic information that, according to LOP theory, the prices are expected to be equal in the long run. The corresponding prespecified cointegration vector is thus (1,-1). In other words, if the prices are tied together as implied by the LOP the price spread should be a stationary series that reverts to a mean of zero. Therefore, we can test for cointegration by testing whether the price spread series contains a unit

[4] In addition, as can be seen in Lo and Zivot, a log-log specification implies arbitrage conditions that we do not consider as a convincing representation of the real circumstances. In specific, their arbitrage conditions imply that the threshold, representing transaction costs and other frictions, is proportional to the revenues from arbitrage.

root.[5] In our case, the null hypothesis that the price spread $s_t = p_{UK,t} - p_{US,t}$ is a non-stationary, unit root process is tested against the alternative hypothesis of a linear autoregressive process. Thus, the cointegration test is essentially an ADF test that is based on the estimation of the following equation.

$$\Delta s_t = \varrho(s_{t-1}) + \sum_{i=1}^{M} \Delta s_{t-i} + \epsilon_t \tag{3.9}$$

with $-1 < \varrho < 0$ to ensure stationarity of the process.

If ϱ is not significantly different from zero, the price spread s_t will be exclusively driven by its current error term. In other words, s_t follows a unit root process and the US and the UK prices are considered as not cointegrated. If ϱ is different from zero at conventional significance levels, the null hypothesis of non-stationarity will be rejected and the process is stationary. Stationarity of the price spread implies a stable a long-run equilibrium relationship of the two prices. According to the Engle and Granger (1987) representation theorem, cointegration requires a significant error correction process. In this context, error correction describes the behavior that whenever a shock drives the prices out of their long-run equilibrium, at least one of the prices adjusts in order to restore the equilibrium.

Table 3.2: Cointegration Test for the Prespecified Cointegration Vector (1,-1)

	Price spread 2000-2012	Price spread 2000-2008	Price spread 2009-2012
Linear Cointegration	-3.4325 ***	-5.8921 ***	-0.0859

Results for unit root tests with a null hypothesis of a unit root and a maximum lag length of $T^{1/3} = 9$ Critical values are taken from Engle and Yoo (1987). The Schwarz Criterion is used to select the lag lengths for the ADF unit root test. A rejection of the null hypothesis of a unit root at the 1, 5 and 10 percent significance level is denoted by ***, ** and *, respectively.

Table (3.2) shows the cointegration test results on the prespecified cointegration relationship $s_t = p_{UK,t} - p_{US,t}$. The tests for the full period 2000 to 2012 and for the first sub-sample 2000 to 2008 reject the null hypothesis of a unit root and thus provide evidence of cointegration. In contrast, the results for the period 2009 to 2012 indicate no cointegration.

As already outlined, standard testing for cointegration with known cointegration vectors can be performed by testing the null hypothesis of a unit root against the alternative hypothesis of linear

[5] We abstain from using Johansen and Juselius (1992) or Engle and Granger (1987) cointegration tests. These tests ignore that theory clearly states a cointegration vector of (1,-1) and instead estimate the parameters of the long-run relationship. Thereby, additional uncertainty is introduced that is reflected in critical values that are too conservative. In our case, the testable long-run relationship (1,-1) can clearly be derived from LOP theory. Therefore, we directly test the known relationship for (non-)stationarity as suggested by Horvath and Watson (1995).

cointegration. In specific, linear in this context means that independent of the size of the deviation from the equilibrium, the respective price adjusts proportionally to the size of the deviation.

In contrast, LOP theory suggests that if the absolute value of the price difference between two markets is smaller than transaction costs arbitrage is not profitable. Therefore, in such periods, there will be no arbitrage and hence no price adjustment. Thus, there should be a set of "small" price spreads where no error correction occurs. As outlined in Section 3.2, this set of price combinations comprises a symmetric band around the long-run relationship that is confined by two symmetric thresholds representing transaction costs and other impediments to arbitrage. Only if the price spread is larger in absolute terms than the threshold value, there will be a reversion of the price spread towards the long-run relationship. Balke and Fomby (1997) and Enders and Granger (1998) argue that traditional tests for unit roots and cointegration have low power in the presence of asymmetric adjustment. Particularly, if a process is characterized by threshold cointegration and threshold error correction this can strongly decrease the power of conventional linear cointegration tests. In order to use a test that explicitly takes account of cointegration in the presence of threshold error correction we follow Enders and Granger (1998) and Enders (2001) and employ a cointegration test that is based on the band threshold autoregressive process. The econometric representation of the BAND-TAR model is shown in equation (3.10).

$$\Delta s_t = \gamma_1(s_{t-1} - \tau)[s_{t-1} > \tau] + \gamma_2(s_{t-1} + \tau)[s_{t-1} < -\tau] + \sum_{k}^{K} \Delta s_{t-k} + \epsilon_t \tag{3.10}$$

Equation (3.10) resembles the BAND-TAR of equation (3.6). The changes in the price spread Δs_t is explained by a tripartite process. In each time period, Δs_t follows only one of the three "sub"-processes, either an autoregressive process $\gamma_1(s_{t-1} - \tau)$ if the lagged price spread s_{t-1} is above the threshold value τ, or an autoregressive process $\gamma_2(s_{t-1} + \tau)$ if s_{t-1} is smaller than $-\tau$, or a random walk if the absolute value of s_{t-1} is smaller than τ, respectively. A number of lags of the dependent variable K is included to avoid residual autocorrelation that would render the parameter estimation inconsistent.

The parameters of equation (3.10) including the value of the threshold parameter τ are estimated with the iterative grid search method proposed by Chan (1993). Accordingly, the equation is estimated several times with OLS - each time using a different fixed value of the threshold variable s_{t-1} for τ.[6] Each time the sum of squared residuals (SSR) is stored. The value of the threshold variable that yields the lowest SSR is regarded as the final estimate of τ. This value of τ is used to actually

[6] As recommended by Chan, the lowest and the highest 15 percent of the values of s_{t-1} are not used as potential threshold values to provide meaningful results.

estimate the parameters. The number of lags of the dependent variable is chosen by the Schwarz information criterion. The approach gives consistent estimates of the parameters of equation (3.10) and a rate T consistent estimate of the threshold parameter τ (Chan 1993).

After estimating equation (3.10), the presence of threshold cointegration can be tested. In general, testing for threshold cointegration is similar to testing for linear cointegration. In specific, equation (3.10) is the BAND-threshold nonlinear counterpart of the ADF equation used for the linear cointegration tests. If the price spread, that is, the assumed long-run relationship (1,-1) between the UK and US natural gas price were non-stationary, we will expect $\gamma_1 = \gamma_2 = 0$. Accordingly, the price spread then is only explained by the error term, s_t is a unit root process and the US price and the UK price of natural gas are not cointegrated. The alternative hypothesis is that $\gamma_1 \neq 0$ and $\gamma_2 \neq 0$. The threshold cointegration test is performed using an F-Test for the joint restrictions implied by the null hypothesis. As the threshold parameter is not identified under the null hypothesis, the critical values for the test are non-standard. Unfortunately, no critical values for the general version of the test equation (3.10) are available. However, approximate critical values are available when we use a more restricted version of the equation for the test that closely resembles the two regime threshold cointegration test developed by Enders and Granger (1998). In order to use the critical values published in Enders and Granger (1998) we restrict equation (3.10) by setting $\gamma_1 = \gamma_2 = \gamma_{arbitrage}$ and hence effectively create a two regime threshold cointegration test.[7] This restriction is equivalent to the mild assumption that the speed of adjustment of the price spread to the attractor τ is equal regardless of whether the price spread is positive or negative.

Further, if evidence of cointegration is found, we can test whether the corresponding error correction process has indeed threshold nonlinear character. Linear adjustment, as opposed to threshold nonlinear adjustment, means that the speed of adjustment in all regimes is equal. Thus, it has to be investigated whether the speed of the adjustment in the arbitrage regime is equal to the speed of adjustment in the no arbitrage regime. The speed of adjustment in the no arbitrage regime is zero as implied by the random walk. Accordingly, the null hypothesis of zero adjustment in the arbitrage regime has to be tested, which is the same as the null hypothesis in the threshold cointegration test. Hence, the corresponding test statistic is the same F-statistic that we obtained for cointegration test. However, if we found evidence for threshold cointegration we have to compare the F-statistic

[7] The autoregressive coefficient for the arbitrage regime is given by $\gamma_{arbitrage}$. The inner no arbitrage regime follows a random walk. Even with this restriction in place our test equation is not perfectly equal to the equation used by Enders and Granger. However, we conjecture the critical values for our case will not differ greatly from the Enders Granger critical values. At least, given the size of our threshold cointegration test statistics obtained in the estimations, small deviations from the appropriate critical values should not be of importance for the validity of the test results. Moreover, we also used the threshold cointegration test with bootstrapped critical values proposed in Chapter 4. The OLS based threshold cointegration test and the LAD based test rejected the null of no cointegration at the one percent significance level. The t-ML based test rejected no cointegration at the 10 percent significance level.

to critical values of the standard F-distribution. If the null of no adjustment, $\gamma_{arbitrage} = 0$, can be rejected this is regarded as evidence of threshold nonlinear error correction.

In the subsequent paragraphs, the threshold estimates and test results are presented. A comprehensive economic discussion of the results follows below in Section 3.3.4. The threshold estimates for the full sample 2000 to 2012 and the two sub-samples 2000 to 2008 and 2009 to 2012 are given in the top row of Table 3.3. The estimates for the other parameters of equation (3.10) are left out for conciseness. In the full sample, arbitrage activity and reversion of the price spread to the threshold only happens when the price spread is above 4.72 US-Dollar per MMBtu in absolute terms. Interestingly, the threshold estimate for the first sub-sample 2000 to 2008 is low at 2.89 US-Dollar per MMBtu. In contrast, the threshold estimate for 2009 to 2012 is by far higher at 6.56 US-Dollar per MMBtu, implying that only very high price spreads in absolute terms lead to a reversion of the price spread towards the long-run equilibrium.

Table 3.3: Tests for Cointegration and Threshold Nonlinearity

		2000-2012	2000-2008	2009-2012
Threshold value		4.72	2.89	6.56
Threshold cointegration and nonlinearity test	Test statistic	67.18	72.51	14.27
Threshold cointegration test	Significance level	1%	1%	1%
Nonlinearity test	Significance level	1%	1%	1%

The estimates of the threshold values are measured in US-Dollar per MMBtu. The critical values used to obtain the significance levels for the threshold cointegration test are from Enders (2001). The critical values for the threshold non-normality test correspond with the usual F-distribution.

The second and the third row of Table 3.3 show the test statistics and the significance levels of the threshold cointegration tests. Accordingly, there is strong evidence of cointegration in the full sample and in each of the sub-samples as the null hypothesis of a unit root can be rejected. The fourth and fifth row of Table 3.3 show the results for the threshold nonlinearity tests. The null hypothesis of a linear process can clearly be rejected at the 1 percent significance level in all sample periods. This can be regarded as strong evidence of threshold nonlinearity and, hence, threshold cointegration of the US and UK natural gas prices.

3.3.3 TVECM Estimation - Adjustment of Individual Prices

Up to now, we found evidence of the UK and the US natural gas prices being cointegrated. In order to get a better understanding of the adjustment process of the two individual price series we estimate a threshold vector error correction model. This model allows us to investigate how the individual price series adjusts when the UK and the US prices are not in the long-run equilibrium and arbitrage is profitable. According to the LOP, there should only be arbitrage and hence statistically significant adjustment of at least one of the two price series if the absolute value of the price spread is greater than the threshold value representing transaction costs. The following system of equations represents the TVECM model applied.

$$
\begin{aligned}
\Delta p_{UK,t} &= \gamma_{UK,high}(s_{t-1} - \tau)[s_{t-1} > \tau] + \gamma_{UK,low}(s_{t-1} + \tau)[s_{t-1} < -\tau] + \sum_{k=1}^{L_1} \Delta p_{UK,t-k} \\
&\quad + \sum_{j=1}^{L_2} \Delta p_{US,t-j} + \epsilon_{UK,t} \\
\Delta p_{US,t} &= \gamma_{US,high}(s_{t-1} - \tau)[s_{t-1} > \tau] + \gamma_{US,low}(s_{t-1} + \tau)[s_{t-1} < -\tau] + \sum_{m=1}^{L_3} \Delta p_{US,t-m} \\
&\quad + \sum_{l=1}^{L_4} \Delta p_{UK,t-l} + \epsilon_{US,t}
\end{aligned}
\tag{3.11}
$$

Equation (3.11) resembles the BAND-TVECM equations (3.7) and (3.8). The changes in natural gas prices for UK and US are each explained by two error correction terms and a set of lagged explanatory variables. The "high" and "low" regimes are defined by the price spread being larger than the positive threshold value τ or smaller than the negative threshold value $-\tau$. In accordance with Enders (2008), the value of τ in each sample period is set equal to the threshold estimates from the TAR model given in Table 3.3. We do not allow for arbitrage and adjustment if the absolute value of the price spread is smaller than the threshold value τ. We estimate a symmetric as well as an asymmetric adjustment version of the BAND-TVECM since economic theory does not make a clear a priori case for one of the two model versions. The more general, asymmetric adjustment specification is represented by equation (3.11). For the symmetric adjustment version, we restrict $\gamma_{l,high} = \gamma_{l,low}$ for $l = UK, US$. The set of lags of the dependent variable and the other explanatory variables is determined by the sequential elimination (SE) algorithm as outlined in Lütkepohl (2004).[8] This procedure is repeated until no further reductions in the information criterion are possible by eliminating any variable lags. The model specification that results from the SE procedure is used in the actual estimations. For each TVECM, the sequential elimination procedure is started with 18 initial lags of the lagged dependent and the other lagged explanatory

[8] The SE procedure leads to a reduced number of parameters that have to be estimated and, thus, to a more efficient estimation. The SE procedure first estimates the system of equations by feasible generalized least squares (FGLS) with a certain maximum lag length for the explanatory variables. Hereafter, the explanatory variable whose elimination leads to the largest decrease in an information criterion is eliminated from the system and the system is estimated again with a zero restriction placed on the respective variable. We use the Schwarz information criterion in the SE procedure. When the lag restrictions placed on the system of equations by the SE procedure using the Schwarz criterion resulted in autocorrelated residuals, then the less restrictive Akaike criterion was employed in the SE procedure. This always resulted in residuals that were free from autocorrelation.

variables in each equation. Due to the partly long lag structure, the final specification is not shown here but can be obtained from the authors upon request. The final model is estimated by GLS.[9] The estimated adjustment coefficients for the full sample, and the two sub-samples for the final specification after the sequential elimination procedure are shown in Table 3.4.[10]

The estimation results for the full sample symmetric BAND model show that once the price spread is above the threshold value in absolute terms, there is significant adjustment to the long-run equilibrium for both the US and the UK prices of natural gas. The adjustment of the UK price of 25 percent per week is much larger than the 5.9 percent adjustment of the US price. The estimates for the more general asymmetric BAND model support the symmetric version in the sense that there is adjustment in both prices. In addition, the asymmetric model provides us with more refined insights about the error correction process. In specific, adjustment in the UK price only takes place when the price spread is above the threshold. In contrast, there is no adjustment if there is a negative price spread, that is, the US price is higher than the UK price. Similarly, the US price only shows significant adjustment when the price spread is negative and below the negative threshold value. The adjustment of the US price in the asymmetric model is much stronger at 28.9 percent than in the symmetric specification.

Table 3.4: Results for BAND-TVECM Estimations

Period	Model	Regime	Δp_{US}	t-value	Δp_{UK}	t-value
2000-2012	Symmetric		0.059***	(2.307)	-0.250***	(-5.034)
	Asymmetric	high	-0.007	(-0.229)	-0.305***	(-5.264)
		low	0.289***	(6.283)	-0.047	(-0.524)
2000-2008	Symmetric		0.094***	(3.532)	-0.460***	(-8.997)
	Asymmetric	high	-0.037	(-0.789)	-1.000***	(-12.012)
		low	0.152***	(4.84)	-0.154***	(-2.676)
2009-2012	Symmetric		0.009	(0.168)	-0.387***	(-3.264)
	Asymmetric	high	0.009	(0.168)	-0.387***	(-3.264)
		low	-	-	-	-

Lag order selected with SE algorithm using the by Schwarz information criterion. Where the Schwarz criterion results in autocorrelated error terms, the Akaike information criterion was used because it allows for more generous lag lengths. */**/*** attached to coefficients signify that the coefficient is significantly different from zero at the 10%, 5% or 1% level, respectively. t-values are given in brackets.

As in the full sample, in the symmetric BAND model for the 2000 to 2008 sub-sample both prices adjust significantly. Again, the UK price with 46 percent shows much stronger error correction dynamics than the US price with only 9.4 percent. The results for the asymmetric BAND-TVECM for 2000 to 2008 are qualitatively similar to the results for the full sample. The UK price is significantly adjusting as long as arbitrage is profitable. The adjustment of the UK price in the positive

[9] GLS estimation is necessary because OLS would lead to a less efficient estimation in the presence of a set of lagged dependent and other explanatory variables that faces different restrictions in each equation of the system.

[10] All estimated models were tested for residual autocorrelation with LM tests as proposed by Lütkepohl (2004). No evidence of autocorrelation could be detected in any of the regressions. However, residual non-normality tests point to a decreased estimation efficiency. Autocorrelation and non-normality test results are not shown here for conciseness.

price spread regime is very strong at 100 percent, meaning that a positive price spread will be fully corrected after one week if the positive price spread is above the threshold value. In contrast, the US price adjusts only significantly at a rate of 15 percent when arbitrage is profitable and the US price is above the UK price, i.e., the price spread is negative.

The results for the 2009 to 2012 sub-sample differ substantially from the results for the period 2000 to 2008. In this period, the results for the symmetric and the asymmetric TVECM are equal as there were no negative price spreads that where below the negative threshold value. The estimation results indicate that there is significant and substantial adjustment only in the UK price.

