Modelling Nonlinearities in Cointegration Relationships

A thesis submitted to the Faculty of Business, Economics and Social Sciences, University of Hohenheim, in partial fulfillment of the requirements for the Degree

Dr. oec.

Karsten Schweikert

Department of Econometrics and Statistics

Date of submission: April 2017

This thesis has been accepted in 2017 as a dissertation to attain the degree of *Doctor oeconomiae (Dr. oec.)* by the Faculty of Business, Economics and Social Sciences at the University of Hohenheim.

Defense of doctoral thesis: July 13, 2017

Professor Dr. Dirk Hachmeister, Dean, University of Hohenheim
Professor Dr. Robert C. Jung, Supervisor, University of Hohenheim
Privatdozent Dr. Thomas Dimpfl, Second assessor, University of Tuebingen
Professor Dr. Aderonke Osikominu, Chair of the doctoral committee, University of Hohenheim

Contents

List of Figures

List of Tables

Acknowledgements

Numerous people helped me in the process of exploring the ideas put forth in this thesis: First, and foremost, I am greatly indebted to my supervisor Professor Dr. Robert C. Jung for giving me the opportunity to join his research group. I thank him for his guidance, advice and motivation. Equally important has been Dr. Karl-Heinz Schild in introducing me to the field of time series analysis. Without his interesting and passionate lectures and his support during the writing of my master thesis I would not have taken up my doctoral studies. I also thank Professor Dr. Karlheinz Fleischer for trusting me with the lecture in time series econometrics which gave me invaluable teaching experience. As my second supervisor, I thank Privatdozent Dr. Thomas Dimpfl. Also, I thank Professor Dr. Aderonke Osikominu for joining the committee for the defence of my thesis. Further, I want to thank my fellow colleagues for interesting discussions and advice: Dr. Robert Maderitsch, Konstantin Kuck, Dr. Stephanie Glaser, Marie-Sophie Vasamiliette, Alexander Schmidt and Tanja Strohäker. Also, I want to thank the organizers and the participants of the Tübingen-Konstanz-Hohenheim doctoral workshops for a stimulating scientific exchange. Finally, I thank my family and friends for their support at all times.

Chapter 1

Introduction

This thesis is concerned with the statistical modelling of long-run equilibrium relationships between economic variables. Widely used concepts in statistical theory to describe long-run equilibrium relationships are cointegration and error correction as defined in Engle and Granger (1987). Many economic variables such as prices, production aggregates and wages exhibit stochastic trends, which means that shocks have a permanent effect on the trajectory of the time series. The implications of stochastic trends is that the joint distribution of the stochastic process that generates the times series is nonstationary and the integrated process does not revert to a stable long-run attractor.

Cointegration is a multivariate concept and refers to the statistical property that two or more nonstationary (integrated) variables (e.g. spot and futures prices, input and output prices) do not diverge without bound from each other. Thus, cointegrated variables form a stable long-run relationship although the individual processes are not stable. Such relationships are often associated with economic equilibria. A disequilibrium is any situation in which the variables are not in exact equilibrium. For the long-run relationship to be stable, deviations from equilibrium, however, have to be of temporary nature. The system has to have a tendency to return to the equilibrium after it is perturbed. Throughout the remainder of this thesis, we refer to this behavior as error correction or adjustment.

While it is sometimes possible to derive the form of the cointegrating relationship from economic theory, its exact form is typically unknown in empirical applications. Cointegration models provide methods to estimate the parameters of the long-run equilibrium equation if it is not known ex ante and model the adjustment behaviour after disequilibrium states. An important part of empirical cointegration models consists of testing for the presence of a cointegration relationship since any results from a cointegration model would be spurious in the absence of cointegration.

The existence of cointegration relationships is of particular interest for economists since many theoretical models are based on long-run relationships between economic variables (e.g. purchasing power parity, Fisher equation, unbiasedness hypothesis). However, the findings in empirical studies are often ambiguous (see, for example, the discussion on purchasing power parity). A reason for this might be the restrictive nature of conventional cointegration models. Turning to nonlinear dynamics might therefore improve estimation and inference.

The original Engle-Granger cointegration model is linear, in that is requires a linear combination of nonstationary variables to be stationary. The stationarity of the linear combination is empirically tested by conventional unit root tests which build on linear autoregressive models. The choice of a linear model specification is easily justifiable considering the availability of statistical tools at the time of Grangers' discovery of cointegration, published in Granger (1981). A general theory of statistical inference for multivariate integrated processes did not arrive until Phillips and Durlauf (1986) using results from Phillips (1986, 1987a,b) and Stock (1987). From a mathematical perspective, estimation and inference involving nonstationary variables are conducted more easily in a linear context. Further, the computational costs are usually lowest for linear models.

Nevertheless, there are good reasons for expecting economic relationships to be nonlinear. For example, business cycle dynamics, transaction costs, diminishing marginal utility or policy interventions potentially lead to mechanisms which are hardly captured in linear mathematical models. The specification of linear econometric models serves as a useful approximation in most cases but might be discarded for meaningful nonlinear models. Also, the evolution of computing power makes it feasible to estimate and evaluate complicated nonlinear models efficiently.

Cointegrated systems can be described in several alternative representations. The two statistical approaches to modelling cointegration relationships used in this thesis are the two-step single equation approach by Engle and Granger (1987) and the vector error correction model by Johansen (1988, 1991). Both approaches are based on linear parametric models and assume the parameters to be constant over time. This presents a natural starting point to relax the restriction of a linear model and move to a nonlinear model specification.

In the following, a selective review of relevant literature on nonlinear extensions of cointegration models is presented. This should help to embed the methods that are developed in this thesis and the existing methods that are used in the empirical applications in the current econometric literature. In the context of the Engle-Granger framework, several approaches have been proposed to extend the original model. Nonlinearities can be introduced to this framework at both steps.

A straightforward extension to the first step, the cointegrating regression, is to include a linear deterministic trend variable as in Engle and Yoo (1987). Strictly speaking, this extension is not a nonlinear one from a technical perspective, although it implies that the cointegration relationship is not static and gradually changes with time. More frequently, the cointegration relationship is thought to change instantly caused by policy changes or events such as economic crises. In this spirit, structural break models deal with changes in the cointegrating vector and any deterministic terms that are present in the initial model. Gregory and Hansen (1996a,b) introduce cointegration tests under the presence of structural breaks. The break date does not have to be known for these tests, which is usually the case in empirical applications. Often it is not even known how many structural breaks have to be accounted for. An extension of the Gregory-Hansen model to two structural breaks is given in Hatemi-J (2008). Arai and Kurozumi (2007) develop tests for the null hypothesis of cointegration with structural break against the alternative of no cointegration.

Gonzalo and Pitarakis (2006a) consider a cointegration regression with threshold nonlinearity. The variable that governs the change from one equilibrium to the next is a stationary threshold variable. If the threshold variable has crossed the threshold in a given period $t-d$, the slope coefficients of the cointegrating vector change in period t. Threshold nonlinearity allows for regime-specific behaviour depending, for example, on the phases of the business cycle.

Xiao (2009) applies quantile regression methods, developed in Koenker and Bassett (1978), to the cointegrating equation to obtain quantile-dependent coefficients. Instead of modelling the conditional expected value in the case of least squares estimation, quantile regression estimates the τth quantile of the dependent variable conditional on the information set in period t. The quantile cointegration model can be understood as a restricted form of a general random coefficient model.

Random coefficient models are a unifying framework for several forms of nonlinearities in which the coefficients depend on a stochastic process (Nicholls and Quinn (1982)). They are not easily estimated with conventional regression techniques without imposing some structure on the behaviour of the coefficients. Quintos and Phillips (1993) test for constancy of the cointegrating vector in a model that allows for a random walk process of the slope coefficient. Another way to deal with random coefficient models in the context of state-space models is detailed in Wagner (2010).

Saikkonen and Choi (2004) develop the asymptotic theory for cointegrating regressions with a smooth transition structure in the spirit of Granger and Terasvirta (1993). Cointegrating smooth transition regressions can be used to describe long-run relations that change smoothly depending on the location of some economic variables. The transition function helps to model a smooth change from one equilibrium to the next instead of an abrupt change from one equilibrium to the next in the threshold model by Gonzalo and Pitarakis (2006a).

The second step of the Engle-Granger procedure, unit root testing of the cointegration residuals, presents an even wider variety of nonlinear models. Theoretically, all types of unit root tests could be applied to the cointegration residuals, to evaluate whether the order of integration has been reduced for a linear combination of the variables. Conventional unit root tests are constructed to test the speed of adjustment after disequilibrium states.

Balke and Fomby (1997) suggest a threshold process for the equilibrium error. The equilibrium error is assumed to follow a threshold autoregression that is mean-reverting outside a range specified by two threshold values and has a unit root inside the range. This type of model accounts for the presence of transaction costs which might prevent adjustment after small shocks. However, if the adjustment behavior in the outer regimes is strong enough, the cointegration relationship is maintained in the long-run. A cointegration model with threshold adjustment was also proposed by Enders and Siklos (2001). Their models is restricted to a regime-specific coefficient of the first lag. The threshold variable is either the lagged equilibrium error series in levels, in a self-exciting threshold autoregression, or the differenced series which is a momentum threshold autoregression in the spirit of Enders and Granger (1998) and Caner and Hansen (2001). Maki and Kitasaka (2015) propose cointegation tests with three-regime threshold autoregressive adjustment.

Kapetanios *et al.* (2003) consider an exponential smooth transition (ESTAR) model for the equilibrium error, where the speed of adjustment is slower when the error is close to zero. The transition function is symmetrically U-shaped around zero. For large values of the smoothing parameter, the ESTAR model collapses to a linear model. In the same sense, a logistic smooth transition model (LSTAR) as proposed by Terasvirta (1994) can be used for the equilibrium error process. In this specification the logistic function links a regime of positive deviations to a regime of negative deviations. For large values of the smoothing parameter, the transition function effectively approaches an indicator function and the LSTAR model reduces to a two-regime threshold autoregressive model.