3.3.4 General Discussion

In this subsection, we interpret and discuss the empirical results obtained from our econometric estimations. As outlined above, the LOP implies price convergence which is reflected in threshold error correction and cointegration in the econometric modeling. In contrast, linear cointegration tests ignore the potential threshold property of an adjustment process caused by transaction costs. This difference may be reflected in the results for linear cointegration tests in Table 3.2 and threshold cointegration tests in Table 3.3. Whereas both linear and threshold tests find cointegration in the full sample and in the 2000 to 2008 sub-sample, the test results for the later sub-sample differ. In contrast to the linear cointegration test that finds no evidence of cointegration the threshold cointegration test strongly supports cointegration in the period 2009 to 2012. This finding provides empirical evidence against the notion of a "decoupling" of gas markets in recent years, at least if decoupling refers to the fact that the LOP does not hold anymore. Threshold nonlinearity tests further indicate that a threshold framework indeed improves previous linear cointegration approaches because it is a more appropriate model to capture adjustment dynamics of benchmark gas prices in different markets.

By and large, LNG trade is the only way of arbitrage between the US and the UK natural gas market. The typical transport cost differential of LNG transports from major exporting countries such as Qatar to US destinations compared to UK destinations is below 2 US dollar per MMBtu.[11] During the last decade this LNG transport cost differential presumably has not changed substantially. For example, Maxwell and Zhu (2011) argue that LNG tanker costs change only gradually over time and that these costs affect LNG imports only in the long run. Therefore, comparing our threshold values between the two sub-samples provides an indication of the magnitude of non-transport cost related

[11] Market information providers such as Platts or ICIS differ somewhat in their measurement of LNG transport cost. However, two US dollar per MMBtu seems to be a reasonable upper bound for the US-UK transport cost differential from the most relevant exporting regions.

impediments for arbitrage activity. Table 3.3 shows that the threshold estimates differ widely (of 2.89 US dollar per MMBtu for the period 2000 to 2008 and 6.56 US dollar per MMBtu for the period 2009 to 2012, respectively). This can be regarded as evidence of increasing impediments to arbitrage in the Atlantic gas market other than transport costs in the later period. Potential sources of these impediments are capacity constraints for LNG import and export. This could be an explanation between 2009 and 2012 when UK LNG imports reached unprecedented levels and possibly also UK LNG import capacity limits. Another potential explanation may be increased Japanese LNG demand and prices after the Fukushima disaster in the year 2011. The spike of Asian gas prices may have created higher opportunity costs of delivering LNG quantities to the Atlantic basin. In addition, also contractual structures and other non-fundamental factors can be regarded as a potential explanation for the increase in the transatlantic price spread. Further, technical constraints may also play a role.

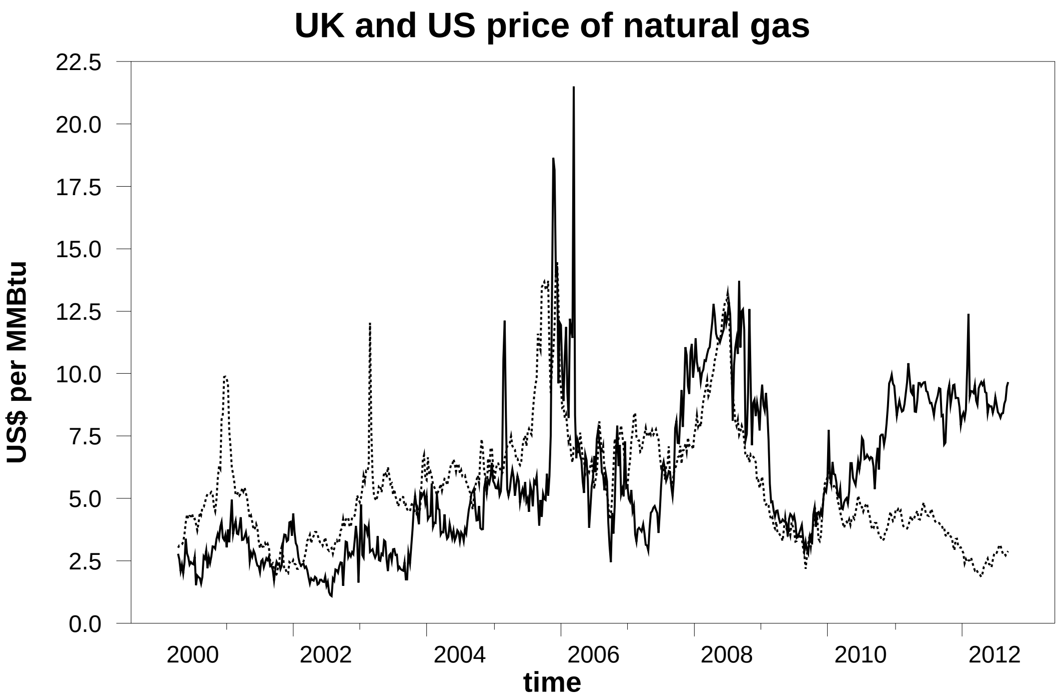


Figure 3.3: Prices for UK National Balancing Point (Solid Line) and US Henry Hub (Dashed Line)

The results from the TVECM estimation presented in Table 3.4 allow for a more detailed look on the adjustment behavior of individual prices. In the symmetric model versions in all (sub-)samples, the UK price is adjusting stronger than the US price. Further, the results for the asymmetric adjustment models reveal an interesting pattern. The adjustment of prices tends to be statistically significant and strongest when the disequilibrium is resulting from a market-specific positive price shock in the respective market. As shown in Figure 3.3, positive price shocks can be frequently observed in both the US and the UK natural gas markets. Most of these price spikes result from increased demand resulting for instance from unexpected cold spells or economic upswings as well as from perilous events on the supply side such as the US hurricane season in the year 2005. In line with the results from the symmetric model, the asymmetric adjustment in the UK price is substantially stronger for the UK price than for the US price in the full sample and the sub-sample 2000 to 2008. In the period 2009 to 2012 the US price shows no significant adjustment which is in line with the fact that US LNG imports and US gas prices have dwindled to low levels in that period due to the shale gas boom. However, we observe significant downward adjustment of the UK price when the UK price is above the US price. This situation was particularly prevalent in the late year 2011 and during the year 2012, when UK gas prices and UK LNG imports were both high. Thus, in the 2009 to 2012 period, price convergence in accordance with the LOP is mainly induced by downward pressure on the UK price.

3.4 Conclusion

The term “decoupling of gas markets” has been frequently used to describe the observation that benchmark natural gas prices such as the US and the UK price have diverged since the beginning of the year 2009. Given that transport costs have been far lower than these observed price spreads, this situation seems to contradict the LOP at first glance. In this study, we used a threshold cointegration framework to empirically investigate whether there is empirical evidence of the LOP still to hold between the US and UK natural gas market.

After we outlined the properties that made a threshold cointegration model more appropriate than linear cointegration to study the LOP in gas markets, we used a linear and a threshold cointegration model to test for cointegration as well as the presence of threshold nonlinearity. In contrast to the linear cointegration tests, we find strong evidence in favor of cointegration and price convergence as well as threshold nonlinearity in all our (sub-)samples using the threshold cointegration approach. Thus, although transatlantic natural gas prices seem to have become disconnected in recent years, they are still pulled together by arbitrage dynamics as implied by the LOP. Our threshold estimates constitute a measure for transaction costs and other impediments to arbitrage. Interestingly, the threshold in the period 2009 to 2012 is found to be much larger than the typical LNG transport cost differential between US and UK market. This potentially indicates the emergence of very high non transport transaction costs and other arbitrage impediments such as capacity constraints in LNG market. Hence, the decomposition of the threshold values into their different components is a very promising field for further research. The estimates of the threshold vector error correction models suggest that between 2000 and 2008, both US and UK prices adjusted significantly to restore an equilibrium. In contrast, between 2009 and 2012, when UK prices were mostly far above US prices, we find downward pressure on the UK price, but no adjustment in the US price. Taken together, our econometric estimations provided evidence of the LOP to hold even in the period 2009 to 2012 when US and UK gas prices seemed to have decoupled. Moreover, our empirical results indicate the presence of high non-transport transaction costs and other impediments to arbitrage in recent years that substantially decrease the speed of price convergence compared to the period 2000 to 2008.

Chapter 4

Robust Testing of the Law of One Price in Natural Gas Markets

4.1 Introduction

The law of one price (LOP) states that due to arbitrage activity the prices of a homogenous good in two connected markets should converge towards a price spread of zero. When the LOP holds the markets under consideration are said to be integrated. In the presence of transaction costs, the LOP further states that the price spread stops to converge when the spread has become as low as the transaction costs and arbitrage has become unprofitable. Then, prices move independently until the price spread eventually exceeds the transaction costs and arbitrage activity leads to price convergence again. Econometrically, the economic long run equilibrium of price equality corresponds to a cointegration relationship of $p_i = p_j$ for the prices in the markets i and j. According to the LOP with transaction costs, the adjustment of prices towards this long run equilibrium, i.e., the rate of error correction is non-linear with respect to the price spread. Prices may only adjust and the prices spread only converges towards zero if the price spread is above a certain threshold value that represents transaction costs.

A threshold autoregressive (TAR) model (see Tong 1978) is a simple econometric model for this threshold non-linear adjustment of the price spread. TAR models can be specified such that the dependent variable, i.e., the price spread only reverts towards zero if it is larger than a threshold value representing transaction costs, see Balke and Fomby (1997) and Lo and Zivot (2001). In the context of the LOP, TAR models can be used to investigate whether two markets are integrated. If a unit root in the price spread can be rejected against the alternative of a TAR model, we can conclude that the prices are (threshold) cointegrated and that the two markets are economically integrated.

Moreover, TAR models can be used to estimate the transaction costs of arbitrage. Transaction cost estimates are explicitly given by the estimates of the threshold parameter.

In markets for natural resources, agricultural products and many other commodities price spikes are a frequent phenomenon which, statistically, are reflected in non-normal error distributions, e.g., fat tailed or additive outlier distributions. Such occasional extreme price movements potentially create large small sample biases in OLS estimates of TAR parameters. Particularly, the threshold estimates, representing transaction costs, may be severely biased with usual sample sizes. Correspondingly, extreme prices movements may result in a low power of TAR based cointegration tests and researchers may draw false conclusions about the economic integration of two markets. Therefore, in this essay we consider the problem of TAR based transaction cost estimation and cointegration testing that is robust to the presence of occasional extreme price movements.

Our paper draws upon and is situated in the intersection of roughly three fields of econometric research: non-linear adjustment models, cointegration testing and outlier robust estimation. Several publications have investigated two of the three fields at a time. First, estimation and cointegration testing with nonlinear adjustment models. Examples which address this issue with threshold non-linear adjustment models are Caner and Hansen (2001), Enders and Granger (1998) or Seo (2008). Second, robust estimation of nonlinear adjustment models in the presence of outliers (e.g. see Dijk et al. 1999, Chan and Cheung 1994 or Zhang et al. 2008).[1] Third, robust testing for unit roots as in Lucas (1995b), Lucas (1995a), Herce (1996) or Carstensen (2003) as well as robust testing for linear cointegration in the presence of outliers (see Franses and Lucas 1998 or Nielsen 2004). Against this background, our approach to estimate the transaction costs of arbitrage and to test for market integration in the presence of occasional extreme price movements involves the outlier robust estimation of threshold non-linear adjustment models as well as the using these estimates to test for cointegration.

We use Monte Carlo simulation to show that occasional extreme variable movements - interpreted as additive outliers or fat tails - result in large small sample biases of OLS based TAR parameter estimates and, subsequently, in a decreased power of related (threshold) cointegration tests. As a remedy we suggest the use of robust estimators, namely the least-absolute-deviation (LAD) estimator and the maximum likelihood estimator based on the Student t-distribution (t-ML). We demonstrate that robust estimators lead to considerably less small sample bias, particularly for the threshold parameter representing transaction costs, and to more power of (threshold) cointegration tests. When outliers or fat tails are present, the robust cointegration tests also often outperform the Augmented Dickey Fuller (ADF) test at small sample sizes. Finally, we exemplify this phenomenon

[1] General treatments on robust estimation are Huber (1981) and Hampel et al. (1986).

by estimating the transaction costs of arbitrage between the US and the UK natural gas market and by testing for market integration.

In Section 4.2 we review the LOP, the corresponding TAR model as well as cointegration testing for market integration and propose two robust TAR estimators. In Section 4.3 we use Monte Carlo simulation to study the small sample biases of robust TAR estimators and the power of robust cointegration tests. In Section 4.4 we apply robust estimators and tests using US and UK natural gas prices. We conclude in Section 4.5.

4.2 Methodology

4.2.1 Testing for Spatial Arbitrage

The LOP in the presence of transaction cost τ states that arbitrage is profitable if one of the following conditions holds:

$$\begin{aligned} P_{it} &> P_{jt} + \tau \\ P_{it} &< P_{jt} - \tau. \end{aligned} \tag{4.1}$$

where P_{it} and P_{jt} denote the prices of a tradable, homogenous good in markets i and j, respectively, at time $t = 1, \ldots, T$.

The price spread $s_t = P_{it} - P_{jt}$ should revert towards zero whenever arbitrage is profitable. This type of adjustment towards a long run equilibrium corresponds to the notion that the prices are cointegrated with a cointegration vector of $(1, -1)$. Hence, in the absence of transaction costs, i.e., when $\tau = 0$, the LOP can be tested with a unit root test applied to s_t. This assumes a linear adjustment (or error correction) to the long run equilibrium irrespective of the size of the price spread. When transaction costs are present, i.e., $\tau > 0$ the LOP implies that the error correction towards the long run equilibrium stops as soon as $|s_t| < \tau$. Hence, the adjustment of the price spread follows a threshold non-linear process and a unit root based cointegration test should be modified to capture this type of or threshold cointegration or, more precisely, threshold error correction.[2]

[2] The term threshold cointegration was introduced by Balke and Fomby (1997) and, since then, has often been used to describe the phenomenon of non-linear adjustment to the long run equilibrium. However, it should be noted that the term threshold error correction is a more precise description as the threshold non-linearity is in the adjustment process towards the long run equilibrium, but not in the cointegration vector itself. For an investigation of threshold effects in the cointegration relation see Gonzalo and Pitarakis (2006). In the remainder of the text we follow Balke and Fomby (1997) and mean threshold error correction when we use the term threshold cointegration.

Following Balke and Fomby (1997) we use a BAND- and an EQ-TAR model for the price spread that allows for this type of threshold non-linear adjustment. In a BAND-TAR model the price spread reverts towards the transaction cost or threshold band given by τ and $-\tau$. However, this reversion only happens if the price spread (in case it has a positive value) is larger than the threshold value τ or if the price spread (in case it has a negative value) is smaller than $-\tau$. In contrast, in an EQ-TAR model the price spread reverts towards zero if the absolute value of the price spread is larger than the threshold value representing transaction costs.

Economically, the BAND- and EQ-TAR mainly differ in the their assumption about the arbitrage process. The BAND-TAR reflects a situation where economic agents can trade arbitrary small quantities. This enables them to maximize their profits from arbitrage by only importing a quantity of the good that is just enough to make the price spread equal to transaction costs. In the EQ-TAR economic agents may not be able to trade less than a certain minimum quantity, e.g. at least one container ship.[3] This might lead to a reduction of the price spread to less then the transaction costs, which in turn could be reflected in price spreads that revert not towards τ or $-\tau$, but towards points closer to the point of price equality. A EQ-TAR assumes that this point of attraction is a price spread of zero. Hence, a BAND-TAR reflects a world where very small quantities can be traded, whereas a EQ-TAR reflects a world where only quantities of a certain miminum size can be traded.

Formally, a BAND-TAR model is given by:

$$\Delta s_t = \phi(s_{t-1} - \tau) I_{\{s_{t-1} > \tau\}} + \phi(s_{t-1} + \tau) I_{\{s_{t-1} < -\tau\}} + \sum_{l=1}^{L} \alpha_i \Delta s_{t-i} + \varepsilon_t \tag{4.2}$$

An EQ-TAR model is given by:

$$\Delta s_t = \phi(s_{t-1}) I_{\{|s_{t-1}| > \tau\}} + \sum_{k=1}^{K} \alpha_i \Delta s_{t-i} + \varepsilon_t \tag{4.3}$$

where $\phi = \rho - 1$ (with $0 < \rho < 1$), $I_{\{.\}}$ denotes an indicator function and ε_t has mean zero and is weakly stationary. The standard way to estimate equation (4.2) and (4.3) was developed by Chan (1993). Accordingly, the model is estimated with OLS repeatedly for each unique value of s_{t-1} except the $(100 \times p)\%$ smallest and largest values of s_{t-1}. In the literature it is typically suggested to set the so called trimming parameter p equal to 0.05, 0.10 or 0.15.[4] The value of τ that minimizes

[3] This restriction can lead to lower profits and welfare gains from arbitrage than in a BAND-TAR world, when arbitrage forces the price spread over and beyond the transaction cost threshold.

[4] We use $p = 0.05$ throughout the paper in all estimations.

the sum of squared residuals (SSR) yields a rate T consistent estimator for the threshold parameter and thereby for the transaction costs of arbitrage.

Following Enders and Granger (1998), we test for threshold cointegration, i.e., market integration by computing the F-statistic for the null hypothesis of a unit root $\phi = 0$ against the alternative hypothesis of the unrestricted BAND-TAR (4.2) or EQ-TAR model (4.3), respectively. As τ is not identified under the null the testing problem is non-standard. For this reason, critical values are obtained using a residual-based bootstrap. Similar testing problems have been considered in Enders and Granger (1998), Enders (2001) or Caner and Hansen (2001). For the bootstrap, we draw with replacement from the residuals of model (4.2) or (4.3) and use these residuals as innovations to simulate the process under the null hypothesis of a unit root. The test statistic is computed for each bootstrap sample and the procedure is repeated 1000 times.

Because OLS minimizes the sum of squared residuals extreme values generally have a large effect on the OLS based estimates of TAR parameters. As we demonstrate in Section 5.6 when there are occasional extreme values in the s_t series OLS based estimation procedures suffer from considerable small sample biases for the threshold parameter. As a corollary, this often makes OLS based TAR estimates sensitive to the choice of the so called "trimming parameter" p. Moreover, it decreases the power of cointegration tests that rely on such OLS based TAR estimates. In other words, occasional spikes in the price spread series lead to a lower probability that the alternative hypothesis of cointegration and market integration is accepted in case it is correct.

4.2.2 Robust Estimation of Threshold Autoregressions

To remedy the poor small sample properties of TAR based estimation and cointegration testing in the presence of outliers we propose to estimate equation (4.2) and (4.3) with two robust estimators. The first robust estimator is the least-absolute-deviation (LAD) estimator (see Basset and Koenker 1978 and Pollard 1991 for asymptotic properties, and Koenker and Basset 1982 for hypothesis testing in this framework). Similar to the OLS based estimation procedure the equations are estimated for a grid of τ values that includes each unique value of s_{t-1} except the $(100 \times p)\%$ smallest and largest values of s_{t-1}.[5] The parameter vector $\beta = (\phi, \alpha_1, \ldots, \alpha_p)$ that minimizes the sum of the absolute residual values (SAR) gives the estimate of the parameters of (4.2) or (4.3):

$$\hat{\beta}_{LAD} = \text{argmin} \sum_{t=1}^{T} |\hat{\varepsilon}_t|. \tag{4.4}$$

[5] Alternatively, an equally spaced grid between the smallest and the highest permissible values of s_{t-1} (i.e. ignoring the $(100 \times p)\%$ smallest and largest values of s_{t-1}) can be used. This option is sometimes computationally more efficient than the standard procedure.

Alternatively, we suggest using the maximum likelihood (t-ML) estimator assuming ε_t follows a t-distribution with degrees-of-freedom ν (see e.g. Lucas 1997).

$$(\hat{\beta}_{t-ML}, \hat{\sigma}) = \text{argmax} \sum_{t=1}^{T} \ln \left(\frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sigma\Gamma\left(\frac{\nu}{2}\right)\sqrt{\nu\pi}} \left(1 + \frac{\hat{\varepsilon}_t^2}{\nu\sigma^2}\right)^{-\frac{1}{2}(\nu+1)} \right). \tag{4.5}$$

Although one could in principle estimate ν along with the other parameters we suggest estimating the model with one or several fixed values. Following Lucas (1995b), throughout this investigation we estimate the model with a fixed degrees-of-freedom parameter $\nu = 3$. As above for the OLS and LAD estimator a grid search is used to choose the parameter values that correspond to the highest log-likelihood and to obtain the estimate of the other TAR parameters. Both of the above estimators fall into the class of M-estimators in robust statistics (see Hampel et al. 1986) and alternatives such as the Huber estimator could be considered.[6] The test of the null hypotheses of a unit root is tested similarly as in the OLS case. However, instead of an F-statistic a likelihood ratio statistic is used for the test.

4.3 Monte Carlo Study

4.3.1 Small Sample Biases

In this section we present Monte Carlo simulation results concerning the small sample properties of OLS and robust estimators of the threshold parameter τ (transaction costs) as well as the power of threshold cointegration tests (market integration) in the presence of fat tails and outliers.

To study small sample biases of different estimators we simulate the TAR models given in the equations (4.2) and (4.3), respectively. For sample sizes T = 100, 125, ..., 200, 250, 300, 500 we generate ε_t sequences from a set of T uncorrelated pseudo-random draws from three different distributions: (a) the standard normal distribution, (b) a t-distribution with three degrees of freedom and (c) a normal distribution with additive positive outliers. More precisely, the ε_t sequence in the "additive outliers" case (c) is obtained by pseudo-random draws from a standard normal distribution and by adding the number 5 or -5 to a draw with a 2,5 % probability each.

For each error series (a), (b) and (c) and after randomizing the initial value s_1 the next $T-1$ values of s_t were generated using the equations (4.2) and (4.3) with $\rho = 0.9$, $\tau = 2$ as well as $L = 0$ or $K = 0$, respectively, as a base specification. Further, four alternative specifications were considered to study how the results change across the parameter space: $\tau = 2$ and $\rho = 0.85$, $\tau = 2$ and 0.95,

[6] In fact, the LAD estimator corresponds to the Huber estimator when the tuning constant is chosen to be equal to zero.

$\rho = 0.90$ and $\tau = 1$ as well as $\rho = 0.90$ and $\tau = 3$. For each parametrization and TAR-model (BAND-TAR and EQ-TAR) the three series $s_{t,(a)}$, $s_{t,(b)}$ and $s_{t,(c)}$ resulted. Using these series we estimated the TAR model parameters employing the OLS based procedure outlined in Section 4.2.1 as well as the t-ML and the LAD estimators proposed in Section 4.2.2. This procedure was repeated $M = 10000$ times for each TAR-model, parameter specification and sample size T. The average of the threshold estimates over these 10000 trials was recorded and the true threshold value $\tau = 2$ was subtracted to obtain the small sample bias. The left panels of Figure 4.1 (BAND-TAR) and Figure 4.2 (EQ-TAR) show the small sample biases at different T for each of the three estimators used on each of the three error distributions for the base specification $\tau = 2$ and $\rho = 0.90$. For conciseness we present the results for the alternative specifications of the BAND-TAR and the EQ-TAR in the left panels of the Figures shown in Appendix B in Section B.1 and Section B.2, respectively.

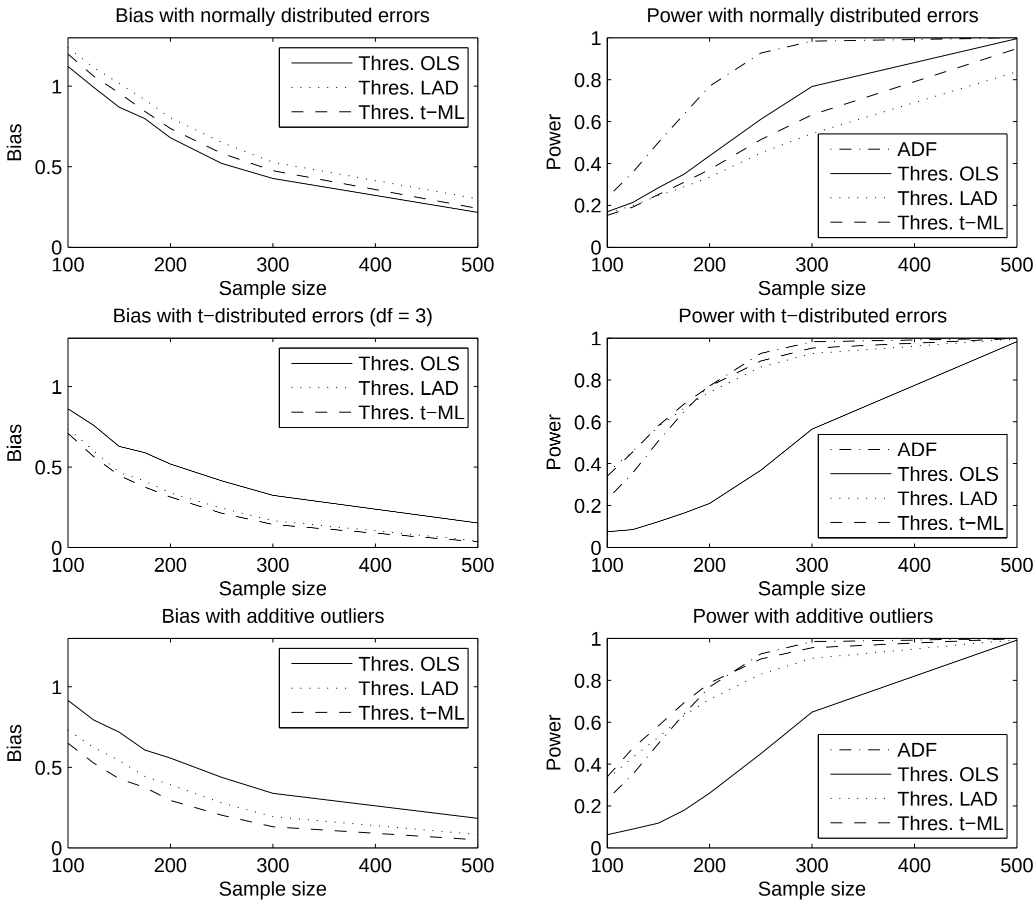


Figure 4.1: BAND-TAR: Bias in threshold estimates (left) and power of cointegration tests (right) in the base specification

In all settings the three TAR estimators of the threshold parameter appear asymptotically unbiased, but exhibit sizable positive biases in small samples. In the case of normally distributed errors the bias of the OLS estimator for both the BAND- and the EQ-TAR model is slightly smaller than the

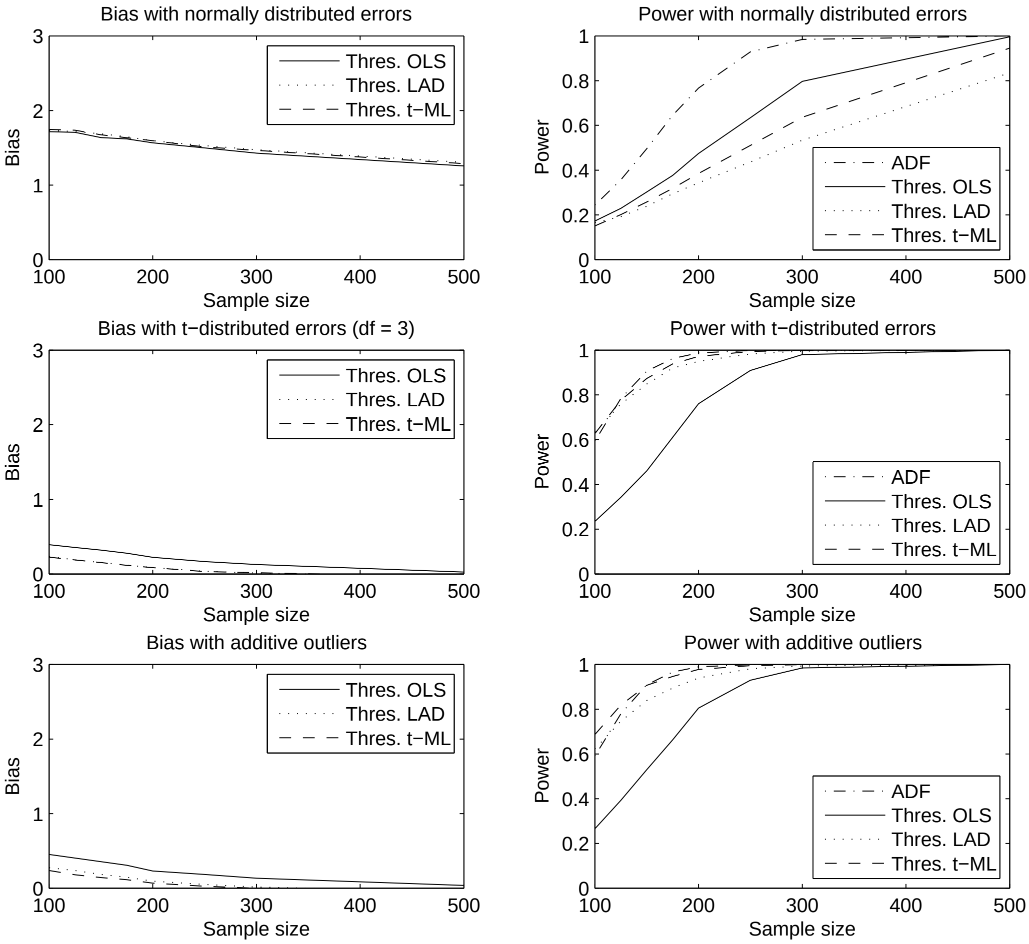


Figure 4.2: EQ-TAR: Bias in threshold estimates (left) and power of cointegration tests (right) in the base specification

bias of the robust estimators. In the case of t-distributed and additive-outlier-contaminated errors, however, OLS performs worse than the robust estimators. Further, the t-ML estimator always has the lowest bias with the non-normal error distributions. Interestingly, the bias for the case of additive outliers and t-distributed errors is generally smaller than for the case of a normal error distribution. By inspection of the Figures shown in Appendix B Section B.1 and B.2 we further find that this phenomenon becomes more distinct with higher values of τ and seems to be more pronounced for the EQ-TAR. Further, the small sample bias generally increases the closer the persistence parameter ρ is to one and the closer the true threshold τ is to zero. In conclusion, these large small sample biases in $\hat{\tau}$ are concerning as they imply that given the usual sample sizes empirical results on the threshold parameter, in general, and TAR based transaction cost estimates, in particular, should be interpreted with great care. For diligent applied work we recommend to carefully discuss the economic reasons for using a BAND-TAR or an EQ-TAR or to test which model is more appropriate. Moreover, we recommend the inspection of the residuals and to test the properties of the error distribution.

4.3.2 Critical Values and Power of Cointegration Tests

In this Section we study the power of robust threshold cointegration tests as well as the power of OLS based (threshold) cointegration tests when the errors are drawn from additive outlier or t-distributions. To obtain critical values and the size-adjusted power for OLS based and robust threshold cointegration tests we performed a similar Monte Carlo experiment as in Section (4.3.1). More specifically, to obtain the distribution of the threshold cointegration test statistic we simulate the BAND-TAR (4.2) and the EQ-TAR (4.3) under the null hypothesis of a random walk $\rho = 1$ (or $\phi = 0$) for each of the three error distributions (a), (b) and (c).[7] For each simulated series we calculate the F-statistic or the likelihood ratio statistic for the restriction $\rho = 1$. This procedure is repeated $M = 10000$ times for each of the three error distributions and both the BAND-TAR and the EQ-TAR model. The 95% percentile of the corresponding distributions of the F-statistic or the likelihood ratio statistic, respectively, is used as the 5% critical value for the threshold cointegration test.

As the next step, the critical values are used to obtain the statistical power of cointegration tests for the BAND-TAR and the EQ-TAR specification based on the three estimators for each of the three error distributions (a), (b) and (c). We use the simulated s_t series corresponding to the base specification ($\tau = 2$ and $\rho = 0.90$) from Section 4.3.1 as the outcomes under the alternative hypothesis of the threshold cointegration tests i.e. $s_{t,(i)} = s_{t,(i),H_A}$, for $i = a, b, c$. Using the simulated critical values each series was tested for a unit-root using the OLS based threshold cointegration test and the robust threshold cointegration tests. For each error distribution (a), (b) and (c) and for each of the four cointegration tests the process was repeated $M = 10000$ times. The percentage of instances in which each test correctly rejected the null hypothesis of a unit-root for the base specification is depicted in the right panel of Figure 4.1 for the BAND-TAR and Figure 4.2 for the EQ-TAR for growing sample sizes T =100, 125, ..., 200, 250, 300, 500 and a test size of 5%.

For comparison, Figure 4.1 and Figure 4.2 also show the power of a standard Augmented Dickey Fuller (ADF) test using the same series $s_{t,(i),H_A}$ as for the threshold cointegration tests. Enders and Granger (1998) and Enders (2001) demonstrate in Monte Carlo Simulations that in the case of normally distributed errors the ADF test performs equally well and sometimes even better than the OLS based threshold cointegration test.[8] Hence, a comparison of the robust tests and the ADF based based cointegration test in the case of outliers and fat-tailed error distributions seems worthwhile. Additionally, we also use the simulated series of the alternative specifications from Section 4.3.1

[7] As before $T = 100$, 125, ..., 200, 250, 300, 500 as well as $L = 0$ and $K = 0$, respectively.

[8] The authors suggest that although the ADF test is misspecified it performs so well because the ADF test avoids the estimation of the additional parameter τ.

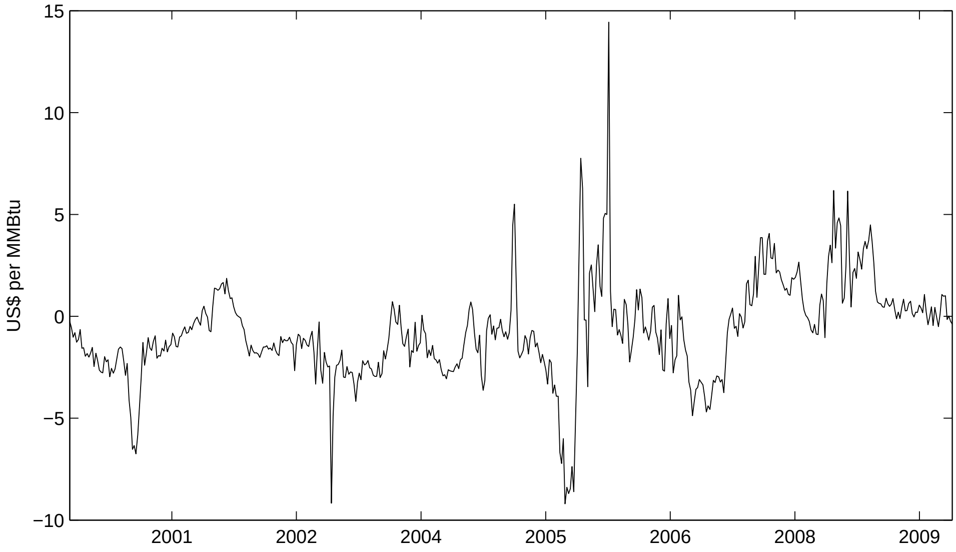


Figure 4.3: Price spreads of US and UK gas markets

for power simulations. The right panels of the Figures shown in Appendix B Section B.1 and B.2, respectively, present the results.