Hall *et al.* (1997) and Psaradakis *et al.* (2004) propose a Markov-switching model for the adjustment process. This framework allows for periods of strong adjustment behavior as well as periods in which the system can diverge temporarily from the long-run equilibrium. The equilibrium error follows a Markov-switching autoregression where a latent state variable governs the regime switches.

The Granger representation theorem guarantees that, if some variables are cointegrated, an error correction representation of the variables exists. A vector error correction model (VECM) developed in Johansen (1988, 1991) can be used to simultaneously test for cointegration and investigate the contribution of the individual variables to maintaining the long-run equilibrium. This framework is based on an autoregressive representation of a cointegrated system. It is augmented with nonlinearity concepts similar to the Engle-Granger case. However, the structure of the VECM naturally requires a very different implementation of these concepts.

Quintos (1997) considers a general time-varying structure for the reduced-rank matrix of error correction coefficients so that both the cointegrating vector and the adjustment dynamics may change over time. Seo (1998) develops a model for a one-time change of the cointegrating vector and of the adjustment coefficients at a potentially unknown change point. Johansen *et al.* (2000) and Inoue (1999) analyze breaks in the deterministic terms of a VECM. Further studies on cointegrationg test under structural breaks in a VECM are conducted in, for example, Saikkonen and Lütkepohl (2000), Lütkepohl *et al.* (2003), Lütkepohl *et al.* (2004), Trenkler *et al.* (2007) and Harris *et al.* (2016).

Threshold models in a multivariate framework were first examined in Tsay (1998). Hansen and Seo (2002) describe an estimation and testing procedure for a VECM with unknown cointegrating vector and unknown threshold value. Seo (2006) discusses bootstrap tests of the null hypothesis of no cointegration in threshold VECM, while Gonzalo and Pitarakis (2006b) develop an asymptotic theory for testing the existence of a threshold effect in a VECM. Krishnakumar and Neto (2015) extend the threshold VECM to more than one cointegrating relation.

Kapetanios *et al.* (2006) and Kiliç (2011) provide testing methodology and asymptotic theory for exponential and logistic smooth transition VECM. Saikkonen (2005, 2008) discusses stability results for nonlinear VECM and in particular discusses the statistical properties of smooth transition VECM. Krolzig (1997) develops the estimation and testing methodology for Markov-switching VECM.

More recently, Kristensen and Rahbek (2010), Seo (2011) and Kristensen and Rahbek (2013) discuss estimation and testing procedures related to a general class of VECM that allows for a wide range of nonlinear adjustment processes.

After reviewing the above existing studies, we turn to the original research conducted in this thesis. The main part of this thesis comprises of four chapters - each representing a standalone research paper - that can be read independently. The connecting thread is the use of nonlinear cointegration models but each chapter deals with a particular aspect of these approaches. Chapter 2 and Chapter 4 have a theoretical focus and propose extensions of the Engle-Granger framework to capture nonlinear dynamics, whereas Chapter 3 and Chapter 5 employ nonlinear cointegration models to study the commodity market. More precisely:

Chapter 2, **Asymmetric price transmission in the US and German fuel markets: A quantile autoregression approach**, proposes a new econometric model for asymmetric price transmissions. Long-run equilibrium equations between upstream and downstream prices are estimated and quantile autoregression is applied to estimate a quantile-dependent adjustment behavior for lower and upper quantiles of the residual process. We develop a bootstrap cointegration test which is suitable for cointegration relationships that exhibit quantile-dependent adjustment. Furthermore, we introduce the appropriate statistical tests for across-quantile comparisons and overall quantile effects. The methodology is applied to the US and German gasoline and diesel markets. Our empirical results suggest that asymmetries can be found in the early stages of the production chain but are not completely transferred to retail prices.

Chapter 3, **Are gold and silver cointegrated? New evidence from quantile cointegration**, revisits an earlier study on the long-run relationship between gold and silver by Escribano and Granger (1998). We apply a quantile cointegration model to gold and silver prices and to prices of the corresponding futures contracts. Whereas cointegration models, assuming a constant cointegrating vector, fail to detect a cointegration relationship between gold and silver, we are able to show that a nonlinear long-run relationship exits. The cointegrating vector is modelled as state-dependent and time-varying in our framework and the quantile cointegration estimates reveal substantial asymmetry in the relationship. The results suggest that the pronounced role of precious metals as investment opportunities in times of financial turmoil leads to comovement of gold and silver in these periods.

Chapter 4, **Testing for cointegration with SETAR adjustment in the presence of structural breaks**, develops a new cointegration test with SETAR adjustment allowing for the presence of structural breaks in the equilibrium equation. Since the timing of structural breaks is usually unknown, we propose a simple procedure to simultaneously estimate the breakpoint and test the null hypothesis of no cointegration. Thereby, we extend the well-known residual-based cointegration test with regime shift introduced by Gregory and Hansen (1996a) to include SETAR adjustment. We derive the asymptotic distribution of the test statistic and demonstrate its finite-sample performance in a series of Monte Carlo experiments. We find a substantial decrease of power of the conventional cointegration tests with SETAR adjustment caused by a shift in the slope coefficient of the equilibrium equation. The proposed test performs superior in these situations. An application to the ‘rockets and feathers’ hypothesis provides empirical support for this methodology.

Chapter 5, **A Markov regime-switching model of crude oil market integration**, is a joint paper with Konstantin Kuck.[1] This paper revisits the globalization-regionalization hypothesis for the world crude oil market. We examine long-run equilibrium relationships between major crude oil prices – WTI, Brent, Bonny Light, Dubai and Tapis – and focus on the adjustment behaviour following disequilibrium states. We account for a changing adjustment behaviour over time by using a Markov-switching vector error correction model. Our overall findings suggest that the crude oil market is globalized. Dubai turned out to be the only weakly exogenous price in all regimes, indicating its important role as a benchmark price. Furthermore, an interesting finding of our study is that the degree of market integration seems to be connected to global economic uncertainty.

Chapter 6 summarizes the key findings, critically assesses the studies and provides concluding remarks.

[1]This article has been originally published as *Kuck, K. and Schweikert, K. (2017): A Markov regime-switching model of crude oil market integration, Journal of Commodity Markets, 6, 16–31.*

Chapter 2

Asymmetric price transmission in the US and German fuel markets: A quantile autoregression approach

1 Introduction

The relationship between upstream and downstream fuel prices is one of the most commonly studied topics in asymmetric pricing. Starting with Bacon (1991) and Manning (1991), a steadily growing literature has emerged (see, among others (Kaufmann and Laskowski, 2005; Grasso and Manera, 2007; Al-Gudhea *et al.*, 2007; Meyler, 2009; Douglas, 2010; Douglas and Herrera, 2010; Fosten, 2012)), trying to determine whether price decreases in upstream markets are adjusted in downstream markets differently to price increases. Previous empirical studies find mixed evidence for price asymmetries depending on the methodology used, on the country or regional market under investigation and on the stage of the supply chain. Perdiguero-García (2013) conducts a meta-analysis of empirical studies on price asymmetries in the oil market from 1991 until 2011. He finds that the research design contributes substantially to finding asymmetries. Also, the level of competition seems to be a key factor for the existence of asymmetries in the market.

Several concepts of asymmetry in price transmissions are found in the literature (see Meyer and Cramon-Taubadel (2004) for a comprehensive survey on asymmetric pricing and Frey and Manera (2007) for an overview of econometric approaches). The specific type of asymmetry we focus on in this paper is long-run asymmetry, where we investigate the reaction times of a cointegrated system after equilibrium errors. Because the cost function for retail fuel is primarily determined by the price of crude oil, we expect fuel markets to be strongly vertically linked. Hence, upstream and downstream prices are expected to maintain a long-run equilibrium which implies that either the upstream or the downstream prices have to adjust in response to equilibrium errors. In this context, asymmetric pricing refers to a situation in which the rate of price adjustment differs, depending on the size or the sign of the deviation from equilibrium. Long-run asymmetry has a negative effect on consumer welfare if positive equilibrium errors (downstream prices are too high relative to the long-run equilibrium) are not adjusted as quickly as negative equilibrium errors (upstream prices are too high relative to the long-run equilibrium).

Most studies on asymmetric pricing are conducted under a similar framework: A long-run relationship between upstream and downstream prices is estimated by least squares as the first step of the Engle-Granger two-step cointegration procedure. The resulting residual process is separated into two or more regimes and the speed of adjustment in each regime is measured. Significantly different adjustment rates over at least two regimes may be considered as evidence for long-term asymmetry in the cointegrating relationship. The methodological aspects of testing for cointegration with threshold effects have been developed by Enders and Siklos (2001). Although the latter framework is appealing due to its straightforward implementation, it yields contradictory results in a number of studies.[1] These ambiguities may be related to difficulties for the researcher in correctly determining the boundaries of the regimes. Chan (1993) postulates that searching over the set of possible threshold values so as to minimize the sum of squared residuals yields a consistent estimate of the threshold parameter. However, it is possible that multiple local extrema can be found and the global extremum might not necessarily be the only reasonable parameter choice from an economic perspective. Additionally, it is not quite clear how many regimes should be used to quantify the degree of asymmetric pricing. Taking into account the existence of transaction costs, it might be reasonable to model the price adjustment process with three regimes - one regime for small equilibrium errors with weak or insignificant adjustment and one regime for large positive and negative equilibrium errors, respectively. However, the standard literature on threshold cointegration (Enders and Siklos, 2001; Hansen and Seo, 2002) tends to restrict the analysis to only two regimes. Therefore, a certain degree of subjective judgement is involved in all threshold cointegration models.