In accordance with the literature on unit root tests, the plots show that at small sample sizes the test power is low when the persistence parameter ρ is close to one. Further, a higher value of τ also generally leads to a loss of power. In line with Enders and Granger (1998) and Enders (2001), the ADF test has always more power than both the OLS based and robust threshold cointegration tests when the errors are normally distributed. However, for additive outliers and t-distributed errors the LAD and t-ML based tests outperform the ADF and the OLS based cointegration test at small sample sizes. However, the test based on the LAD estimator has almost always slightly less power than the t-ML based test. At larger sample sizes the power of the ADF test becomes similar to the robust tests.[9] Generally, the sample size at which the power difference between the ADF and the robust tests disappears is larger for the BAND-TAR and grows with τ and ρ. The OLS based threshold cointegration test tends to have very low power even for larger sample sizes when the errors are not normally distributed. Finally, the test power for the EQ-TAR specifications tends to be higher than when cointegration is tested with a BAND-TAR model.

4.4 Application to US and UK Natural Gas Prices

In this section, we illustrate the importance of robust TAR methods for estimating the transaction costs of arbitrage between the US and UK markets for natural gas which recurrently exhibit price

[9] An exception is the EQ-TAR with a low $\tau = 1$. Here the ADF test has a higher power than the other tests even at small sample sizes.

spikes. Further, we test for market integration using the OLS based and the robust threshold cointegration tests. We use weekly natural gas prices at the UK National Balancing Point (NBP) and the US Henry Hub (HH) for the period from April 2000 to December 2009 (509 observations). Figure 4.3 shows the price spread $s_t = p_{NBP,t} - p_{HH,t}$ and illustrates our notion of occasional extreme price movements. Multiple positive and negative spikes in the price spread series can be observed. Some spikes in the price spread can be attributed to external shocks such as the US hurricane seasons, but for other spikes it is hard to unequivocally relate them to singular events. Generally, after shocks to the price spread strong adjustments towards zero can be observed which in part may be caused by arbitrage activity. Econometrically, the spikes in the price spread can be interpreted as resulting from a fat tailed error distribution or due to the presence of outliers. This likely leads to a large small sample bias in any transaction cost estimates that rely on the OLS based TAR estimator and to a low power of related test for market (co-)integration.

We argue that a BAND-TAR specification is adequate to model arbitrage between the Atlantic gas markets. The minimum quantities transported in natural gas tankers are small compared to the market size of the HH or NBP market. They are certainly not large enough to move the price spread to a point within the transaction cost band. Therefore, we use the NBP-HH gas price spreads to estimate Equation (4.2) with OLS and with the robust TAR estimators. The Bayesian information criterion selects a lag length of $L = 2$. First of all, one can expect that the value of the threshold is overestimated for the OLS based TAR estimator, because extreme observations have a large effect on the sum of squared residuals and a large threshold may mitigate that effect. Second, as a consequence of an overestimated threshold the persistence parameter may be underestimated. In order to discuss how the small sample bias of the estimators compare we plot the potential values of τ against the value of the estimators objective function i.e. SSR, SAR or LL, respectively (see Figure 4.4). The OLS based SSR decreases in the threshold value. Thus, the optimal threshold will always be on the boundary of the permissible value range and the threshold estimate is dependent on the trimming parameter p. This is clearly not a desirable feature. Because of this sensitivity in OLS based estimates researchers may even conclude that the TAR model is inappropriate and dismiss it entirely.

As the estimate of ϕ and, thus, the cointegration test statistic is dependent on the threshold estimate $\hat{\tau}$ this may also lead to a decreased power of OLS based threshold cointegration tests. In contrast, the robust estimators have clear inner minima for the SAR and LL values.

Table 4.1 shows the estimation results using the three estimators and the p-values for the cointegration tests. The smaller threshold and adjustment parameter estimates of the robust estimators confirm the graphical analysis. The robust threshold estimates for the transaction costs of arbitrage

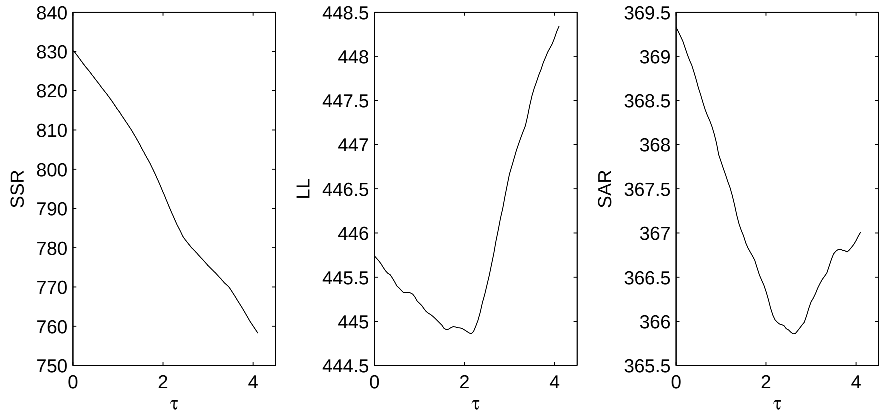


Figure 4.4: Potential threshold values τ and the value of the objective function for the OLS, t-ML and LAD estimators.

Table 4.1: Parameter estimates and test results

	$\hat{\tau}$	$\hat{\phi}$	$\hat{\alpha}_1$	$\hat{\alpha}_2$	p-val unit root test
OLS	4.106	-0.643	-0.007	-0.085	0
		(0.073)	(0.044)	(0.041)	
Student	2.135	-0.156	0.034	-0.111	0
		(0.034)	(0.029)	(0.029)	
LAD	2.631	-0.232	0.020	-0.083	0
		(0.032)	(0.027)	(0.025)	

between the US and the UK natural gas market are considerably smaller than the OLS based estimates. Furthermore, the estimated adjustment is much slower for the robust estimators. In contrast, the OLS estimator suggests much larger transaction costs and a much faster rate of mean reversion towards the no-arbitrage band. The results of the cointegration test, however, are identical for all three approaches.

4.5 Conclusion

In this essay, we have shown that extreme price movements, either interpreted as outliers or as fat tails lead to large positive small sample biases in the OLS estimation of TAR models. In the context of market integration studies where the threshold estimates are used as a measure of transaction costs, this means that the transaction costs of arbitrage are overestimated when TAR model is estimated with OLS. Moreover, conventional (threshold) cointegration tests that are based on such TAR estimates, e.g. for testing the LOP, can have very low power in the presence of occasional extreme values. Monte Carlo simulations demonstrate that robust methods decrease the small sample biases considerably and increase the power of threshold cointegration tests. An application to natural gas prices further illustrates the differences between OLS based and robust TAR estimation and cointegration testing.

A number of issues need to be addressed in future research. From an empirical perspective, applications as seen in Lo and Zivot (2001) may be reconsidered with our findings in mind. However, great attention should be paid to the properties of the error distribution and to the choice of the appropriate TAR model. From a methodological perspective, more general specifications of the non-linear model should be considered. These could be asymmetric adjustment models, which straightforwardly generalize the model used here. But also ESTAR models for which the distinction between regimes is not abrupt but continuous should be studied. Moreover, as transaction costs between markets change over time models that take account of sharp or smooth structural breaks in the threshold parameter should be investigated. Finally, robust methods should be considered for the estimation of cointegrating relationships when the integration of markets with goods of heterogeneous quality is investigated.

Chapter 5

The Role of Experience Curves for the Shale Revolution

5.1 Introduction

Since less than one decade the shale revolution reshapes the North American crude oil and natural gas market. The term "shale revolution" describes the dramatic increase in US hydrocarbon production from shale formations and other unconventional reservoirs as well as its economic implications.[1] Shale oil and gas production has not only increased as a share of total US oil and gas production, but has also substantially raised total US production of both hydrocarbons. This has led to low US gas prices and to lower US oil imports which subsequently contributed to the substantial, recent decrease in global oil prices. Moreover, in the US public debate large economic benefits are attributed to the shale revolution.[2] The shale revolution is regarded to have positive effects on US economic growth by increasing the supply of energy and decreasing energy prices. Further, positive employment effects are mentioned as well as higher energy security for the US (see IHS (2013) and Medlock et al. (2011)). Multi-billion investments in new liquified natural gas export infrastructure and the current vote in the US Senate energy panel to lift the 40-year-old oil export ban even hold out the prospect of large scale hydrocarbon exports from the US before the year 2020. Finally, recent reductions in US CO_2 emissions are associated with the shale revolution due to a replacement of coal by less CO_2-intensive natural gas in electricity generation.

Despite these implications little research has been done on the drivers of the massive oil and gas production gains of the US shale revolution. "Hydraulic fracturing" and the growing share of

[1] Hydrocarbons are an umbrella term for crude oil and natural gas, but also for other chemical compounds that are composed entirely of hydrogen and carbon.

[2] For example, see http://marcelluscoalition.org/2015/01/20-obama-administration-quotes-about-u-s-shale/ for a comprehensive overview of quotes of the Obama administration on the economic benefits of the shale revolution.

"horizontal drilling" operations are deemed the core technologies behind the sweeping rise of US oil and gas production during the shale revolution. However, the enormous number of US shale drilling projects and the necessity to learn an optimized use of specialized drilling technology in unfamiliar shale rock formations also stress the importance of experience for shale oil and gas extraction. Moreover, statements of US industry representatives like "*shale drilling has much of an art form*" and the acquisition of forerunner "shale firms" by the formerly more reluctant multi-national companies in order to "*catch up on the shale boom*" highlight the importance of experience as a decisive production factor.[3] More specifically, these observations may indicate that "local" drilling experience might be of high relevance for shale oil and gas extraction. In the definition used here, overall experience can be subdivided into local experience and global experience. Global experience is the part of experience that can spill over to other firms, industries or countries. Hence, global experience has the characteristics of a public good. In contrast, local experience is a private good that can be used only by firm within the industry, production region or country under consideration. For the case of the shale industry, this local experience may be incorporated in the experience workers or firms have accumulated with the specific regional shale geology and the corresponding operational designs that are tailored to the characteristics of a certain shale region.

When local experience is a decisive production factor, this might have important implications for the possibility of replicating the successes of the US shale revolution in other countries. Whereas the technology level in other countries can be adjusted to US levels by upgrading the drilling rig fleet, local experience has to be accumulated by a large number of drilling operations. In some countries such as several European countries there is a comparatively low potential to set up such a large number of drilling operations due to limiting factors such as a high population density, public acceptance or stringent regulation. Hence, if local experience is an important production factor for shale oil and gas production, limits to the accumulation of local experience might, ceteris paribus, hamper the growth of shale oil and gas production in such countries.

This paper empirically investigates the role of both drilling technology and local drilling experience for the production of shale oil and gas. To address this, we estimate oil and gas production functions that account for drilling technology and drilling experience. To control for experience spillovers and to identify the effect of local drilling experience, we use the common correlated effects mean group estimator (CCEMG) and the pooled common correlated effects (CCEP) estimator introduced by Pesaran (2006). These estimators account for the presence of common correlated effects, i.e., processes that are common to the cross-sectional units in the panel and that are potentially driving both the explanatory variables and the error term. Common correlated effects capture cross-sectional

[3] It is certainly no problem for multi-national companies to buy and use drilling rigs that are capable of horizontal drilling and hydraulic fracturing. However, how productive these drilling rigs are may depend on experience gathered with these technologies in a certain region.

spillovers and other forms of cross-sectional dependency. Hence, these estimators control for experience spillovers, i.e., for global experience and allow the identification of the role of local drilling experience in the production of shale oil and gas.

We find robust evidence for strong local experience effects in shale oil and gas production. Further, by disaggregating drilling and experience into different technology classes we can confirm that horizontal drilling contributes most to the production of shale oil and gas. Moreover, the impact of experience gathered by using the horizontal drilling technology has the largest impact on shale oil and gas production compared with other production factors.

The remainder of the paper is organized as follows: Section 2 discusses the role of technology and related experience for the US shale revolution. In Section 3 a theoretical and empirical model is presented. In Section 4 the estimation strategy is outlined. Section 5 describes the data, Section 6 presents the estimation results and Section 7 concludes.

5.2 The Shale Revolution and the Role of Experience Curves

5.2.1 The Shale Revolution as a Result of Growing per Rig Production

The shale revolution is characterized by a steep increase in the crude oil and natural gas production from formerly untapped geological shale formations in several US regions.[4] The increase in production was so strong that shale oil or gas extraction increased from around 10 percent of total US oil and gas production in 2007 to almost 50 percent in 2014. Correspondingly, in the same period total US oil and gas production increased by 70 percent and more than 30 percent, respectively. See Figure 5.1 for plots of total US oil and gas production as well as shale oil and gas production.

The production increase from shale sources was not only caused by more oil and gas drilling activity, but also by dramatic increases in production per drilling operation in US shale regions. Figure 5.2 shows the initial oil and gas production per drilling rig for the seven major US shale regions in the period 2007 to 2015.[5] In recent years the initial oil production per shale drilling rig has increased strongly, e.g., from below 50 barrels to almost 700 barrels per day in the Eagleford region. Similarly, initial gas extraction per drilling rig has increased from around half a million cubic feet to more

[4] The term "shale revolution" seemingly implies that the growth in hydrocarbon production is exclusively from shale rock. The type of rock that is the source of production growth, however, should, more precisely referred to as low-permeability sedimentary rock. Besides shale, this type of rock also includes sandstone and limestone. Shale is, however, the most important of these rocks with respect to hydrocarbon extraction. Hence, the name shale revolution.

[5] Ideally, total life cycle production of each drilling operation should be used to illustrate the production increase per drilling operation. However, such data are rarely available. Therefore, we follow the US Energy Information Administration (EIA) and uses initial oil and gas production per drilling operation to approximate production per drilling operation.

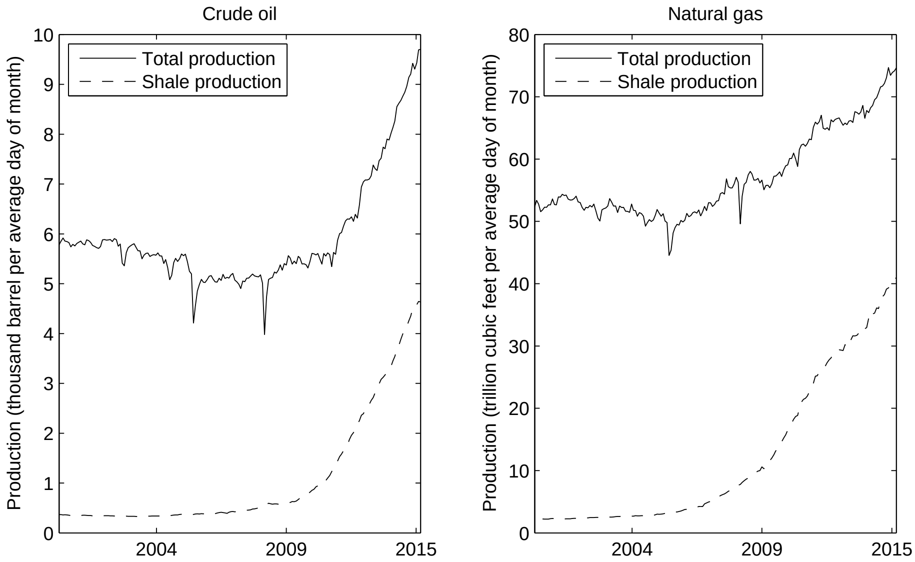


Figure 5.1: US total and shale production of crude oil and natural gas. Source: US Energy Information Administration

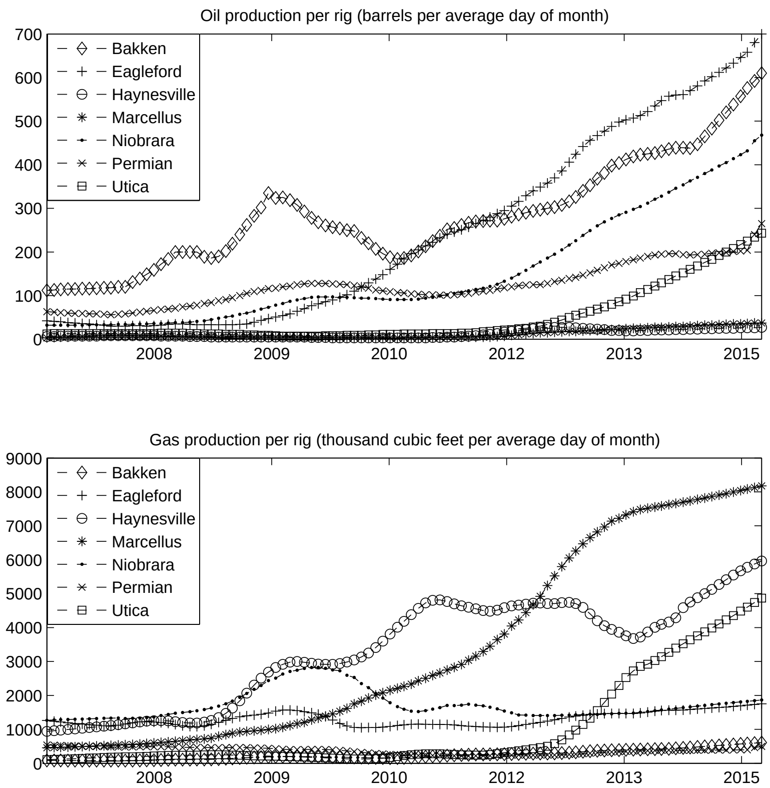


Figure 5.2: Average initial production per drilling rig. Source: US Energy Information Administration

than 8 million cubic feet per day in the Marcellus region. The graph also shows a significant timely and regional variation of the most substantial gains. Moreover, there are regions that have a high per rig output only in terms of oil such as the Bakken region or only in terms of gas such as the Marcellus region. But there also regions such as Utica or Niobrara which have a more balanced mix of oil and gas output per drilling operation.

In summary, Figure 5.2 indicates that the mere quantity of drilling operations cannot explain the increase in shale oil and gas production. The next section explains the role of drilling technology and drilling experience as potential drivers of this growth in the production (per drilling operation).

5.2.2 Drilling Productivity and the Role of Technology and Experience

Both the technology, i.e., the use of drilling rigs that are capable of hydraulic fracturing and horizontal drilling as wells as the accumulated experience on how to optimally use these technologies in unfamiliar shale rock are deemed to be the predominant drivers behind the large gains in shale oil and gas production.[6] In the following, technological aspects of hydraulic fracturing and horizontal drilling, their implications for the organization of drilling projects as well as the associated opportunities to collect production enhancing local and global experience are discussed.

Hydraulic fracturing means to crack fissures into the target rock formation at multiple locations along a well borehole - the so called fracturing stages. The fissures are then widened by inserting the so called fracturing fluids consisting of water, sand and a mix of chemicals. This treatment induces the flow of oil and natural gas out of formerly isolated pockets into the borehole. Due to the low permeability of shale rock hydraulic fracturing is a prerequisite for extracting hydrocarbons from shale rock. Hydraulic fracturing offers a large range of opportunities to gather production enhancing experience. There is a multitude of different procedures and potential technical specifications for cracking the rock. One of the most important choices with respect to the hydrocarbon output of a borehole is the number and distance of the fracturing stages along that borehole. The composition of the fracturing fluid also strongly influences the hydrocarbon output of the well. Due to far reaching fissures and the inserted fracturing fluids, the choices made for one borehole can have effects on the output of nearby boreholes and wells. Most of the corresponding choices have to be optimized with respect to the specific geology of a certain region. Hence, whereas a certain part of the hydraulic fracturing experience can be used in other shale regions, a substantial part of the gathered experience can only be used locally.

[6] Actually, both hydraulic fracturing and horizontal drilling were developed and use decades before the shale revolution. However, the combined use of both technologies in geologically unfamiliar shale formations was novel for the industry.

Horizontal drilling means to begin a borehole vertically, but to deviate horizontally when the drill head approaches target depth. Horizontal drilling was developed as an extreme case of directional drilling, a category that includes any drilling that deviates from a vertical trajectory.[7] Horizontal drilling facilitates the extraction of hydrocarbons from low-permeability formations that are narrow and run horizontally through the ground such as shale rock. Similar to hydraulic fracturing, horizontal drilling involves a large range of choices to be optimized in order to maximize production per borehole and well. Important choices include the trajectory of a borehole, the distance between neighboring boreholes as well as the timing of subsequent, nearby boreholes. Again a part of the gathered experience spills over to other regions, but large part of the experience is specific to the geology of a certain shale region.

The geological properties of shale also entail challenges and opportunities for well planing and field development strategies that oil and gas market stakeholders often summarize as "industrial drilling".[8] It turned out that when hydraulic fracturing and horizontal drilling are combined well output could be further increased by drilling multiple boreholes from a single well pad into a confined underground target space.[9] The quantity, trajectories and other specifications of these boreholes have an impact on the production of the respective well, but also on nearby wells. This necessitates the development of new well design and field development strategies which in turn relies on collected experience. Moreover, local experience is of major importance, because the optimal borehole specification and field development strategy is specific to the regional geology.

In summary, hydraulic fracturing and horizontal drilling are certainly the core technologies behind the high production growth rates during the shale revolution. However, the vast set of new technical and organizational choices that are to be coordinated in shale drilling stress the importance of learning and experience.

5.3 Theoretical and Empirical Model

5.3.1 Empirical Model

To investigate the role of experience for the production of shale oil and gas we use a Cobb-Douglas production function.

$$Y_{rkt} = AE_{rt}^{\gamma} D_{rt}^{\beta} \tag{5.1}$$

[7] More precisely, horizontal drilling is defined as directional drilling with a wellbore that exceeds a bend of 80 degrees. However, due to its expanding use in the production of shale oil and gas most data publishers, e.g., Baker Hughes count horizontal drilling as a separate category - next to directional and vertical drilling.

[8] The term "industrial drilling" has become widespread use among market stakeholder. However, there is no clear definition of the term as different stakeholders include different aspects. Hence, we only present a choice of aspects that is relevant for the relationship of output growth (per drilling operation), technology and experience.

[9] Shale boreholes can be drilled in much closer proximity without a "cannibalization" of production rates than boreholes in conventional fields due to the low permeability of shale rock. This forces drilling operators to re-adjust their solutions to such proximity problems to the properties of shale geology.

The dependent variable Y_{rkt} is the sum of initial output of $k = oil$ or $k = gas$ produced by all wells in region r that are drilled by a number of drilling rigs D_{rt}.[10] The output elasticity of the drilling rigs is given by β. Note that D is measured by the quantity of active drilling rigs and therefore includes both capital and labor input. However, because rig crews work seven days a week and 24 hours per day during a drilling operation, the number of workers is very closely tied to the number of drilling rigs. Correspondingly, given a certain number of drilling rigs in a region r well output cannot be increased by adding more workers. Hence, in the drilling industry capital and labor are complements rather than substitutes.

Production experience E_{rt} is incorporated in 5.3.1 as a part of total factor productivity $\hat{A}$ which is given by $\hat{A} = AE_{rt}^{\gamma}$ with γ being the learning elasticity and A representing residual total factor productivity. This functional form assumes decreasing marginal returns to experience and is in line with time-honored "power rule" learning curve model proposed by Wright (1936) and introduced to economics by Arrow (1962).[11] Accordingly, equation (5.1) describes a case of so called "unbounded" learning i.e. production does not stop to increase at a certain level of experience. The assumption of unbounded learning is sometimes too strong for mature markets. However, it is adequate in the case of shale drilling where the accumulation of productivity enhancing drilling experience is still very likely to take place as new fields are developed and the optimization of shale specific field development strategies, well design and drilling specifications is far from being finished.

The experience variable is constructed as the cumulative sum of all earlier drilling operations

$$E_{rt} = \sum_{\tau=1}^{t-1} D_{\tau} i^{(t-1)-\tau} \tag{5.2}$$

where i is the depreciation factor of experience. In the base specification $i = 0.99$, i.e., experience depreciates one percent per period.[12] Only cumulative drilling activity up to $t-1$ is included in the experience variable. Hence, we assume that useful experience for the drilling projects measured in D_t can only emerge from drilling projects that were completed earlier than period t. In equation (5.2) we use accumulated drilling activity, i.e., accumulated input as a measure of experience because in oil and gas drilling most production enhancing experience can be collected during the well drilling phase and not during the actual production phase. During the drilling phase most manpower is used

[10] More specifically, initial well production is the production of wells in their first full production month. Notice, that initial well production is used as the dependent variable instead of total life cycle well production. As outlined above no comprehensive data measuring total life cycle well production are available. However, after a well has been drilled and completed well production rates are determined by the loss in geological pressure and cannot be significantly changed by human interference. Therefore, initial new well production rates are closely related with total life cycle well production rates and can be considered as a useful dependent variable with respect to choices made by drilling companies. This view is in accordance with the EIA who use the growth in initial well production per rig to discuss productivity development during the shale revolution in their drilling productivity report. See Section 5.5 for details on the EIA drilling productivity report.

[11] See Berndt (1991) for a derivation of the power rule experience curve model from a Cobb-Douglas production function where total factor productivity is constructed as shown above.

[12] Since, we use monthly data for the estimation further below, a one percent depreciation rate seems appropriate.

as well as most technical and organizational choices have to be made. Hence, most of the experience is collected before a well starts to produce oil and gas. After a well is completed oil and gas flow mostly without human interference and at a rate that is almost exclusively determined by choices in the drilling and completion phase. Accordingly, there is very little opportunity for learning after well completion. Therefore, using the cumulative input variable drilling activity as the experience proxy is much more reasonable than using output variables such as cumulative oil and gas production.[13] Finally, notice that both D and E lack the subscript for the hydrocarbon type k because many shale wells produce both oil and gas.

By taking the logarithms of equation (5.1) a linear empirical model is obtained. This linearized version is augmented by the equations (5.4) and (5.5) which state assumptions about the process that is driving the explanatory variables and the error term, respectively. This setup will serve as the most general specification to be accounted for in the estimation.

$$y_{rkt} = \alpha_{rk} + \beta_{rk} d_{rt} + \gamma_{rk} e_{rt} + u_{rkt} \tag{5.3}$$

with

$$d_{rt} = \boldsymbol{f}_t \boldsymbol{\delta}_r + \boldsymbol{g}_t \boldsymbol{\pi}_r + \eta_{rt} \tag{5.4}$$

$$u_{rkt} = \boldsymbol{f}_t \boldsymbol{\gamma}_{rk} + \epsilon_{rkt} \tag{5.5}$$

where lowercase letters in equation (5.3) denote logarithms of the inputs in equation (5.1). More generally than in equation (5.1) the output elasticities β and γ as well as the time constant part of residual total factor productivity α_{rk} are allowed to vary across the cross-sectional groups rk. Allowing for parameter heterogeneity is appropriate given the regional variation of output per rig time series as shown in Figure 5.2 above. The equations (5.4) and (5.5) supplement the production function (5.3) with assumptions about the processes that drive the explanatory variables and the error term. In particular, we assume that the error term as well as the regressors potentially contain a finite number of unobserved common processes $\boldsymbol{f}_t$ and $\boldsymbol{g}_t$ (also known as "common correlated effects" or "(common) factors", whose impact may vary across cross-sectional groups.[14] The common factors represent unobserved, processes that appear in a subset or all of the cross-sectional groups and therefore provide a quite general framework to capture cross-sectional dependencies. As experience spillovers can be regarded as a type of cross-sectional dependency, estimation methods that take account of the common factors also control for experience spillovers and enable the identification of local experience effects γ_{rk}. Moreover, to account for spillovers that come into effect only after

[13] Using cumulative inputs as a measure of experience is in line with Arrow (1962) who uses cumulative capital investment as a measure for experience. But, it differs from a large part of the experience curve literature which mostly uses accumulated output as a proxy for experience. See e.g. Liebermann (1984).

[14] Whereas the $\boldsymbol{g}_t$-type common factors only enter the process that is driving the explanatory variables, the $\boldsymbol{f}_t$-type is contained in both the error term and the explanatory variables and leads to endogeneity problems if not accounted for.

a longer time period the common factors are allowed to be related over time by stationary or non-stationary processes.[15] Apart from spillovers, common factors also account for other types of cross-sectional dependency. Examples are general technological progress, trends in drilling rig costs or movements in oil and gas prices that drive drilling activity. Another example are singular events such as the US hurricane seasons or the financial crisis which resulted in the temporary shut down of drilling operations.

The fixed effects α_{rk} in equation (5.3) account for cross-sectional heterogeneity in residual total factor productivity. Hence, the α_{rk} mainly accounts for geological and other geographic differences between regions and between oil and gas extraction that affect initial well production. Finally, the explanatory variables and the error term are assumed to contain the stochastic, idiosyncratic shocks η_{rt} and ϵ_{rkt}, respectively.

5.3.2 The Role of Drilling Technology and Technology Specific Experience

In former decades, drilling was predominantly performed by the use of vertical drilling technology. As outlined in Section 5.2 during the shale revolution drilling technology gradually shifted to directional and eventually horizontal drilling technology. To investigate how much certain drilling technologies contribute to production and how experience effects differ for these technologies, we divide the drilling rigs and the experience variable into the technological subcategories shown in Table 5.1.

Table 5.1: Technology categories

m	Technology
1	horizontal
2	vertical
3	directional

We disaggregated the explanatory variables of equation (5.3) according to Table 5.1 and obtain the modified equation

$$y_{rkt} = \alpha_{rk} + \sum_{m=1}^{3} \beta_{mrk} d_{mrt} + \sum_{m=1}^{3} \gamma_{mrk} e_{mrt} + u_{rkt} \tag{5.6}$$

This equation will be used to estimate the contribution of different drilling technologies β_{mrk} and technology specific experience effects γ_{mrk} to the production of shale oil and gas.

Deeper wells tend to have higher pressure and therefore tend to have higher initial production rates. We account for this, by disaggregating drilling activity not only by drilling technology, but also by target depth. Accordingly, we split the drilling variable into the four depth categories shown in

Table 5.2: Target depth categories

h	Target depth (feet)
1	< 5000
2	5000 - 10000
3	10000 - 15000
4	> 15000

Table 5.2.

Hence, using both the split by technology and by target depth equation (5.3) extends to

$$y_{rkt} = \alpha_{rk} + \sum_{h=1}^{4} \sum_{m=1}^{3} \beta_{hmrk} d_{hmrt} + \sum_{m=1}^{3} \gamma_{mrk} e_{mrt} + u_{rkt} \tag{5.7}$$

The disaggregation scheme (5.7) controls for both the variation in technological as well as geological production characteristics. Hence, the estimation of (5.7) will serve as a robustness check for the estimation results of equation (5.3) and (5.6). As an additional sensitivity check, we investigate whether the results are sensitive to the choice of the experience depreciation rate. Hence, we alternatively use a depreciation rate of $i = 0.97$ to construct the experience variable. Furthermore, the Permian region is the only mature region in the sample since it has been developed and exploited for several decades. In addition, in the Permian region there is a considerable number of drilling operations which are not targeted at shale rock, but at conventional reservoirs within the same geographical area. These conventional sources are most often exploited using vertical drilling technology. Hence, to remove the potentially confounding effect of the large number of conventional drilling operations in the Permian region we exclude of the Permian region from the sample as an additional robustness check.

5.4 Estimation Strategy

The main interest of this essay is to investigate the effect of local experience and drilling technology on the production of shale oil and gas. Given our econometric model presented in the equations (5.3), (5.4) and (5.5) several aspects have to be considered for the identification and consistent estimation of the parameters. First, in order to identify the effect of local experience on oil and gas production we need to account for cross-sectional dependencies including experience spillovers. More specifically, we need to take account of the common factors $\boldsymbol{f}$ and $\boldsymbol{g}$ which may follow a (non-) stationary process

[15] If some elements of $\boldsymbol{f}$ follow unit root processes and enter both Equation (5.4) and Equation (5.5) with non-zero coefficients, the variables y, d and $\boldsymbol{f}$ have a common stochastic trend and are cointegrated. More precisely, for cointegration of y, d and $\boldsymbol{f}$ the other elements of u_{rkt} may not follow a unit root process.

over time. Second, the potential presence of the common correlated effects of the $\boldsymbol{f}$-type may lead to the correlation of the error term u_{rkt} and the explanatory variables. If unaccounted for, this would render the parameter estimates inconsistent. Third, we primarily interested in the overall or average effect of the explanatory variables on shale oil and gas production. However, as discussed in Section 5.3 and stated in equation (5.3) there might be regional variation in the model parameters. The false assumption of cross-sectional parameter homogeneity can result in inconsistent estimates when the regressors are not strictly exogenous (see Pesaran and Smith 1995). On the other hand, estimators which correctly assume parameter homogeneity typically provide more efficient estimates by pooling the information of all cross-sectional groups. Hence, it seems advisable to compare the results of estimators which account for parameter heterogeneity and of estimators which assume parameter homogeneity. Fourth, the presence of non-stationary common factors in both the error term and the explanatory variables is equivalent to a common stochastic trend that potentially establishes a long run relationship between the dependent variable and the explanatory variables. Hence, we need to take account of potential cointegration of the variables.

To address these four issues we use the common correlated effects mean group (CCEMG) estimator and the common correlated effects pooled (CCEP) estimator proposed by Pesaran (2006). Both approaches yield consistent estimates of the average parameter values in the presence of the four outlined econometric problems.[16] Most importantly, by accounting for common factors the CCEMG and CCEP estimators are designed to "purge" the explanatory variables from cross-sectional dependency and, thus, experience spillovers. Therefore, the CCEMG and CCEMP estimates of the experience parameter identifies the effect of local experience on shale oil and gas production. Moreover, according to Pesaran (2006) the two estimators yield consistent and asymptotically normal parameter estimates, even if an unknown number of common factors is present. Group specific errors are allowed to be serially correlated and heteroscedastic. Further, the common factors are allowed to exhibit an arbitrary degree of correlation among themselves and they may have a heterogeneous impact across cross-sectional groups - both within the error term as well as a component of the regressors. Moreover, Kapetanios et al. (2011) demonstrate that CCEMG and CCEP estimates are robust to an independent, stationary or non-stationary as well as a nonlinear evolution of the common factors. Accordingly, both estimators accounts for potential common stochastic trends, i.e., cointegration.

To obtain the CCEMG estimates, equation (5.3) is augmented by adding the cross-section averages of the dependent variable and the explanatory variables to the equation. This augmented version of equation (5.3) is estimated separately for each cross-sectional group ik with ordinary least

[16] In other word, the two methods are designed to provide estimates for the cross-sectional average of the cross-section specific coefficients.

squares (OLS). Thereby, the CCEMG estimator implicitly accounts for parameter heterogeneity across cross-sectional groups. Finally, the CCEMG estimates of each parameter are obtained as the cross-section averages of the group specific parameter estimates. To obtain the CCEP estimates, dummies for cross-sectional groups are interacted with the cross-section averages of the dependent and explanatory variables. These interaction terms and dummies for the cross-sectional groups are used to augment equation (5.3). This augmented version of equation (5.3) is estimated with OLS. In contrast to the CCEMG estimator, the CCEP estimator assumes homogenous parameters across groups. Consequently, using both estimators provides enables addressing the question of whether the explanatory variables have an impact on oil and gas production that varies across groups. Standard errors for the CCEMG and the CCEP estimates are constructed following Pesaran and Smith (1995) and Pesaran (2006). For comparison, we also estimate equation (5.3) with the standard fixed effects estimator (FE) and with the first difference estimator with time dummies (FD_t).[17]

5.5 Data

We use monthly data that cover the period January 2011 to March 2015 for the seven major US shale regions. In this period the largest expansion of drilling and production in US shale regions could be observed. Hence, a huge potential for the accumulation of productivity enhancing drilling experience is present in the sample period. For each region, we use time series of drilling activity, crude oil production and natural gas extraction. Oil and gas production data for the seven major US shale regions are published in the drilling productivity report of the US Energy Information Administration (EIA). In particular, the report provides monthly data on total oil and gas production of newly drilled wells for each of the seven regions.