Typically, the comparison of adjustment rates between regimes is based on a comparison of conditional-means. Because the analysis is restricted to the mean behaviour of the residual process in each regime, specifying the threshold parameter correctly exerts a substantial influence on the outcomes. Consider, for instance, a residual process that exhibits gradually increasing mean-reversion starting with low mean-reversion for negative deviations up to high mean-reversion for positive deviations, i.e. the adjustment rates do not follow a piecewise linear step-function but rather a monotonically increasing continuous function. In this case, the threshold cointegration approach is not able to produce robust results since the aforementioned adjustment process requires a large number of regimes and hence a correspondingly large number of thresholds to be estimated (Honarvar (2010)). Alternatively, the class of smooth transition autoregressive (STAR) models may be used for modelling nonlinear regime-dependent processes (see (Terasvirta, 1994; van Dijk *et al.*, 2002) for an overview). In particular, a logistic transition function could provide an adequate fit for the above described process. However, the recent literature points to severe identification problems associated with STAR models (Ekner and Nejstgaard (2013)).

In line with the majority of papers on the subject, we use Engle-Granger cointegration as a starting point and focus on the mean-reversion of the residual process. But instead of piecewise linear models, we propose a quantile autoregression model. This model expresses the τ-th conditional-quantile function of the response as a linear function of the lagged values of the response. Using quantile autoregression, we are able to analyze different parts of the response distribution and thereby use information that would not be accessible in a conditional-mean paradigm. This is also done without separating the process into sub-processes in a subjective manner. Since the equilibrium error series - obtained as least squares residuals from the cointegrating regression - are centered around zero by construction, a natural interpretation for the conditional-quantiles applies: Lower quantiles correspond to large negative deviations from the long-

[1] Compare for example the results in (Al-Gudhea *et al.*, 2007; Douglas, 2010)

run equilibrium and upper quantiles to large positive deviations. A comparison of quantile-dependent autoregressive coefficients enables us to assess the degree of asymmetry more thoroughly.

We apply this new approach to price relationships in the US and German fuel markets. So far it has not been possible to draw any conclusive statement about whether or not prices are adjusted asymmetrically in these fuel markets. We consider the two major fuel types, gasoline (regular grade for the US market and Euro Super95 for Germany) and diesel, and follow the supply chain disaggregation by Grasso and Manera (2007) to track the price transmission at the different stages of the production chain from crude oil to retail prices. The German fuel market has a distinctly different market structure as compared to the US market hence we seek to provide new insights as to how the potential asymmetries are formed.

This article provides two main contributions. First, we develop a new methodology that is able to model asymmetric price adjustments in a more flexible way. Second, we apply this new methodology to the US and German fuel markets and study the price transmission channel between different stages of the supply chain. Comparing the results for two major fuel markets allows us to draw conclusions on how the different market structures may be related to the potentially different degrees of asymmetric price transmission.

The remainder of the paper is organized as follows. Section 2 summarizes the unique characteristics of the US and German fuel markets. Section 3 outlines the quantile regression methodology by Koenker and Xiao (2006) and discusses its applicability in a cointegration model for asymmetric pricing. In Section 4 , we apply these techniques to assess the degree of asymmetric price transmission in the US and German fuel markets and Section 5 offers a conclusion.

2 A brief description of the US and German fuel markets

Gasoline and diesel play a primary role in transportation and the economy in general. As liquid fuels, they are derived from crude oil in a refinery process, are stored in fuel depots and are finally distributed to local filling stations. To reveal the potentially asymmetric price transmission in the fuel markets, we follow Grasso and Manera (2007) and analyze individual steps of the transmission chain. At the first stage of the production chain, the price transmission occurs from crude oil to ex-refinery prices. We refer to this as the *first stage* or refining stage price transmission. The *second stage* price transmission then occurs when wholesale price changes affect the cost structure for retailers. We refer to this as the *second stage* or distribution stage price transmission. The refined fuel is transported to the filling stations and priced depending on the fuel grade. Additionally, we consider a *single stage* transmission, directly from crude oil prices to retail prices. Concerning the retail price, one has to distinguish between prices that exclude (PTD) and prices that include tax and duty (ITD). Hence, the taxation structure might have an influence on whether price transmissions are asymmetric.

In this study, we examine two fuel markets which are geographically separated and feature distinct market structures. The US fuel market is characterized by a large dependence on gasoline, with 137.8 billion gallons of gasoline consumption in 2010 whereas diesel consumption amounted to only 49.2 billion gallons (US Energy Information Agency (2015)). The share of diesel-engined retail car sales is generally low in the US.[2] The preference for gasoline can in parts be explained by a higher federal excise tax burden on diesel fuel (24.4 cents per gallon) in comparison to gasoline (18.4 cents per gallon). State

[2]The share of diesel cars sales rose to an all-time high with 2.94% in 2009 but then dropped back down to 0.33% in 2012 (U.S. Department of Energy (2013): Transportation Energy Data Book Edition 32)

and local state taxes and fees amount to a national average total of 49.44 cents per gallon for gasoline and 55.41 cents per gallon for diesel (American Petroleum Institute (2017)). Diesel is almost exclusively consumed by professional users (e.g. truck companies, heavy-duty machinery). Approximately 85% of gasoline sold is of regular grade, therefore we do not consider midgrade and premium gasoline.

US refineries mostly use North American crude oil that is considered light and sweet making it a high quality crude. The price for North American crude oil (WTI) is formed in a trading hub in Cushing, Oklahoma. An ex-refinery price can be stated for the West coast (Los Angeles), East Coast (New York Habor) and the Gulf Coast region. The retail price is then derived from a sample of filling stations throughout the country.

Northern and Central European countries utilize primarily crude oils for which the North Sea crude oil Brent serves as a benchmark. The crude oil production is delivered to the Antwerp-Rotterdam-Amsterdam (ARA) oil hub and transported to nearby refineries. For the retail price of fuel we concentrate on Germany as a major automotive market in Europe and analyze the country-specific fuel prices. A Europe-wide analysis would only be feasible as a panel of individual country data (see Grasso and Manera (2007) or Meyler (2009)) since the market structures and taxing schemes vary greatly. European transportation relies much more on diesel-powered engines than the US. Around half of all new passenger cars sold in 2013 were diesel-powered (Eurostat (2017)). Including industrial use, the overall diesel consumption of 31.3 million tons in 2009 was higher than the gasoline consumption of 20.2 million tons (Statista (2010)). The retail fuel tax in Germany is a compound of a fixed mineral oil tax (diesel 47.04 Cent per litre, gasoline 65.45 Cent per litre) and a value added tax applied to both the fuel itself and the mineral oil tax.

3 A quantile-dependent error correction mechanism

The starting point for the empirical analysis of asymmetric price adjustments in this paper is the residual-based cointegration framework developed by Engle and Granger (1987). Two individually integrated time series, y_t and x_t, are said to be cointegrated if they form a linear combination that is stationary. In our empirical application, y_t describes the downstream price and x_t corresponds to the upstream price. In the first step, the long-run equilibrium equation

$$y_t = \beta_0 + \beta_1 x_t + z_t \tag{2.1}$$

is estimated by least squares to obtain the cointegrating vector. In the second step, a stationarity test is applied on the least squared residual series z_t to ascertain whether the latter indeed constitutes a stationary equilibrium error.[3] The ADF-type Engle-Granger cointegration test assesses the significance of the reversion of the residual process towards its mean.

The majority of studies on asymmetric price adjustment focusses on the mean-reversion property of the cointegration residuals. In order to allow for asymmetric adjustment, the residual process is divided into sub-processes at one or more threshold values. Instead, we propose a quantile autoregression model that is able to measure nonlinear effects in the adjustment process using repeated estimation of a linear model. We assume an autoregressive process of order p and use the following linear function (see

[3]The disequilibirum series, although estimated, will be denoted z_t for simplicity.

Koenker and Xiao (2006)) for the residuals z_t,

$$z_t = \mu_0 + \alpha_{1,t} z_{t-1} + \alpha_{2,t} z_{t-2} + \cdots + \alpha_{p,t} z_{t-p} + u_t \tag{2.2}$$

with $\mu_0 = E[\theta_0(U_t)]$, $u_t = \theta_0(U_t) - \mu_0$ and $\alpha_{j,t} = \theta_j(U_t)$ for $j = 1,\ldots,p$. The θ_j's are real-valued functions $[0,1] \to \mathbb{R}$ of standard uniform random variables U_t. The functions are unknown and have to be estimated. u_t is a sequence of independently identical distributed random variables with distribution function $F(\cdot) = \theta_0^{-1}(\cdot + \mu_0)$. The autoregressive coefficients $\alpha_{j,t}$ depend on the quantile $\tau \in [0,1]$ of the error term via the function $\theta_j(U_t)$, allowing them to change from one period to the next.

The residual process z_t is assumed to follow a globally covariance-stationary process under the alternative that is allowed to exhibit some locally persistent or even explosive behavior. However, significant mean-reversion is required in some quantiles to ensure overall stability of the process. Estimation of (2.2) requires solving

$$\min_{\alpha_t \in \mathbb{R}^{p+1}} \left[\sum_{t \in \{t: z_t \geq \mathbf{X}_t \boldsymbol{\alpha_t}\}} \tau |z_t - \mathbf{X}_t \boldsymbol{\alpha_t}| + \sum_{t \in \{t: z_t < \mathbf{X}_t \boldsymbol{\alpha_t}\}} (1-\tau)|z_t - \mathbf{X}_t \boldsymbol{\alpha_t}| \right] \tag{2.3}$$

with $\mathbf{X}_t = (1, z_{t-1}, \ldots, z_{t-p})$ and $\boldsymbol{\alpha_t} = (\mu_0, \alpha_{1,t}, \ldots, \alpha_{p,t})'$ by using linear programming techniques (see (Koenker and d'Orey, 1987; Portnoy and Koenker, 1997)).