The dependent variable initial production of newly drilled wells (Y_{rkt} in equation 5.3) is measured as the extracted amount of oil or gas in the second and only the second month a well is producing.[18] This is due to the fact that wells are often completed at some point of time within a month (and typically not at the end of a month). In addition, wells often produce at less than their full production rate during the first week of production because fracturing fluids are flowing back into the wellbore and inhibit the flow of oil and gas. Consequently, production data from the first partial-month are not representative of the initial production of a well. To mitigate this problem and to have a measure

[17] The fixed effects estimator assumes parameter homogeneity, controls for group specific effects, but does not control for cross-sectional dependency and a resulting correlation between the error term and the regressors. The first difference estimator with time dummies assumes parameter homogeneity, controls for group specific effects and - through the inclusion of time dummies - at least captures simple forms of cross-sectional dependency.

[18] Consequently, there is actually a lag of two month between the measured activity of drilling rigs and the measured initial oil and gas output that is produced with the newly drilled wells. Accordingly, the explanatory variables are actually predetermined. This probably alleviates, but most likely does not eliminate potential correlation between error term and the explanatory variables.

for new well production the drilling report exclusively uses extraction data from the second month of production. Drilling activity data are obtained from the Baker Hughes North America Rotary Rig Count (BH). The BH data are differentiated by type of drilling technology (horizontal, vertical or directional) and target depth of each drilling operation. The original BH data are measured with a weekly frequency and on the county level. Hence, to obtain monthly data on the level of shale regions as defined in the EIA report the BH drilling data were aggregated to the regional level.[19] In some months and regions there were no rigs actively drilling. For these observations the number zero was replaced by one in order to enable taking logarithms for the equations (5.3), (5.6) and (5.7). For the experience variable it was done in the same way. Table C.1 in Appendix C shows the share of observations where this replacement altered an observation for each of the variables used in the estimation. Replacements only needed to be done when the disaggregated drilling data outlined in Subsection 5.3.2 were used.[20] Table 5.3 shows descriptive statistics for the main variables.

Table 5.3: Descriptive statistics

Variable	Notation	Mean	Std. Dev.	Min	Max
Oil new well production [thousand barrels]	$Y_{oil,rt}$	39,189	44,096	60	166,944
Gas new well production [million cubic feet]	$Y_{gas,rt}$	237,765	207,586	1,260	840,284
Drilling	D_{rt}	161.9	143.1	5	563
Experience	E_{rt}	4388.9	4895.2	6	24237
Drilling: Horizontal	$D_{hor,rt}$	109.4	77.2	0	349
Drilling: Vertical	$D_{ver,rt}$	42.9	95.3	0	359
Drilling: Directional	$D_{dir,rt}$	9.4	10.4	0	46
Experience: Horizontal	$E_{hor,rt}$	2807.9	2583.0	0	10156
Experience: Vertical	$E_{ver,rt}$	1312.4	3171.3	4	13757
Experience: Directional	$E_{dir,rt}$	266.3	322.0	1	1204

The table presents descriptive statistics for each month and region. The production data represent new well production per average day of a month in a given region as published in the EIA drilling productivity report. The descriptives for the technology and depth split shown in Table 5.2 are left out for conciseness.

5.6 Results

In this section, the empirical results on the contribution of local experience and different drilling technologies to crude oil and gas production are presented. First, the results of the FE, FD_t, CCEMG and CCEP estimator for the base specification Equation (5.3) are presented in Table 5.4. The first column shows the results of the FE estimator which does not take account of cross-sectional dependencies. The second column presents the results of the FD_t estimator which accounts

[19] This is possible by using the appendix of the EIA data where the counties included in each EIA region are listed.
[20] Alternative specifications where zero drilling activity was replaced by 10^{-3} or 10^{-6} instead of one were also estimated. There was almost no change in the estimation results.

for simple forms of cross-sectional dependency through the inclusion of time dummies. The last two columns show the results of the CCEMG and the CCEP estimators which account for general forms of cross-sectional dependency and the resulting potential endogeneity of the regressors. Only the CCMEG estimator implictly allows for parameter heterogeneity across cross-sectional groups. For this estimator the average of the group specific parameters is shown in all tables.

The drilling coefficients are significant and range between 0.59 and 0.72. For instance, the CCEP method estimates the effect of drilling on output to be 0.59. This means that a one percent increase in drilling activity leads to a 0.59 percent increase in output. Only the FE estimator using fixed effects for each cross-sectional group is as high as 0.91 which can certainly be discarded as an unreasonably high estimate. All estimates of the experience coefficient are between 0.34 and 0.55 indicating an

Table 5.4: Results base specification

Estimator	(1) FE	(2) FD_t	(3) CCEMG	(4) CCEP
Drilling (d)	0.91**	0.72**	0.66**	0.59**
	(3.31)	(16.49)	(7.56)	(9.04)
Experience (e)	0.48**	0.49**	0.34	0.55**
	(5.51)	(2.83)	(0.61)	(3.24)
CRS	0.00	0.25	0.99	0.45
CD test	0.00	0.45	0.82	0.03
Observations	686	672	686	686

Values in parentheses indicate t-statistics. *,** indicate significance at the 5 percent and the 1 percent level. CRS: Wald test for the null hypothesis of constant returns to scale of drilling and experience (p-values are reported). CD test: Pesaran (2004) test for H_0 of cross-sectionally independent residuals (p-values are presented). The estimates of the augmentation term coefficients and constants are left out for clarity of exhibition.

important role of local experience. For example, the CCEP estimator shows an output elasticity of local experience of 0.55. This means, a one percent increase in experience leads to a 0.55 percent increase in output. All estimates except the CCEMG estimates are statistically significant. The lack of significance for the CCMEG estimate could be resulting from the augmentation. Particularly, there could be a strong correlation of the cross-sectional averages of the experience variable with the group specific values the experience variable.

Cross-sectional independence is rejected for the FE and the CCEP estimators. This indicates that both estimators are not able to totally account for spillovers and should therefore be interpreted with care. Evidence of constant returns to scale is found for the FD_t, CCMEG and CCEP estimators which all account for cross-sectional dependency. Only for the FE estimator constant returns to scale are rejected. In conclusion, we found evidence for considerable local experience effects during the shale revolution.

As a robustness check the alternative specifications discussed in Section 5.3.2 are estimated. Table 5.5 shows the estimation results of the FD_t, CCEMG and CCEP estimators when a stronger depreciation of experience of $i = 0.97$ is assumed (see Columns 1, 3 and 5) or when the mature Permian region is excluded from the sample (see Columns 2, 4 and 6). Taken together, the results of the alternative specifications are qualitatively the same as in the base specification. A higher experience depreciation rate decreases the experience elasticity slightly to 0.47 for FD_t and to 0.31 for the CCMEG estimator, respectively, but increases the experience elasticity to 0.58 for the CCEP estimator. The exclusion of the Permian region increases the output elasticity of experience to 0.52, 0.81 and 0.61 for the FD_t, CCMEG and CCEP estimators, respectively. This relatively strong change could be caused by the comparatively stagnant production development, but persistent drilling activity in the Permian region over the sample period. Only the CCEMG estimates in column 4 seem to go far beyond the range from the base specification, albeit with no significance at conventional levels. In all cases constant returns to scale cannot be rejected. Similar to the results of the base specification above, cross-sectional dependency can only be rejected for the FD_t and the CCEMG estimates.

Table 5.5: Alternative specifications

Estimator	(1) FD_t	(2) FD_t	(3) CCEMG	(4) CCEMG	(5) CCEP	(6) CCEP
Drilling (d)	0.72**	0.73**	0.66**	0.70**	0.59**	0.60**
	(16.45)	(16.96)	(7.66)	(7.52)	(9.22)	(8.84)
Experience (e)	0.47**	0.52**	0.31	0.81	0.58**	0.61**
	(3.00)	(3.00)	(0.65)	(1.68)	(3.88)	(3.56)
Depreciation factor	0.97	0.99	0.97	0.99	0.97	0.99
Permian excluded	no	yes	no	yes	no	yes
CRS	0.25	0.17	0.96	0.29	0.26	0.25
CD test	0.44	0.21	0.77	0.35	0.03	0.02
Observations	672	576	686	588	686	588

Depreciation factor: The experience variable was constructed using a monthly depreciation factor of $i = 0.99$ in the base specification or $i = 0.97$ in the alternative specification. Permian excluded: "yes" ("no") if the Permian region is (not) excluded from the sample. See Table 5.4 for other explanations.

In order to address the role of drilling technologies and local experience in using these technologies, we estimate the disaggregated equation (5.6). As outlined in Section 5.3.2 the explanatory variables are now disaggregated into technology specific subcategories. Columns 1 and 2 of Table 5.6 show the estimation results for the CCEMG estimator with and without the Permian region. Likewise, the results for the CCEP estimator are shown in Columns 4 and 5.

As an additional robustness check Column 3 shows the CCEMG estimates of the experience elasticities when the explanatory variables are disaggregated by both technology and target depth (according to equation (5.7)).[21] Accordingly, the drilling rig variable is disaggregated into three technology

[21] Due to the extreme number of augmentation variables the CCEP estimator cannot be reasonably applied when using the disaggregation scheme shown in equation (5.7).

Table 5.6: Disaggregated specification

Estimator	(1) CCEMG	(2) CCEMG	(3) CCEMG	(4) CCEP	(5) CCEP
Drilling: Horizontal	0.39**	0.46**	-	0.32**	0.34*
	(6.01)	(8.36)	-	(3.02)	(2.92)
Drilling: Vertical	0.09**	0.06**	-	0.07**	0.07**
	(4.67)	(6.97)	-	(5.99)	(6.07)
Drilling: Directional	0.06**	0.06**	-	0.04*	0.04*
	(3.49)	(3.42)	-	(2.46)	(2.24)
Experience: Horizontal	0.63*	0.78**	0.45	0.36*	0.45*
	(2.35)	(2.65)	(1.58)	(2.35)	(2.46)
Experience: Vertical	-0.31	-0.09	-0.42	-0.28	-0.23
	(-1.80)	(-0.59)	(-1.46)	(-1.64)	(-1.26)
Experience: Directional	-0.14	-0.21	-0.06	-0.02	0.00
	(-0.77)	(-0.99)	(-0.43)	(-0.33)	(0.03)
Permian excluded	no	yes	no	no	yes
Depth disaggregation	no	no	yes	no	no
CRS	0.25	0.84	0.21	0.17	0.35
CD test	0.51	0.47	0.51	0.00	0.00
Observations	686	588	686	686	588

See Table 5.4 and Table 5.5 for explanations.

and four target depth subcategories.[22] All coefficients of the drilling variables are significantly different from zero. However, in terms of output elasticity horizontal drilling by far exceeds vertical and directional drilling ranging from 0.32 for the CCEP estimator in Column 4 to 0.46 for the CCEP estimator when the Permian region is excluded from the sample. This result confirms that horizontal drilling can indeed be regarded as the main drilling technology for the extraction of shale oil and gas. Further, only the local experience accumulated with the horizontal drilling technology has a significant effect on output. The coefficients of horizontal drilling experience range from 0.36 for the CCEP estimate in Column 4 to 0.78 for the CCEMG estimate when the Permian region is excluded (Column 2). The CCEMG experience coefficients of horizontal drilling are significant and substantially higher than the effect of aggregated experience in the base specification. The CCEP coefficients of horizontal drilling are somewhat lower than for the aggregated experience variable in Table 5.4 and 5.5. Moreover, the importance of local experience with horizontal drilling rises and the (insignificant) coefficient of vertical drilling experience shrinks when the stagnant Permian region is excluded. As before cross-sectional dependence cannot be rejected for the CCEP results, which could indicate that experience spillovers might not be completely purged from the estimates. Hence, the CCEMG estimates should be regarded as the most appropriate estimates of local experience effects on the production of shale oil and gas.

In summary, the results provide robust evidence for strong local experience effects for the production

[22] For the clarity of presentation, the estimates of the corresponding twelve drilling rig variables are not shown in Table 5.6.

of shale oil and gas. In comparison, Kellogg (2011) investigated experience effects on the cost side of conventional oil drilling. He focused on how much experience, measured as the duration of drilling activity, lowered the time and thereby cost to drill individual oil wells.[23] Overall, he found that a year of additional drilling experience lowered drilling time in the low double digit percentage range. Even if the results are not directly comparable, our results indicate a substantially larger role of local experience for the production of shale oil and gas.[24] Whereas Kellogg (2011) and most of the petroleum engineering literature focus on the input-per-well side of drilling experience effects, our results highlight that the shale revolution is characterized by strong local experience effects on the output-per-well side.

5.7 Conclusion

This paper investigated the drivers of the sweeping production gains of shale oil and gas during the US shale revolution. In particular, we focused on the role of local experience and drilling technology. We estimated this relationship using panel time series estimators which take account of experience spillovers and other cross-sectional dependencies and thereby allow the identification of local experience effects. The results exhibit strong local experience effects. Our results also demonstrate the importance of horizontal drilling technology for the production of shale oil and gas. Moreover, the local experience gathered with horizontal drilling rigs has a very high impact on production.

These results indicate that the dramatic gains in US shale oil and gas production are a result of both using horizontal drilling technology as well as building a large local experience stock with that technology. Hence, a replication of the US shale revolution in other countries not only depends on updating the technological level of the drilling rig fleet, but also on the possibility to accumulate a large local experience stock using these technologies. In several countries with shale resources (e.g. in Europe), limiting factors such as a high population density, stringent regulation and public acceptance problems make it unlikely that a large number of drilling operations can be realized. This constraint to the accumulation of local experience could hamper the growth in shale oil and gas production in these countries.

A suggestion to accumulate local experience faster could be to focus on what Dar-El (2000) calls the "learning organization". The new concept of "industrial drilling" and the other idiosyncracies

[23] Kellogg (2011) used data for Texas in the period 1991 to 2005, i.e., before the shale revolution. In his paper, the well was interpreted as the output, whereas drilling time constituted the input.

[24] In contrast to Kellogg (2011) we interpret the drilling rigs as inputs and extracted oil and gas as outputs.

of shale drilling discussed in Section 5.2.2 entail new approaches to increase the speed of collecting information. "Exporting" these new approaches for information procurement could lead to a faster accumulation of local experience in other countries and hence to less difficulties in replicating the shale revolution. Further research could address the size and patterns of interregional as well as intraregional technology and experience spillovers. Moreover, the relationship between the costs of shale drilling, experience and drilling technology could be investigated. Thereby, a more complete picture of the role of experience and technology in the hydrocarbon industry could be drawn. With more disaggregated well and firm level data different types of learning effects such as individual worker experience effects in contrast to organizational firm level experience effects could be identified.[25]

[25] Such data, however, are only available at very high prices from private information agencies.

Appendix A

Supplementary Material for Chapter 2

Table A.1: Variable description and sources

Variable	Details	Source
Gas drilling	Number of active rigs drilling for gas	US Energy Information Agency, Baker Hughes Inc (2013)
Oil drilling	Number of active rigs drilling for oil	US Energy Information Agency, Baker Hughes Inc (2013)
Gas price	Real US wellhead price, base year 2005	US Energy Information Agency (2013)
Oil price	Real US WTI oil price, base year 2005	US Energy Information Agency (2013)
Rig utilization	Ratio of rigs actively drilling to total number of onshore rigs in the oil rig fleet	Guiberson-AESC Well Service Rig Count (2013)
Macroeconomy	Industrial production index measures real output for all facilities located in the US in manufacturing, mining, and electric, and gas utilities, base year 2007	Federal Reserve Economic Data (2013)

Table A.2: Descriptive statistics

Variable	Observations	Mean	Standard Deviation	Minimum	Maximum
Gas drilling	170	954.61	318.03	371.00	1585.00
Oil drilling	170	372.38	344.40	108.00	1423.00
Gas price	168	4.64	1.83	1.65	10.25
Oil price	168	53.47	24.63	13.20	123.62
Rig utilization	171	0.65	0.07	0.48	0.76
Macroeconomy	171	92.62	4.30	83.46	100.74

Notes: The table shows the basic statistics for all variables over the time period October 1998 to September 2012 used in the models above.

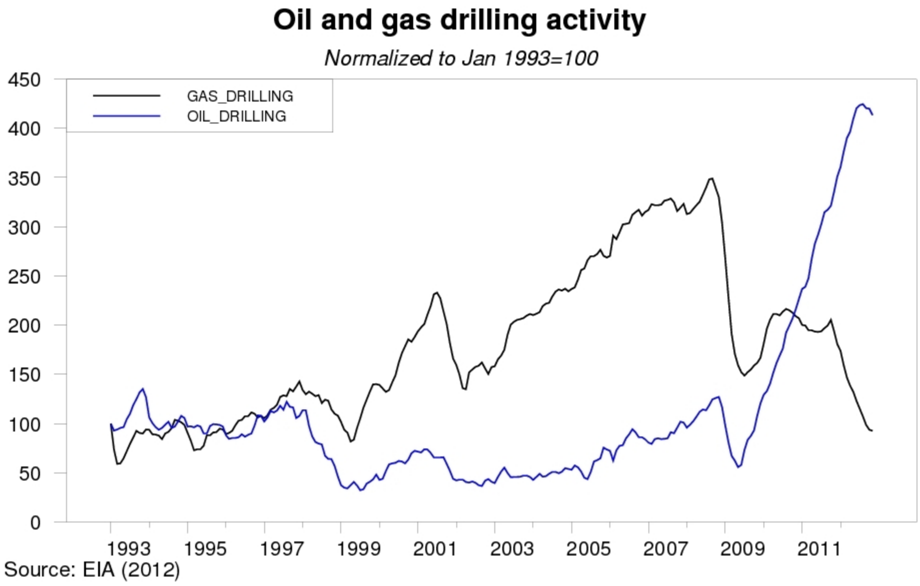


Figure A.1: Oil and gas drilling activity [Number of Rigs] 1993-2012

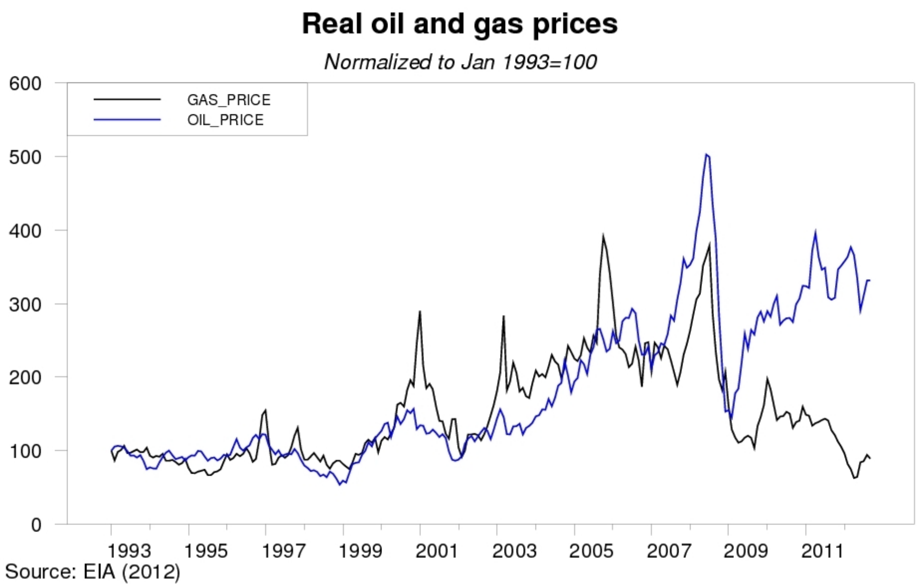


Figure A.2: Real oil and gas price [normalized to Jan 1993 = 100] 1993-2012

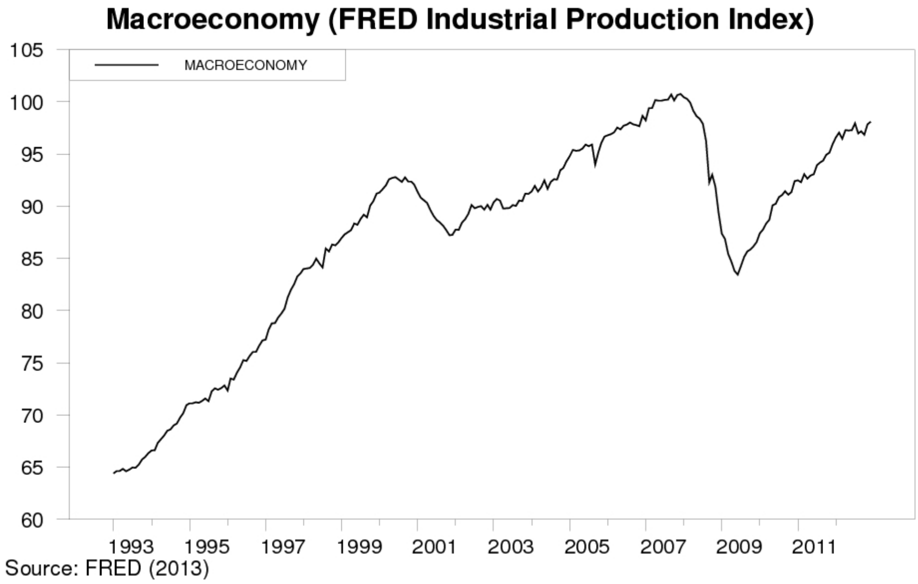


Figure A.3: Macroeconomy 1993-2012

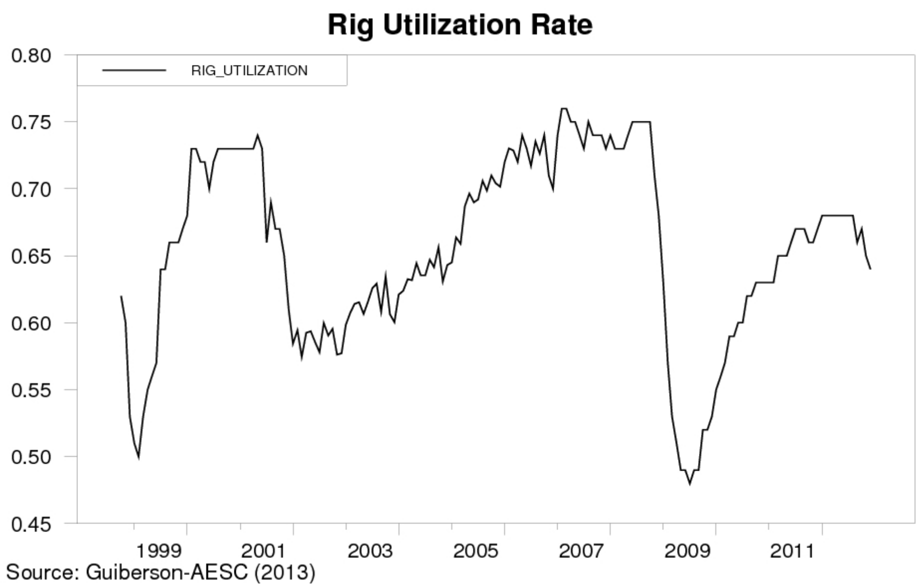


Figure A.4: Rig utilization rate 1998-2012

Table A.3: Deterministic regressor and event dummy variables

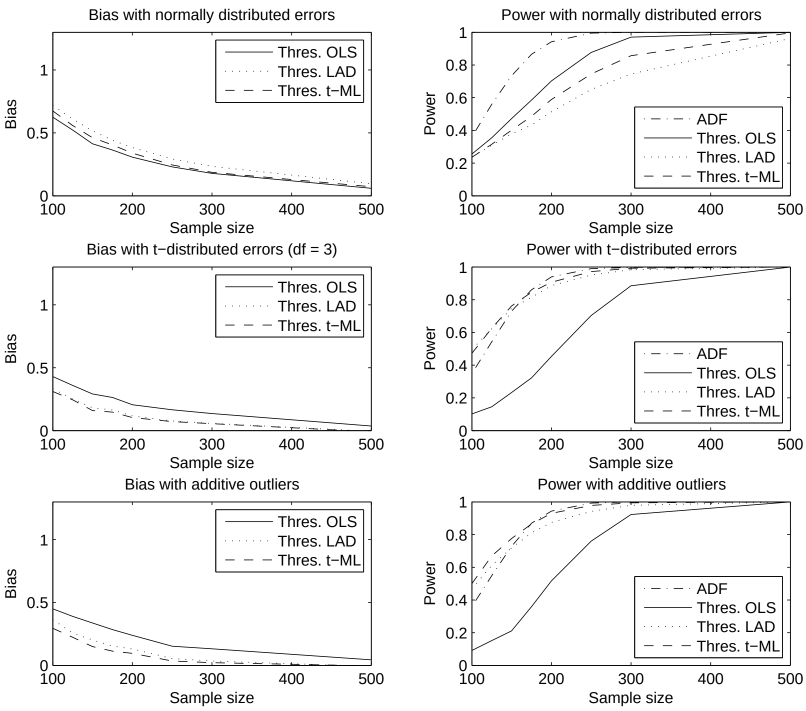
Event	Date	Variable type	Details
Storage gas price hike Winter 2000/2001	January 2001	Impulse dummy	California energy crisis, low gas storage levels and a very cold winter led to a major gas price spike
Gas price spike 2003	March and April 2003	Impulse dummy for each month	Major price spike particularly for gas prices
Hurricane Season 2005	August to October 2005	Impulse dummy for each month	Hurricanes Katrina, Wilma and Rita
US financial crisis	August 2008 to April 2009.	Impulse dummy for each month	
Oil drilling trend shift after economic crisis	May 2009	Linear trend shift	Oil drilling changes to a steep upwards trend after the economic crisis likely corresponding to the advent of unconventional oil drilling technology
Seasonal Dummies	Monthly	Impulse Dummies	

Appendix B

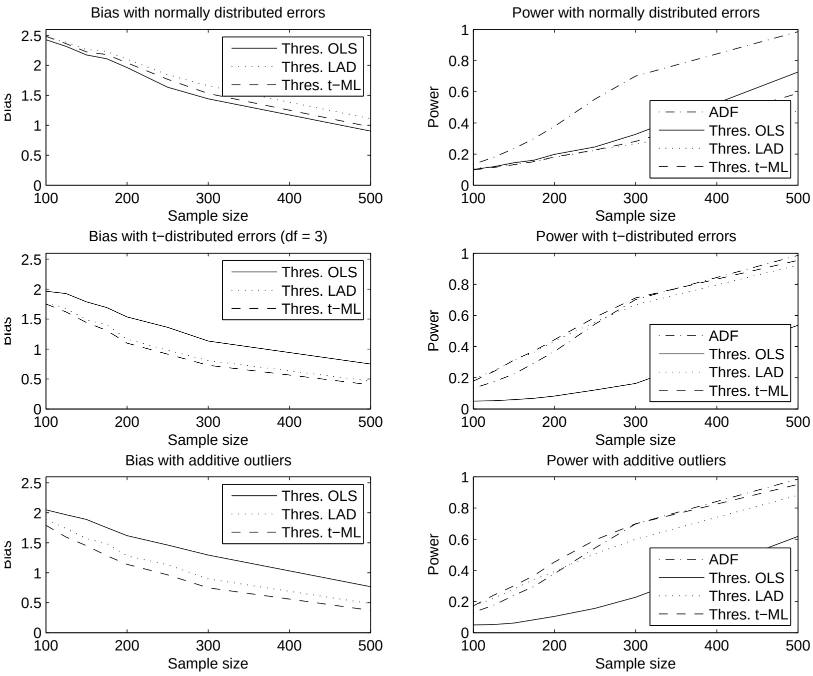
Supplementary Material for Chapter 4

B.1 Monte Carlo Alternative Specifications: Bias and Power for BAND-TAR models

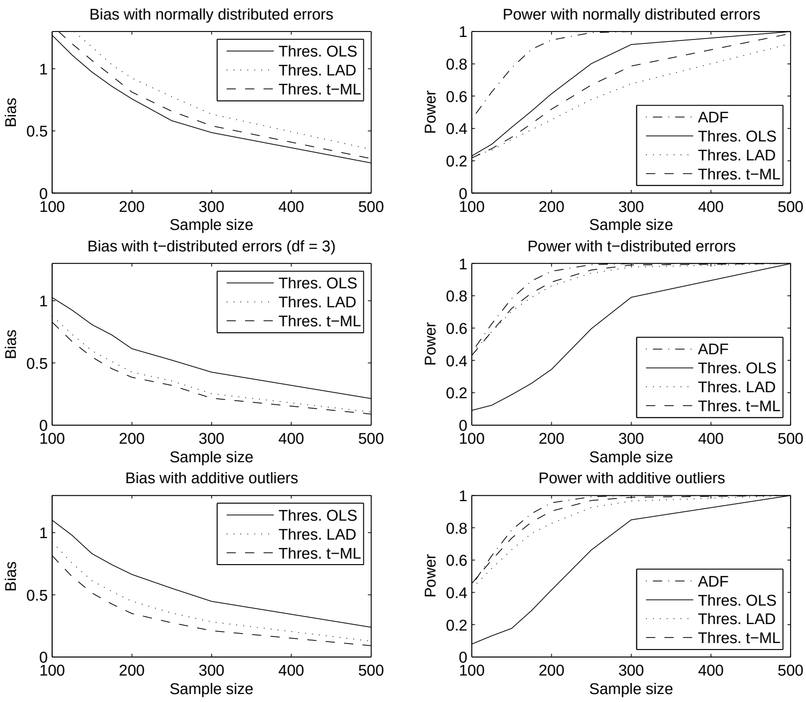
BAND-TAR specification: $\tau = 2$ and $\rho = 0.85$



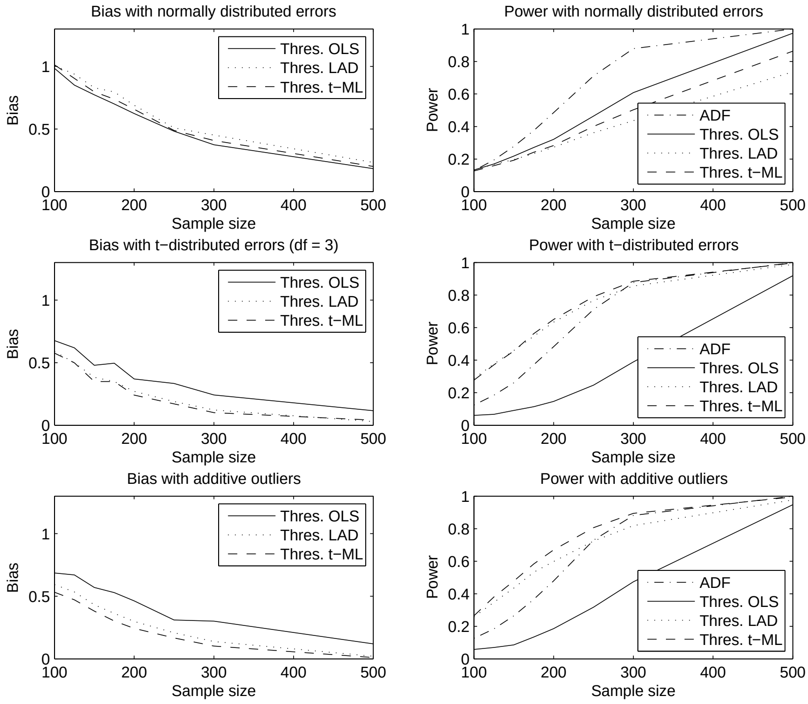
BAND-TAR specification: $\tau = 2$ and $\rho = 0.95$