The quantile autoregression can equivalently be written in the random-coefficient notation which will be hereafter referred to as the QAR(p) model,

$$z_t = \mu_0 + \rho_t z_{t-1} + \sum_{j=1}^{p} \gamma_{j,t} \Delta z_{t-j} + \varepsilon_t \tag{2.4}$$

where the additional p lags are included to accommodate the dynamics of the process. The analysis continues to focus on the quantile-dependent autoregressive coefficient ρ_t or equivalently the mean-reversion $1-\rho_t$ of the τth conditional-quantile of z_t.[4] Since we are interested in a quantile-dependent error correction mechanism, we apply the QAR(p) model to the least squared residuals resulting from the long-run equation in (2.1). The coefficient ρ_t is estimated for a sequence of quantiles so that the mean-reversion behaviour can be studied for disequilibria of different signs and magnitudes.

3 .1 Testing for cointegration

We test for stationarity of the residual series z_t by applying a modified version of the quantile unit root test developed by Koenker and Xiao (2004). For that purpose, equation (2.4) is estimated for a range of quantiles (in our case $\mathscr{T} = (0.01, 0.02, \ldots, 0.99)$) and the t-statistic for the null hypothesis of no cointegration, $\rho_t(\tau) = 1$, is computed by

$$t_n(\tau) = \frac{\widehat{f(F^{-1}(\tau))}}{\sqrt{\tau(1-\tau)}} (\mathbf{Z_{-1}}'\mathbf{P}_{\Delta}\mathbf{Z_{-1}})^{1/2} (\hat{\rho}_t(\tau) - 1) \tag{2.5}$$

[4]Note that the quantile autoregression should not be estimated in the mean-reversion notation since the application of the nonparametric quantile function on the response Δz_t is not equivalent to the application on the response z_t. The former could be used to model momentum shifts in the adjustment process.

where $\mathbf{Z_{-1}}$ is the vector of the lagged variable z_{t-1} and $\mathbf{P}_\Delta$ is the projection matrix onto the space orthogonal to $\Delta = (1, \Delta z_{t-1}, \ldots, \Delta z_{t-p})'$. $f(\widehat{F^{-1}(\tau)})$ can be written as $f(\widehat{F^{-1}(\tau)}) = (\tau_i - \tau_{i-1})/(\widehat{Q}_{z_t}(\tau_i|\mathbf{X}_t) - \widehat{Q}_{z_t}(\tau_{i-1}|\mathbf{X}_t))$ where $\widehat{Q}_{z_t}(\tau_i|\mathbf{X}_t)$ represents the conditional-quantile of z_t given the information set at point t. The difference quotient, $f(\widehat{F^{-1}(\tau)})$, estimates the conditional density of y_t for some appropriately chosen sequence of τ's. Since the residual process maintains stationarity in the long-run despite the fact that it may display persistence for some quantiles, we use a test statistic that focuses on the overall mean-reversion. For that matter, we employ a quantile Kolmogorov-Smirnov test

$$QKS = \sup |t_n(\tau)| \tag{2.6}$$

for the t-ratios in (2.5). Large values of QKS signal a strong overall mean-reversion behaviour of the residual process and should therefore lead to a rejection of the hypothesis of no cointegration.

The limiting distributions of the individual t-statistics are nonstandard so that we follow Koenker and Xiao (2004) and use a re-sampling procedure for inference based on the QKS statistic. A bootstrap design has to account for the fact that residuals from the cointegrating regression in (2.1) are used. The existing literature on bootstrapping cointegrating regressions points to some difficulties related to nuisance dependencies between the error term and the regressor(s) in the cointegrating regression (see (Li and Maddala, 1997; Chang *et al.*, 2006)). However, bootstrapping cointegrating regressions is mostly used to test linear hypothesis on the cointegrating vector, whereas in our study we seek to test whether the variables are cointegrated with a potentially time-varying mean-reversion behaviour. The error term in (2.1) is not well defined under the null of no cointegration so that a contemporaneous dependence structure between z_t and the x_t variable(s) cannot exist. We therefore propose a modification of the bootstrap unit root test in Koenker and Xiao (2004) in order to make it applicable in cointegration testing. In step (4) of the bootstrap algorithm (see below) the cointegrating regression is re-estimated to mimic the data more closely. The algorithm then proceeds as follows:

(1) Fit the pth order autoregression

$$\Delta z_t = \sum_{j=1}^{p} \eta_j \Delta z_{t-j} + u_t \tag{2.7}$$

by least squares and obtain the parameter estimates $\hat{\eta}_j$ as well as the residuals $\hat{u}_t$.

(2) Draw iid variables u_t^* from the centered residuals $\hat{u}_t$ and generate Δz_t^* using the estimates from the fitted autoregression so that

$$\Delta z_t^* = \sum_{j=1}^{p} \hat{\eta}_j \Delta z_{t-j}^* + u_t^*. \tag{2.8}$$

(3) Generate z_t^* under the null restriction of a unit root

$$z_t^* = z_{t-1}^* + \Delta z_t^* \tag{2.9}$$

with $z_1^* = z_1$.

(4) Regard the exogenous cointegration variables as fixed and generate $y_t^* = \hat{\beta}_0 + \hat{\beta}_1 x_t + z_t^*$. Estimate

$$y_t^* = \beta_0^* + \beta_1^* x_t + z_t^{**} \tag{2.10}$$

by least squares and obtain the residuals z_t^{**}.

(5) Estimate

$$z_t^{**} = \mu_0 + \rho_t z_{t-1}^{**} + \sum_{j=1}^{p} \gamma_{j,t} \Delta z_{t-j}^{**} + \varepsilon_t \tag{2.11}$$

to obtain the bootstrap estimates and test statistics.

The bootstrap estimates for *QKS* allow to construct p-values for the empirically observed statistic. If the QKS test confirms global stationarity of the residuals we assume a long-run cointegrating relationship and proceed with the analysis of the degree of asymmetry in the adjustment path, especially as to how the mean-reversion parameter differs for different signs and sizes of the shock.

3 .2 Testing for quantile effects

Inferential evidence for an asymmetric adjustment behavior is obtained by evaluating the difference in the autoregressive coefficients across quantiles. Least squares residuals are centered around zero by construction so that lower quantiles of z_t refer to large negative and upper quantiles of z_t to large positive deviations. Thus we seek to test the equality of two autoregressive coefficients at the left and right tail of the conditional distribution, for example, according to the null hypothesis $H_0 : \rho_t(\tau_5) = \rho_t(\tau_{95})$ or more generally, we compare a range of coefficients across quantiles with $H_0 : \rho_t(\tau_5) + \cdots + \rho_t(\tau_l) = \rho_t(\tau_u) + \cdots + \rho_t(\tau_{95})$. In both cases, we use a Wald statistic that imposes the corresponding restrictions on the coefficients. The computation of the test statistic requires estimation of the covariance matrix of the estimators.

Cointegration residuals, although covariance-stationary, potentially display a large degree of dependence. Therefore, to account for potentially autocorrelated errors in (2.4), we suggest a block bootstrapping procedure to estimate the covariance matrix.[5] Evaluating the Wald statistic becomes a direct test for asymmetric adjustment in the cointegration relationship.[6]

Furthermore, we are interested in a comparison of the quantile-dependent coefficients with the conditional-mean coefficient. The corresponding null hypothesis of the constancy of the autoregressive coefficient can be formulated as $\rho_t(\tau_i) = \rho_M$ for all $\tau_i \in [\tau_L, \tau_U] = \mathscr{T}$, where ρ_M is the least squares estimate for ρ in (2.4). Following Bera *et al.* (2014), we estimate a sequence of Wald tests with the null hypothesis $\rho_t(\tau_i) = \rho_M$ and compute a Kolmogorov-Smirnov type statistic. The practical application requires an estimate of the joint covariance matrix for the QAR- and AR-parameters. For that purpose we use the above outlined block bootstrap set-up and include the calculation of the least squares estimate for ρ_M. Through resampling we can then calculate the bootstrap variance for ρ_M and subsequently the covariance, $\text{cov}(\rho_M, \rho_t(\tau_i))$ for $L \leq i \leq U$. The Wald statistic is computed for each i. To evaluate the resulting Wald process, we consider the supremum statistic,

$$W_n := \sup_{\tau \in \mathscr{T}} W(\tau), \tag{2.12}$$

[5] We intend to retain the dependence structure of the data by choosing a replication with an average block length of $l = 2m$ where m is the most distant lag that still shows a significant impact in the autocovariance function of z_t (see Politis and Romano (1994)). We use 600 replications of the disequilibrium series z_t to estimate the covariance matrix.