BAND-TAR specification: $\tau = 1$ and $\rho = 0.90$

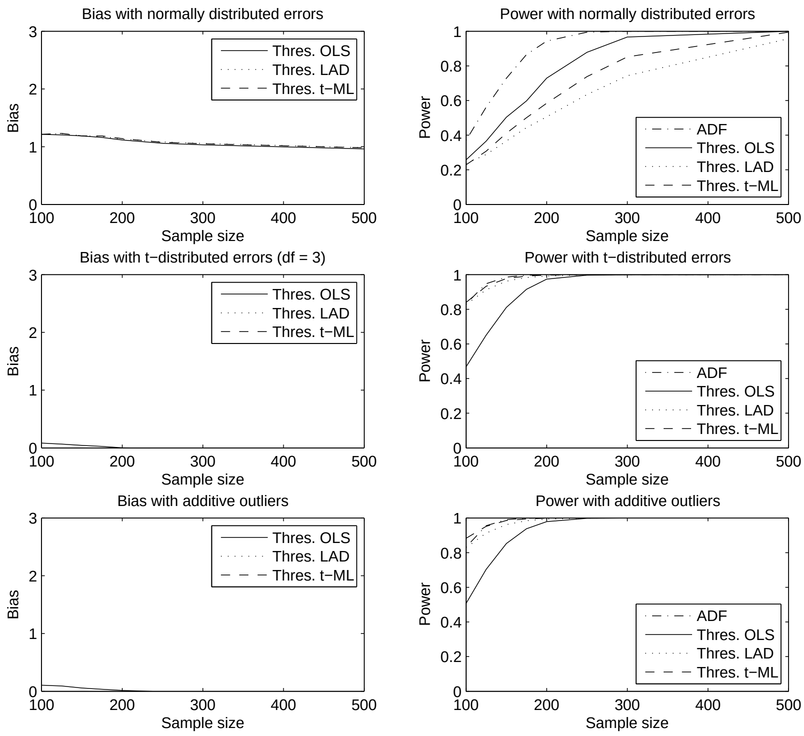


BAND-TAR specification: $\tau = 3$ and $\rho = 0.90$

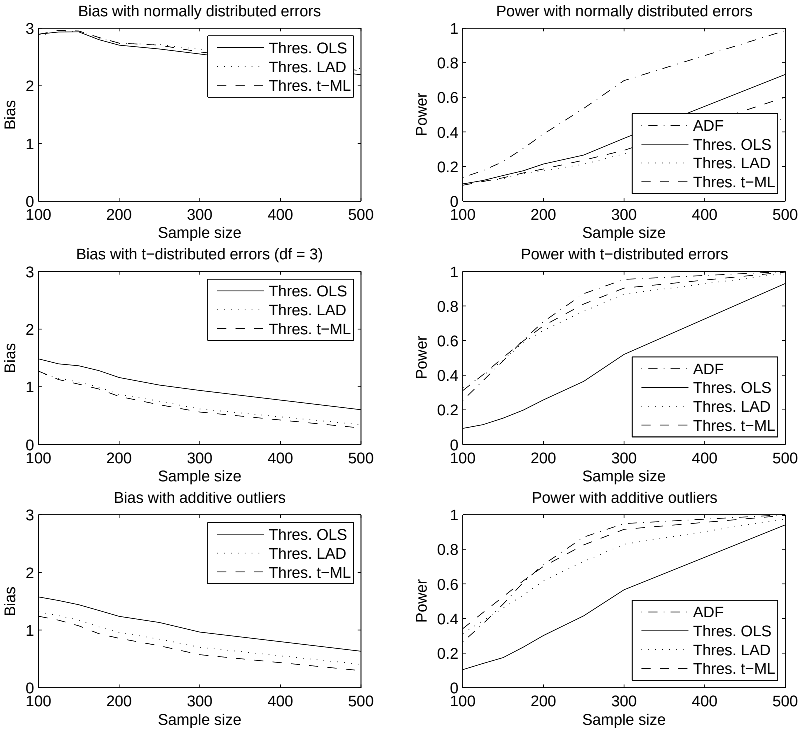


B.2 Monte Carlo Alternative Specifications: Bias and Power for EQ-TAR models

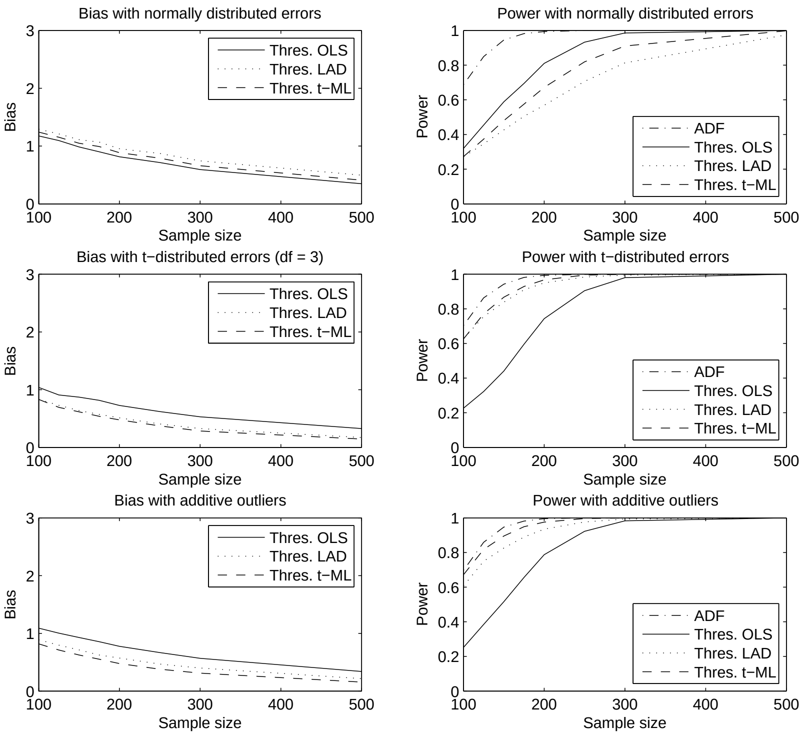
EQ-TAR specification: $\tau = 2$ and $\rho = 0.85$



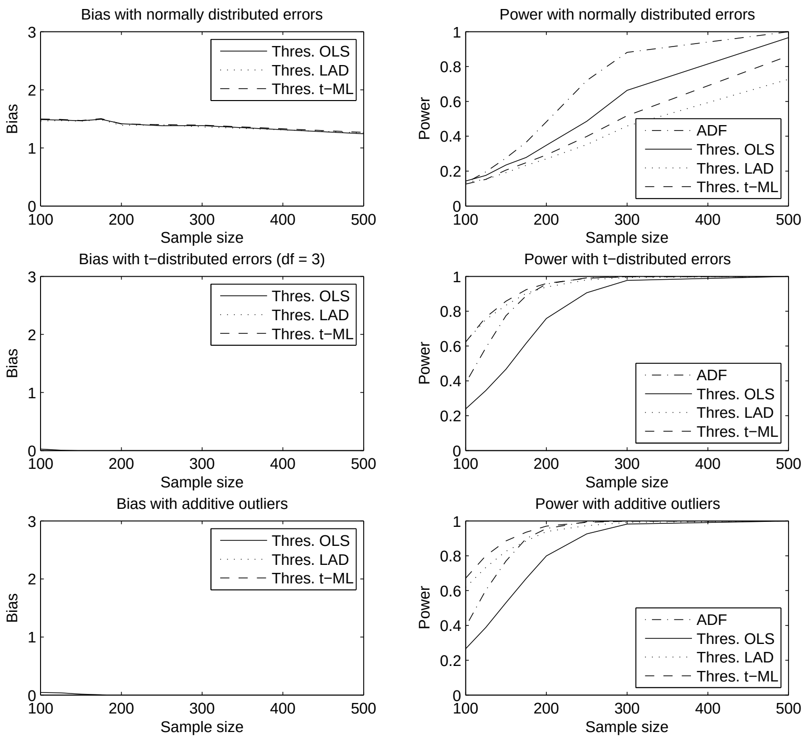
EQ-TAR specification: $\tau = 2$ and $\rho = 0.95$



EQ-TAR specification: $\tau = 1$ and $\rho = 0.90$



EQ-TAR specification: $\tau = 3$ and $\rho = 0.90$



Appendix C

Supplementary Material for Chapter 5

Table C.1: Zero-to-one replacements in the drilling and experience variables

Variable	Notation	Percentage share of replacements
Drilling	d	0.0
Experience	e	0.0
Drilling: Horizontal	d_{hor}	1.1
Drilling: Vertical	d_{ver}	1.1
Drilling: Directional	d_{dir}	0.0
Experience: Horizontal	e_{hor}	6.1
Experience: Vertical	e_{ver}	0.0
Experience: Directional	e_{dir}	12.2
Drilling: Horizontal,depth <5000	$d_{h=1,hor}$	58.0
Drilling: Horizontal,depth 5000-10000	$d_{h=2,hor}$	15.5
Drilling: Horizontal,depth 10000-15000	$d_{h=3,hor}$	1.7
Drilling: Horizontal,depth >15000	$d_{h=4,hor}$	17.8
Drilling: Vertical,depth <5000	$d_{h=1,ver}$	59.5
Drilling: Vertical,depth 5000-10000	$d_{h=2,ver}$	14.0
Drilling: Vertical,depth 10000-15000	$d_{h=3,ver}$	43.4
Drilling: Vertical,depth >15000	$d_{h=4,ver}$	80.5
Drilling: Directional,depth <5000	$d_{h=1,dir}$	81.9
Drilling: Directional,depth 5000-10000	$d_{h=2,dir}$	36.2
Drilling: Directional,depth 10000-15000	$d_{h=3,dir}$	51.3
Drilling: Directional,depth >15000	$d_{h=4,dir}$	73.8

Bibliography

Amisano, G. and Giannini, C. (1997). Topics in Structural VAR Econometrics. *Berlin/Heidelberg*.

Arrow, K. J. (1962). The Economic Implications of Learning by Doing. *The Review of Economic Studies*, 29(3):155–173.

Balke, N. S. and Fomby, T. B. (1997). Threshold Cointegration. *International Economic Review*, 38(3):627–645.

Basset, G. and Koenker, R. (1978). The Asymptotic Distribution of the Least Absolute Error Estimator. *Journal of the American Statistical Association*, 73:618–622.

Berndt, E. R. (1991). *The Practice of Econometrics: Classical and Contemporary*. Addison-Wesley Publishing Co.

Boyce, J. R. and Nostbakken, L. (2011). Exploration and Development of U.S. Oil and Gas Fields, 1955-2002. *Journal of Economic Dynamics and Control*, 35(6):891–908.

Breitung, J., Brüggemann, R., and Lütkepohl, H. (2004). Structural Vector Autoregressive Modeling and Impulse Responses. In Lütkepohl, H. and Krätzig, M., editors, *Applied time series econometrics*, Themes in modern econometrics, chapter 4, pages 159–195. Cambridge University Press, Cambridge [u.a.].

Brown, S. P. A. and Yücel, M. K. (2009). Market Arbitrage: European and North American Natural Gas Prices. *The Energy Journal*, Special Issue:167–185.

Brüggemann, R. and Lütkepohl, H. (2000). Lag Selection in Subset VAR Models with an Application to a U.S. Monetary System. Number 0821.

Caner, M. and Hansen, B. E. (2001). Threshold Autoregression with a Unit Root. *Econometrica*, 69:1555–1596.

Carstensen, K. (2003). The Finite-Sample Performance of Robust Unit Root Tests. *Statistical Papers*, 44(4):469–482.

Chan, K. (1993). Consistency and Limiting Distribution of the Least Squares Estimator of a Threshold Autoregressive Model. *The Annals of Statistics*, 21(1):520–533.

Chan, W.-S. and Cheung, S.-H. (1994). On Robust Estimation of Threshold Autoregressions. *Journal of Forecasting*, 13:37–49.

Dar-El, E. M. (2000). *Human Learning: From Learning Curves to Learning Organizations*. Kluwer Academic Publishers, Dordrecht.

Dijk, D. V., Franses, P. H., and Lucas, A. (1999). Testing for Smooth Transition Nonlinearity in the Presence of Outliers. *Journal of Business & Economic Statistics*, 17(2):pp. 217–235.

Elliott, G., Rothenberg, T. J., and Stock, J. H. (1996). Efficient Tests for an Autoregressive Unit Root. *Econometrica*, 64(4):813–836.

Enders, W. (2001). Improved Critical Values for the Enders-Granger Unit-Root Test. *Applied Economics Letters*, 8:257–261.

Enders, W. (2008). *Applied Econometric Time Series*. John Wiley & Sons.

Enders, W. and Granger, C. W. J. (1998). Unit-Root Tests and Asymmetric Adjustment with an Example Using the Term Structure of Interest Rates. *Journal of Business & Economic Statistics*, 16(3):304–311.

Engle, R. F. and Granger, C. W. J. (1987). Co-Integration and Error Correction: Representation, Estimation, and Testing. *Econometrica*, 55(2):251–76.

Engle, R. F. and Yoo, B. S. (1987). Forecasting and Testing in Co-Integrated Systems. *Journal of Econometrics*, 35:143–159.

Farzin, Y. H. (2001). The Impact of Oil Price on Additions to US Proven Reserves. *Resource and Energy Economics*, 23(3):271–292.

Franses, P. H. and Lucas, A. (1998). Outlier Detection in Cointegration Analysis. *Journal of Business & Economic Statistics*, 16(4):pp. 459–468.

Gonzalo, J. and Pitarakis, J.-Y. (2006). Threshold Effects in Cointegrating Relationships. *Oxford Bulletin of Economics and Statistics*, 68:813–833.

Gospodinov, N., Herrera, A., and Pesavento, E. (2013). Unit Roots, Cointegration and Pre-Testing in VAR Models. In Fomby, T., L., K., and Murphy, A., editors, *VAR Models in Macroeconomics, New Developments and Applications: Essays in Honor of Christopher A. Sims*, Advances in econometrics; 32. Bingley [u.a.] : Emerald, 1. ed. edition.

Hall, P. (1992). *The Bootstrap and Edgeworth Expansion*. New York [etc.]: Springer.

Hamilton, J. (2009). Understanding Crude Oil Prices. *The Energy Journal*, Volume 30(Number 2):179–206.

Hampel, F., Ronchetti, P., Rousseeuw, P., and Stahel, W. A. (1986). *Robust Statistics. The Approach Based on Influence Functions*. Wiley, New York.

Herce, M. A. (1996). Asymptotic Theory of LAD Estimation in a Unit Root Process with Finite Variance Errors. *Econometric Theory*, 12:129–153.

Horvath, M. T. K. and Watson, M. W. (1995). Testing for Cointegration When Some of the Cointegrating Vectors are Prespecified. *Econometric Theory*, 11(5):984–1014.

Huber, P. (1981). *Robust Statistics*. Wiley, New York.

IGU (2014). IGU World LNG Report 2014. Igu report, International Gas Union, Fornebu, Norway.

IHS (2013). Americaâs New Energy Future: The Unconventional Oil and Gas Revolution and the US Economy. Ihs report, IHS Inc., Washington, D.C.

Iledare, O. O. (1995). Simulating the Effect of Economic and Policy Incentives on Natural Gas Drilling and Gross Reserve Additions. *Resource and Energy Economics*, 17(3):261–279.

Johansen, S. (1988). Statistical Analysis of Cointegration Vectors. *Journal of economic dynamics and control*, 12(2):231–254.

Johansen, S. and Juselius, K. (1990). Maximum Likelihood Estimation and Inference on Cointegration with Applications to the Demand for Money. *Oxford Bulletin of Economics and statistics*, 52(2):169–210.

Johansen, S. and Juselius, K. (1992). Testing Structural Hypotheses in a Multivariate Cointegration Analysis of the PPP and UIP of UK. *Journal of Econometrics*, 53:211–244.

Kao, C.-W. and Wan, J.-Y. (2009). Information Transmission and Market Interactions Across the Atlantic: An Empirical Study on the Natural Gas Market. *Energy Economics*, 31:152–161.

Kapetanios, G., Pesaran, H. M., and Yamagata, T. (2011). Panels with Non-Stationary Multifactor Error Structures. *Journal of Econometrics*, 160(2):326 – 348.

Kellogg, R. (2011). Learning by Drilling: Interfirm Learning and Relationship Persistence in the Texas Oilpatch. *The Quarterly Journal of Economics*, 126(4):1961–2004.

Kellogg, R. (2014). The Effect of Uncertainty on Investment: Evidence from Texas Oil Drilling. *American Economic Review*, 104(6):1698–1734.

Koenker, R. and Basset, G. (1982). Tests of Linear Hypothesis and L1 Estimation. *Econometrica*, 50:1577–1583.

Kwiatkowski, D., Phillips, P. C., Schmidt, P., and Shin, Y. (1992). Testing the Null Hypothesis of Stationarity Against the Alternative of a Unit Root: How Sure are we that Economic Time Series Have a Unit Root? *Journal of Econometrics*, 54(1):159–178.

Li, R., Joyeux, R., and Ripple, R. D. (2014). International Natural Gas Market Integration. *The Energy Journal*, 35(4):159–179.

Liebermann, M. (1984). The Learning Curve and Pricing in the Chemical Processing Industries. *The Rand Journal of Economics*, 15:213–228.

Lo, M. C. and Zivot, E. (2001). Threshold Cointegration and Nonlinear Adjustment to the Law of One Price. *Macroeconomic Dynamics*, 5:533–576.

Lucas, A. (1995a). An Outlier Robust Unit Root Test with an Application to the Extended Nelson-Plosser Data. *Journal of Econometrics*, 66:153–173.

Lucas, A. (1995b). Unit Root Tests Based on M Estimators. *Econometric Theory*, 11:331–346.

Lucas, A. (1997). Robustness of the Student t Based M-Estimator. *Communications in Statistics, Theory and Methods*, 26:1165–1182.

Lütkepohl, H. (2004). Vector Autoregressive and Vector Error Correction Models. In Lütkepohl, H. and Krätzig, M., editors, *Applied time series econometrics*, Themes in modern econometrics, chapter 3, pages 86–158. Cambridge University Press, Cambridge [u.a.].

Lütkepohl, H. (2005). *New Introduction to Multiple Time Series Analysis*. Springer, Berlin.

Maxwell, D. and Zhu, Z. (2011). Natural Gas Prices, LNG Transport Costs, and the Dynamics of LNG Imports. *Energy Economics*, 33:217–226.

Medlock, K. B., Myers-Jaffe, A., and Hartley, P. R. (2011). Shale Gas and U.S. National Security. Report, James A. Baker III Institute for Public Policy, Rice University.

Muth, J. F. (1961). Rational Expectations and the Theory of Price Movements. *Econometrica*, 29(3):315–335.

Neumann, A. (2009). Linking Natural Gas Markets - Is LNG Doing its Job? *The Energy Journal*, Special Issue:187–200.

Nielsen, H. B. (2004). Cointegration Analysis in the Presence of Outliers. *The Econometrics Journal*, 7(1):pp. 249–271.

Pesaran, H. M. (2006). Estimation and Inference in Large Heterogeneous Panels with a Multifactor Error Structure. *Econometrica*, 74(4):967–1012.

Pesaran, H. M. and Smith, R. (1995). Estimating Long-Run Relationships from Dynamic Heterogeneous Panels. *Journal of Econometrics*, 68(1):79–113.

Pesaran, M. (2004). General Diagnostic Tests for Cross Section Dependence in Panels. Cambridge Working Papers in Economics 0435, Faculty of Economics, University of Cambridge.

Pollard, D. (1991). Asymptotics for Least Absolute Deviation Regression Estimators. *Econometric Theory*, 7(2):186–199.

Ramberg, D. J. and Parsons, J. E. (2012). The Weak Tie Between Natural Gas and Oil Prices. *The Energy Journal*, 2(33).

Renshaw, E. F. (1989). An Oil Import Fee and Drilling Activity in the USA. *Energy Economics*, 11(2):158–160.

Ringlund, G. B., Rosendahl, K. E., and Skjerpen, T. (2008). Does Oilrig Activity React to Oil Price Changes An Empirical Investigation. *Energy Economics*, 30(2):371–396.

Seo, M. H. (2008). Unit Root Test in a Threshold Autoregression: Asymptotic Theory and Residual-Based Block Bootstrap. *Econometric Theory*, 24(6):1699–1716.

Siliverstovs, B., L'Hagaret, G., Neumann, A., and von Hirschhausen, C. (2005). International Market Integration for Natural Gas? A Cointegration Analysis of Prices in Europe, North America and Japan. *Energy Economics*, 27(4):603 – 615.

Tong, H. (1978). On a Threshold Model. In Chen, C., editor, *Pattern recognition and signal processing*. Sijhoff and Noordhoff, Amsterdam.

Wright, T. P. (1936). Factors Affecting the Cost of Airplanes. *Journal of the Aeronautical Sciences*, 3(4):122–128.

Zhang, L.-X., Chan, W.-S., Cheung, S.-H., and Hung, K.-C. (2008). A Note on the Consistency of a Robust Estimator for Threshold Autoregressive Processes. *Statistics and Probability Letters*, 79:807–813.