[6] The interpretation of the quantile approach, unfortunately, suffers from subjective decision-making in that we have to determine which across-quantile comparison are most relevant. For the empirical part, we therefore display a battery of Wald tests as well as plots of the estimates of ρ_t to depict the adjustment behaviour as accurately as possible.

where W_n does not follow a standard χ_p^2-distribution. The proposed method in Bera *et al.* (2014) uses an approximation by Davies (1987) that provides an upper boundary for the p-value. The boundary takes the form of

$$\Pr(W_n > u) \leq \Pr(\chi_p^2 > u) + \frac{u^{\frac{p-1}{2}}}{e^{\frac{u}{2}} 2^{\frac{p}{2}} \Gamma(\frac{p}{2})} \int_{\mathscr{T}} E \left| \frac{\partial W^{\frac{1}{2}}(\tau)}{\partial \tau} \right| d\tau \tag{2.13}$$

where p denotes the number of restrictions. Davies (1987) estimates $\int_{\mathscr{T}} E \left| \frac{\partial W^{\frac{1}{2}}(\tau)}{\partial \tau} \right| d\tau$ from the total variation of the Wald process,

$$V = \left| W^{\frac{1}{2}}(\tau_1) - W^{\frac{1}{2}}(\tau_L) \right| + \left| W^{\frac{1}{2}}(\tau_2) - W^{\frac{1}{2}}(\tau_1) \right| + \cdots + \left| W^{\frac{1}{2}}(\tau_U) - W^{\frac{1}{2}}(\tau_k) \right|, \tag{2.14}$$

where $\tau_1, \tau_2, \ldots, \tau_k$ are the turning points of $W^{\frac{1}{2}}(\tau)$ and L and U are the lower and upper bound of τ, respectively.

3 .3 Monte Carlo simulation results

In this section, we use Monte Carlo experiments to examine the properties of the modified QKS test applied to residuals of a cointegrating regression. The Engle-Granger cointegration test based on the ADF statistic and the threshold cointegration test with TAR adjustment serve as benchmarks. We generate series of length $T \in \{100, 500\}$ according to the model

$$\begin{aligned} y_t &= 5 + 2x_t + u_t \qquad u_t = \rho_t u_{t-1} + \vartheta_t \qquad \vartheta_t \sim N(0,1) \\ x_t &= x_{t-1} + \varepsilon_t \qquad\qquad\qquad\qquad\qquad\quad \varepsilon_t \sim N(0,1) \end{aligned} \tag{2.15}$$

to investigate the empirical size and power of the cointegration tests and discard additional 100 observations to randomize initial values. The theoretical justification of the Monte Carlo approach rests on asymptotic results which means that the number of replications, R, should be large for the Monte Carlo experiment to approximate the distribution of a test statistic. However, the QKS test involves a bootstrap procedure and the number of bootstrap replications B are required to be large for the test to be valid. Therefore, a Monte Carlo experiment concerned with bootstrap procedures has to fulfil $B, R \to \infty$. Assuming that the number of bootstrap replications is fixed at $B = 600$, every added Monte Carlo iteration contributes multiplicatively to the overall computational cost. To avoid this inefficiency, we refer to the 'Warp-speed' bootstrap described by Giacomini *et al.* (2013). The authors provide formal results that it is sufficient to use only one bootstrap replication in each Monte Carlo replication. The critical values are then computed from the empirical distribution of the R bootstrap test statistics. We draw $R = 5,000$ replications from (2.15) in each experiment.

Setting $\rho_t = \rho = 1$ gives the empirical size of the tests. We compare the power of the tests according to four different choices of the autoregressive coefficient ρ_t: First, we consider constant adjustment $\rho_t = \rho = 0.9$. Second, we generate data according to threshold autoregressive adjustment

$$\rho_t = \begin{cases} \rho_1 = 0.95 & u_{t-1} \geq 0 \\ \rho_2 = 0.75 & u_{t-1} < 0 \end{cases} \tag{2.16}$$

where negative shocks are adjusted at a faster rate. Finally, we specify a quantile-dependent adjustment behaviour. For that matter, we set $\rho_t = \theta(\vartheta_t) = \min\{c + F(\vartheta_t), 0.95\}$, $c \in \{0.7, 0.8.0.9\}$, where $F(\cdot)$ is the standard-normal cumulative distribution function. The speed of adjustment is inversely related to the magnitude of shocks with an upper boundary of $\rho_t = 0.95$. Furthermore, we use the specification $\rho_t = \tilde{\theta}(\vartheta_t) = \min\{c + F(\vartheta_t), 1\}$, $c \in \{0.5, 0.6.0.7\}$. This specification allows for persistence in case of large positive shocks and moderate mean-reversion for negative shocks.[7]

Table 2.1: Empirical size and power of the cointegration tests.

	$T = 100$				$T = 500$			
ρ_t	*EG*	*TAR*	*QKS*	*QKS**	*EG*	*TAR*	*QKS*	*QKS**
1	0.054	0.055	0.036	0.078	0.050	0.050	0.050	0.069
0.9	0.218	0.229	0.067	-	1	1	0.814	-
0.95								
0.75	0.217	0.246	0.078	-	0.998	1	0.738	-
$\theta(\vartheta_t)$								
$c = 0.7$	0.168	0.178	0.114	-	0.996	0.999	0.995	-
$c = 0.8$	0.119	0.122	0.055	-	0.959	0.961	0.825	-
$c = 0.9$	0.103	0.103	0.048	-	0.879	0.880	0.287	-
$\tilde{\theta}(\vartheta_t)$								
$c = 0.5$	0.359	0.392	0.345	-	1	1	1	-
$c = 0.6$	0.190	0.202	0.242	-	0.987	0.992	1	-
$c = 0.7$	0.101	0.104	0.132	-	0.786	0.786	0.998	-

Note: EG denotes the Engle-Granger test. TAR denotes the threshold cointegration test with TAR adjustment. The quantile unit root test by Koenker and Xiao (2006), *QKS**, without a modification for the use of cointegration residuals is only reported for the size experiment. The *QKS* test is accommodated for small sample sizes, i.e. we estimate the deciles for $T = 100$ instead of percentiles for $T = 500$.

The results are reported in Table 2.1. We find that the modified QKS test is slightly undersized for small sample sizes but has correct size for $T = 500$. The quantile unit root test by Koenker and Xiao (2006) without a modification for the use of cointegration residuals is still oversized for $T = 500$. The QKS test lacks power in situations of constant or TAR adjustment. Changing the autoregressive parameter ρ_t to a quantile-dependent adjustment scheme does not lead to a superior performance of the QKS test compared to the benchmark cointegration tests if a mean-reversion tendency is assured over the whole distribution of shocks. However, the QKS test clearly outperforms the Engle-Granger and threshold cointegration tests if large positive shocks persist.

4 Empirical analysis

Economic theory strongly suggests that a cointegrating relationship between prices of upstream and downstream fuel markets exists since the prices of downstream goods are largely influenced by upstream prices. Meyler (2009) decomposes EU petrol and diesel prices from 2008 and finds that 75% of petrol and

[7] Using the symmetry of the standard-normal distribution, we can easily generate data so that positive shocks are reverted and large negative shocks persist. The autoregressive coefficient ρ_t then follows the function $\theta(\vartheta_t) = \min\{c + F(-\vartheta_t), 1\}$. However, the results are virtually identical.

62% of diesel are accounted for by the crude oil price. The decomposition for the US fuel market shows a similar result with crude oil accountable for 72% of petrol and 61% of diesel prices. It is therefore not unrealistic to assume that crude oil and fuel prices share a common stochastic trend. In what follows, we will first have to test the individual series for their order of integration. After confirmation of their $I(1)$ property, we will estimate the first step of the Engle-Granger cointegration procedure to obtain the equilibrium error series z_t on which we will then apply the above outlined quantile autoregression approach to cointegration.

4 .1 Data, unit root and cointegration tests

Our data cover the period from January 1999 until November 2013 with weekly observations. For the crude oil price we use WTI as a proxy for the North American market and Brent for the European market.[8] Both series are taken from the Federal Reserve Economic Database (FRED) and are converted into cents per litre in their respective currencies. The US ex-refinery price for Los Angeles, New York Habor and the Gulf Coast as well as the retail prices for gasoline and diesel are obtained from DATASTREAM. We use the spot prices at the ARA oil hub for the ex-refinery prices in Europe. Since regular gasoline is rarely used in Europe, we focus on premium gasoline. Gasoil, a prestage for diesel, serves as the proxy for the ex-refinery diesel price. The German gasoline (Super95) and diesel prices with taxes excluded/included (PTD/ITD) are taken from the Weekly Oil Bulletin of the European Commission.

The prices for crude oil and its derivatives experienced a sudden slump during the financial crisis. This break in the series may influence the rejection frequency of unit root tests which do not account for structural breaks and could lead to a false rejection of a unit root. Therefore, we choose the unit root test by Busetti and Harvey (2001) which allows for a structural break in the intercept as well as in the slope coefficient in both the null hypothesis and alternative. The test is based on the KPSS framework which tests for random walk components while assuming (trend-) stationarity with a potential break under the null hypothesis. The results for the Busetti-Harvey (BH) test suggests a unit root in all available series.[9] The differenced time series are deemed stationary in all cases. The results of the unit root tests are depicted in Table 2.2.[10]

Next, we estimate the cointegrating regressions (2.1) for each stage of the price transmission and test for the stationarity of the residual process z_t, using the EG test and the modified QKS test. Since prices of US retail fuel excluding tax and duty are not available, we estimate the *second stage* and *single stage* for the US and German fuel market directly for prices that are observed at the pump. Hence, we use a log-transformation of the prices in these regressions to capture the fact that the mark-up is increasing in costs due to the value added tax.[11] However, this does not allow to isolate the effects of the taxation structure. To further investigate this issue, we compare the results for German ITD prices with the German PTD prices (see Subsection 4 .3). The mark-up for spot fuel prices and retail prices

[8] The properties of different crude oil benchmarks have been discussed in the literature (see Fattouh (2006) for an extensive exposition) WTI and Brent have been chosen since they are the crude oils primarily utilized in US and European refineries, respectively. However, switching the benchmarks or using a third benchmark (Dubai) instead, did not change the qualitative interpretation of our results.

[9] Estimated breakpoints become irrelevant if the null hypothesis is rejected

[10] The unit root test results for log-transformed prices lead to the same test decision but are not reported here to conserve space.

[11] Estimating the cointegration regressions in a linear specification yields qualitatively identical results for the asymmetry patterns.

Table 2.2: Unit root tests of individual price series.

	BH			ADF	
	$\xi(l)$	break		t-stat	lags
Crudes					
Brent	0.204***	02/2002	Δ Brent	−19.50***	1
WTI	0.103***	02/2005	Δ WTI	−20.88***	1
Ex-refinery prices					
Diesel (ARA)	0.206***	02/2004	Δ Diesel (ARA)	−18.87***	1
Gasoline (ARA)	0.199***	06/2011	Δ Gasoline (ARA)	−19.37***	1
Diesel (US)	0.156***	10/2004	Δ Diesel (US)	−18.94***	1
Gasoline (US)	0.125***	09/2004	Δ Gasoline (US)	−20.58***	1
Retail prices					
Diesel (US)	0.154***	10/2004	Δ Diesel (US)	−10.58***	2
Gasoline (US)	0.140***	10/2004	Δ Gasoline (US)	−9.89***	2
Diesel (GER)	0.136***	05/2009	Δ Diesel (GER)	−20.86***	1
Gasoline (GER)	0.100***	02/2009	Δ Gasoline (GER)	−19.51***	1
PTD retail prices					
Diesel (GER)	0.228***	09/2009	Δ Diesel (GER)	−21.06***	1
Gasoline (GER)	0.182***	04/2011	Δ Gasoline (GER)	−19.86***	1

Note: BH denotes the Busetti-Harvey test. The BH test equation includes a constant and a linear time trend. Critical values are 10%: 0.033, 5%: 0.041, 1%: 0.054. The ADF test equation includes a constant. The number of lags is based on the Bayesian Information Criterion (BIC). Critical values are 10%: −2.57, 5%: −2.86, 1%: −3.43.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

excluding tax and duty does not increase in costs, hence we use a linear specification in these instances.[12] The results are presented in Table 2.3.

The cointegration tests indicate an overall mean-reversion behaviour, with the exception of the German diesel spot/diesel ITD relationship where we find evidence for EG cointegration but cannot reject the null hypothesis of no quantile-dependent cointegration. This discrepancy can be explained with the Monte Carlo simulation results in Subsection 3 .3 in which the QKS test has lower power than the EG test if adjustment is symmetrical. We therefore conjecture the residual process z_t to be a globally stationary process which implies a cointegrating price relationship. In the next section, we proceed with the estimation of the quantile autoregressive model and test the resulting quantile-dependent coefficients for their degree of asymmetry.

4 .2 Quantile autoregression results

For the empirical analysis, we apply the QAR(p) model in (2.4) to US fuel market data and German fuel market data. The residuals in both cases originate from the estimates of the long-run equilibrium equation (2.1). We use the modified Barrodale and Roberts algorithm for the quantile regression (Koenker and d'Orey (1987)). The estimated quantile-dependent coefficients are plotted for quantiles between 0.05 and 0.95 (see Figure 2.1 and Figure 2.2). The remaining quantiles are not displayed since solving (2.3) results in increasingly inaccurate estimates for tail quantiles and the overall pattern is already sufficiently revealed by the constrained quantile sequence.

[12]Likewise, a log-specification does not alter the results substantially.

Table 2.3: Estimates of the equilibrium equations and residual-based cointegration tests.

	Intercept	Slope	EG	QKS
First stage				
Diesel^{ARA}	1.096	1.128	-5.377^{***}	10.360^{***}
Gasoline^{ARA}	3.412	1.149	-6.609^{***}	8.603^{***}
Diesel^{US}	−2.967	1.289	-5.095^{***}	6.183^{**}
Gasoline^{US}	0.919	1.130	-6.278^{***}	10.530^{***}
Second stage				
Diesel^{GER}	3.034	0.469	-4.311^{***}	4.949
Gasoline^{GER}	3.467	0.381	-4.769^{***}	6.016^{**}
Diesel^{US}	1.576	0.693	-7.408^{***}	6.907^{**}
Gasoline^{US}	1.591	0.682	-10.040^{***}	9.714^{***}
Single stage				
Diesel^{GER}	3.123	0.464	-4.654^{***}	5.581^{*}
Gasoline^{GER}	3.658	0.352	-5.132^{***}	6.365^{**}
Diesel^{US}	1.474	0.755	-5.384^{***}	7.555^{**}
Gasoline^{US}	1.690	0.681	-5.716^{***}	5.751^{**}

Note: EG denotes the Engle-Granger test. The number of lags is based on the Bayesian Information Criterion (BIC). Critical values are taken from MacKinnon (2010), 10%: −3.05, 5%: −3.35, 1%: −3.91. QKS denotes the modified quantile Kolmogorov-Smirnov test with 600 bootstrap replications.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

We begin with the *first stage* of the price transmission chain in the US market. We restrict the empirical analysis to the Gulf coast prices since the US refinery industry is concentrated in this region.[13] The estimated autoregressive coefficients (ρ_t) for disequilibria series of the diesel/WTI and gasoline/WTI relationships are plotted in the upper panel of Figure 2.1. We observe a upward-sloping curve for the quantile-dependent coefficients in both relationships. The estimated autoregressive coefficient is visibly smaller than one for lower quantiles, corresponding to large negative deviations. This means that disequilibria induced by crude oil prices that are higher in relation to the ex-refinery prices are adjusted relatively fast over time. Conversely, the point estimates for upper quantiles are close to one indicating that adjustment is slow when crude oil prices are too low.

Generally, the speed of pass-through is quite slow (see Table 2.4). The half-life period (50% of pass-through reached) of shocks to the diesel/WTI relationship is 7.9 weeks for negative deviations from the long-run equilibrium (25% quantile) and 44.6 weeks for positive deviations from the long-run equilibrium (75% quantile). 90% of a shock is passed through after 26.2 weeks for negative deviations and 148 weeks for positive deviations. Correspondingly, the half-life period of shocks to the gasoline/WTI relationship is 5.4 weeks for negative deviations and 48.8 weeks for positive deviations while 90% of the shock is passed through after 17.9 weeks for negative deviations and 162 weeks for positive deviations.

Interestingly, the point estimates indicate that extreme positive shocks are not reverted at all. The QAR(p) model in principle allows for a locally persistent or locally explosive behavior of z_t as long as the disequilibrium process is globally mean-stationary. However, in this case the confidence bands

[13]The results for Los Angeles and New York Habor prices display a similar pattern.

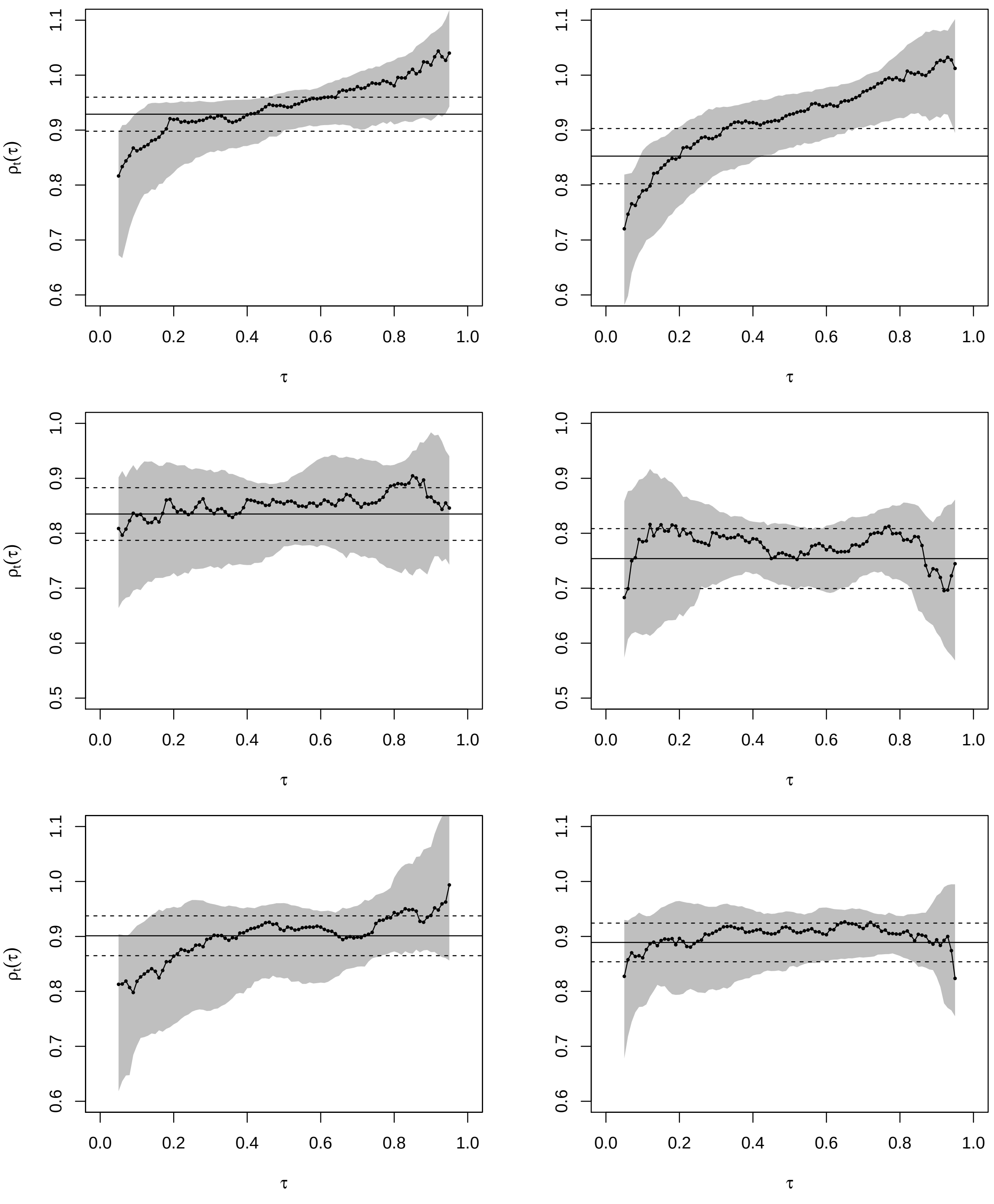


Figure 2.1: Estimation results for the quantile-dependent adjustment coefficient $\rho_t(\tau)$ in the US fuel market. The upper panel, middle panel and lower panel display the *first stage*, *second stage* and *single stage*, respectively. Diesel prices are on the left and gasoline prices are on the right. Shaded areas correspond to a 95% bootstrap confidence interval.

for tail quantiles are relatively wide and include values below one so that we do not find significant statistical evidence for a lack of adjustment. A notion which is supported by the results of the EG and QKS cointegration test rejecting the null of no cointegration for all *first stage* relationships (Table 2.3).

The point estimates for gasoline (right panel) show slightly stronger asymmetric behaviour than the

Table 2.4: Pass-through of long-run equilibrium shocks in weeks.

	lower tail					upper tail				
	50%	60%	70%	80%	90%	50%	60%	70%	80%	90%
First stage										
Diesel^{ARA}	4.5	5.9	7.8	10.4	14.9	253	335	440	588	842
Gasoline^{ARA}	3.6	4.8	6.3	8.4	12.0	10.2	13.5	17.7	23.7	33.9
Diesel^{US}	7.9	10.4	13.7	18.3	26.2	44.6	58.9	77.4	104	148
Gasoline^{US}	5.4	7.1	9.4	12.5	17.9	48.8	64.6	84.8	113	162
Second stage										
Diesel^{GER}	13.9	18.3	24.1	32.2	46.0	7.9	10.5	13.8	18.4	26.3
Gasoline^{GER}	16.4	21.7	28.5	38.2	54.6	6.4	8.5	11.1	14.9	21.3
Diesel^{US}	3.9	5.2	6.8	9.0	12.9	4.4	5.9	7.7	10.3	14.8
Gasoline^{US}	2.9	3.8	5.0	6.7	9.5	3.1	4.1	5.4	7.2	10.3
Single stage										
Diesel^{GER}	8.1	10.7	14.1	18.8	26.9	6.5	8.6	11.3	15.1	21.6
Gasoline^{GER}	9.6	12.7	16.6	22.2	31.8	5.0	6.7	8.8	11.7	16.7
Diesel^{US}	5.2	6.9	9.1	12.1	17.4	8.7	11.5	15.1	20.2	27.0
Gasoline^{US}	6.0	8.0	10.5	14.0	20.0	7.3	9.6	12.7	16.9	24.2

Note: The pass-through durations for the lower tail are based on the 25% conditional-quantile estimations, while the upper tail results are estimated based on the 75% quantile. The durations are computed for the hypothetical case that the quantile-dependent adjustment coefficients stay at the 25% (75%) quantile. It needs to be emphasized that this situation is unrealistic since the coefficients are allowed to change every period.

point estimates for diesel (left panel). The supremum Wald test for equality of conditional-mean and quantile effects, depicted in Table 2.5, signals that the quantile-dependent coefficients are significantly different from the coefficients of the conditional-mean model only for gasoline/WTI. A comparison of the tails of the distribution points towards a strong asymmetry for diesel and gasoline. This is in line with the graphical illustration. The results for the *first stage* suggest that the refinery sector is able to delay the pass-through of price decreases in the US crude oil market, while price increases are passed through at a significantly faster rate.

In the *second stage*, we analyze the transmission from ex-refinery prices to retail prices at the pump. The point estimates, depicted in the middle panel of Figure 2.1, are more concentrated around the baseline conditional-mean value. The conditional-mean estimates indicate that shocks are passed through faster in the gasoline market than in the diesel market. The half-life period of shocks to the diesel/ex-refinery relationship is 3.9 weeks for negative deviations and 4.4 weeks for positive deviations. The half-life period of shocks to the gasoline/ex-refinery relationship is 2.9 weeks for negative deviations and 3.1 weeks for positive deviations. The supremum Wald test supports the hypothesis that the quantile effects are not statistically different from the conditional-mean effect and a comparison at the tails indicates no asymmetries.

In the *single stage* transmission process, we find a slightly upward-sloping curve for the diesel/WTI

Table 2.5: Supremum Wald test for equality of mean and quantile effects and single Wald tests for equality of the autoregressive coefficients across quantiles.

	W_n	$W(\tau_{15}=\tau_{85})$	$W(\tau_{10}=\tau_{90})$	$W(\tau_5=\tau_{95})$	$W(R1)$	$W(R2)$
First stage						
Diesel^{ARA}	17.00***	16.07***	15.34***	12.55***	14.17***	15.54***
Gasoline^{ARA}	10.36**	7.46***	6.99***	9.54***	8.63***	8.37***
Diesel^{US}	6.58	5.48**	5.87**	5.34**	6.47**	6.80***
Gasoline^{US}	18.37***	7.99***	12.44***	13.72***	17.07***	14.30***
Second stage						
Diesel^{GER}	2.27	0.64	1.72	0.59	0.95	1.05
Gasoline^{GER}	6.30	2.10	2.63	1.94	2.68	2.73*
Diesel^{US}	1.89	0.84	0.14	0.16	0.18	0.31
Gasoline^{US}	3.97	0.07	0.28	0.27	0.04	0.17
Single stage						
Diesel^{GER}	3.55	0.00	0.16	0.00	0.07	0.03
Gasoline^{GER}	6.73	0.88	1.56	6.18**	3.11*	2.14
Diesel^{US}	4.54	1.94	1.40	1.94	1.43	1.43
Gasoline^{US}	3.59	0.06	0.29	0.00	0.06	0.08

Note: W_n denotes the supremum Wald test for equality of mean and quantile effects with null hypothesis $\rho_M = \rho_t(\tau_5) = \rho_t(\tau_6) = \cdots = \rho_t(\tau_{95})$. The Wald tests $W(\tau_{15}=\tau_{85})$, $W(\tau_{10}=\tau_{90})$ and $W(\tau_5=\tau_{95})$ test the null hypothesis $\rho_t(\tau_{15}) = \rho_t(\tau_{85})$, $\rho_t(\tau_{10}) = \rho_t(\tau_{90})$ and $\rho_t(\tau_5) = \rho_t(\tau_{95})$, respectively. $W(R1)$ corresponds to a Wald test under the hypothesis $\rho_t(\tau_5)+\cdots+\rho_t(\tau_9) = \rho_t(\tau_{91})+\cdots+\rho_t(\tau_{95})$ and $W(R2)$ to a Wald test under the hypothesis $\rho_t(\tau_5)+\cdots+\rho_t(\tau_{14}) = \rho_t(\tau_{86})+\cdots+\rho_t(\tau_{95})$.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

relationship while the estimated quantile-dependent adjustment coefficients for the gasoline/WTI relationship coincide with the conditional-mean estimate (lower panel in Figure 2.1). The supremum Wald test for equality and the asymmetry tests do not reveal any asymmetries at reasonable significance levels. As expected, the speed of adjustment is slower than in the *second stage*. 50% (90%) of a shock to the diesel/WTI relationship is adjusted after 5.2 (17.4) weeks for negative deviations and 8.7 (28.9) weeks for positive deviations, while 50% (90%) of a shock to the gasoline/ex-refinery relationship is 6.0 (20.0) weeks for negative deviations and 7.3 (24.2) weeks for positive deviations.

We now turn to the German fuel markets. The quantile-dependent adjustment coefficients in the *first stage* transmission are depicted in the upper panel of Figure 2.2 and show a similar pattern compared to their US counterparts. Gasoil and premium gasoline at the ARA hub display a steep upward-directed slope. Since the null hypothesis of equality of the conditional-mean coefficient and all quantile-dependent coefficients is rejected, we find significant quantile effects. Also, the across quantiles comparison are highly significant. The half life of shocks to the gasoil/Brent relationship is 4.5 weeks for negative deviations and 254 weeks for positive deviations. This means that large positive deviations are not effectively adjusted by the system. Premium gasoline is adjusted at a faster rate so that we estimate the half life of shocks to the premium gasoline/Brent relationship to be 3.6 weeks for negative deviations and 10.2 weeks for positive deviations.

A possible source for the strong signs of asymmetry in the *first stage* in Europe and the US might be the fact that the oil refinery market has a relatively small number of competitors due to the capital-

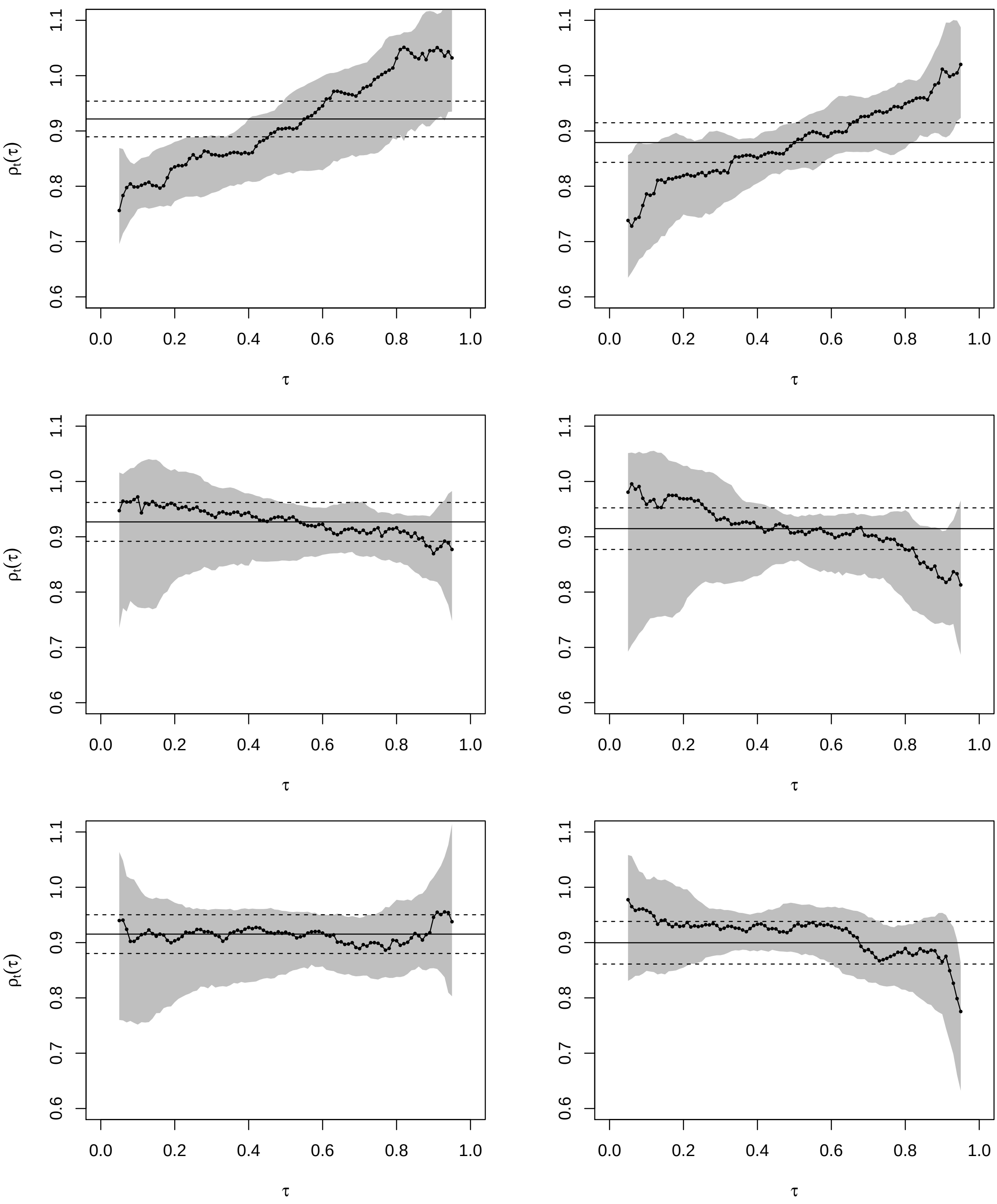


Figure 2.2: Estimation results for the quantile-dependent adjustment coefficient $\rho_t(\tau)$ in the German fuel market. The upper panel, middle panel and lower panel display the *first stage*, *second stage* and *single stage*, respectively. Diesel prices are on the left and gasoline prices are on the right. Shaded areas correspond to a 95% bootstrap confidence interval.

intensive nature of this industry. In 2013, the refining capacity of the US was spread across 57 refinery companies operating 139 refineries (US Energy Information Agency (2013)), while 106 refineries were operated in Europe (FuelsEurope (2014)). Large vertically integrated operations which are involved in several upstream activities might also reduce competition. Additionally, the price formation process in the crude oil and fuel spot markets is unusual. The product is sold in large quantities and trading in the ex-refinery petroleum market depends highly on the benchmark prices provided by price reporting agencies (PRA). Platts, the leading PRA, collects prices by a window or market-on-close process (MOC) in which bids, offers and the trade volume are assessed and prices are published as an end-of-day value. The system has been harshly criticized lately since it rests on voluntary and selective disclosure as well as subjective judgement of the PRA. Even without proclaiming intentional manipulation or collusive action, the MOC price formation is far from a full information pricing and opens up opportunities for delayed price reactions.

The *second stage* transmission does not reveal significant asymmetries. The quantile-dependent adjustment coefficients are depicted in the middle panel of Figure 2.2. The point estimates suggest that, in contrast to expectations, negative deviations are adjusted at a faster rate which corresponds to a situation in which the customers experience an immediate retail price decrease caused by lower crude oil prices, but price increases are delayed. However, the null hypothesis of the quantile effects and asymmetry tests cannot be rejected. The conditional-mean adjustment rates of Super95 and diesel are very similar while the conditional-quantile curve is slightly steeper for Super95. The half life of shocks to the diesel/gasoil relationship is 13.9 weeks for negative deviations and 7.9 weeks for positive deviations. The half life of shocks to the Super95/premium gasoline relationship is 16.4 weeks for negative deviations and 6.4 weeks for positive deviations.

The results for the *single stage* are depicted in the lower panel of Figure 2.2 and whereas the curve is almost flat for diesel, we find a slightly downward-sloping curve for Super95. Equality across quantiles can be rejected only for extreme quantiles. 50% (90%) of a shock to the diesel/Brent relationship is adjusted after 8.1 (26.9) weeks for negative deviations and 6.5 (21.6) weeks for positive deviations while 50% (90%) of a shock to the Super95/Brent relationship is adjusted after 9.6 (31.8) weeks for negative deviations and 5.0 (16.7) weeks for positive deviations.

The results in this section are robust to a sample split at the time of the financial crisis. Furthermore, we find only minor violations of the monotonicity requirement on the conditional-quantile functions (see (Koenker and Xiao, 2006; Chernozhukov *et al.*, 2010)).

4 .3 Effects of the taxation structure on fuel price transmissions

In contrast to the US market, fuel prices excluding tax and duty are available for the German market. Hence, we are now able to investigate whether the tax structure masks any asymmetries in the distribution stages. Greenwood-Nimmo and Shin (2013) study fuel price adjustments in the UK and find that the tax structure masks asymmetries at the pump. However, the UK uses an escalator type fuel duty policy which is different from the fixed sum mineral oil tax in Germany. It is therefore of interest to find out whether the same difference between PTD and ITD prices exist in the German fuel market.

The cointegration equation for PTD prices is estimated in a linear specification and the results for the *second stage* and *single stage* are displayed in the upper and lower panel of Figure 2.3, respectively. The results reveal differences in PTD and ITD prices. Prices before tax and duty are adjusted at a faster rate than prices at the pump. In case of diesel, the half life of shocks in the *second stage* (*single stage*) is 3.1 (2.5) weeks for negative deviations and 1.8 (4.2) weeks for positive deviations. For Super95, the half life of shocks in the *second stage* (*single stage*) is 2.0 (5.0) weeks for negative deviations and 1.6 (3.6) weeks for positive deviations.

In the *second stage*, we find downward-sloping conditional-quantile curves for diesel and Super95, but only the differences across quantiles for Super95 are statistically significant. The reaction to increases in production costs and subsequent adjustment of retail prices seem more difficult in the Super95 market. A higher price elasticity of demand for gasoline could imply that customers postpone refuelling their cars when they use them for expendable activities or they switch to alternative modes of transportation. This pattern is not found in the *single stage* where the conditional-quantile curve is again upward-sloping for diesel and almost flat for Super95. Although it seems that the tax structure in Germany slows down the adjustment rates, we find no evidence that it allows retailers to delay prices decreases.

5 Conclusion

The quantile autoregression approach to asymmetric pricing in the US and German fuel markets leads to new insights about the pricing mechanisms. Using quantile regression techniques, we are able to quantify the degree of asymmetric price transmission without explicitly specifying distinct regimes and estimating the associated threshold values, or without specifying a particular parametric smooth transition framework. Therefore, the estimations are free of subjectivity and the employed model is parsimonious in nature. Applying this methodology to two large, geographically separated fuel markets, we are able to relate potential similarities or differences in the empirical findings to the specific structures of the two markets.

Our results highlight the importance of separating the price transmission chain in individual steps. The price transmission at the *second stage* and *single stage* turn out to be mostly symmetric, while we find evidence for a strong degree of asymmetry in the *first stage* of both markets. This finding might be related to indirect price discovery through a price reporting agency. Furthermore, the literature points to oligopolistic structures and the storage capacity to have some influence on the price transmission process from crude oil prices to the fuel spot markets (Bacon, 1991; Manning, 1991; Kaufmann and Laskowski, 2005). However, we are not able to identify the source of asymmetry in this paper and leave this open for further research.

Interestingly, the asymmetries vanish when we turn to the direct adjustment from crude oil to prices at the pump. This is a surprising result considering that the meta-analysis by Perdiguero-García (2013) reports a greater likelihood of price asymmetries for the retail price segment. A contributing factor might be the fact that we use prices at the pump which include tax and duty. Further analysis of German fuel prices excluding tax and duty reveals a more rapid pass-through. The design of the tax structure seems to contribute to slower reaction times of fuel prices to oil price changes. In terms of asymmetric adjustment behaviour, the retail fuel prices in Germany show a pattern which contradicts the widespread perceptions. Indeed, not the decreases in fuel spot prices are adjusted at a slower rate but rather the increases appear to be delayed. This has a positive effect on customer welfare and signals a highly competitive fuel market. However, the differences in pass-through are only statistically significant for retail gasoline prices.

For the US retail fuel market, we find no statistically significant asymmetry in both gasoline and diesel. This has to be considered a surprising result in the context of previous studies that argue for market power as a possible explanation for empirically observed asymmetric adjustments (Fosten (2012) and Perdiguero-García (2013)). Although the smaller diesel demand side consists almost exclusively of professional users and small-scale enterprises which are usually not able to delay their purchase in times of increasing fuel prices, we find no evidence that retailers are able to exploit the market structure.

In summary, it can be stated that fuel spot prices are asymmetrically adjusted to crude oil prices both in Europe and the US but we find no convincing evidence that those asymmetries are passed on to the retail fuel markets.

A Appendix

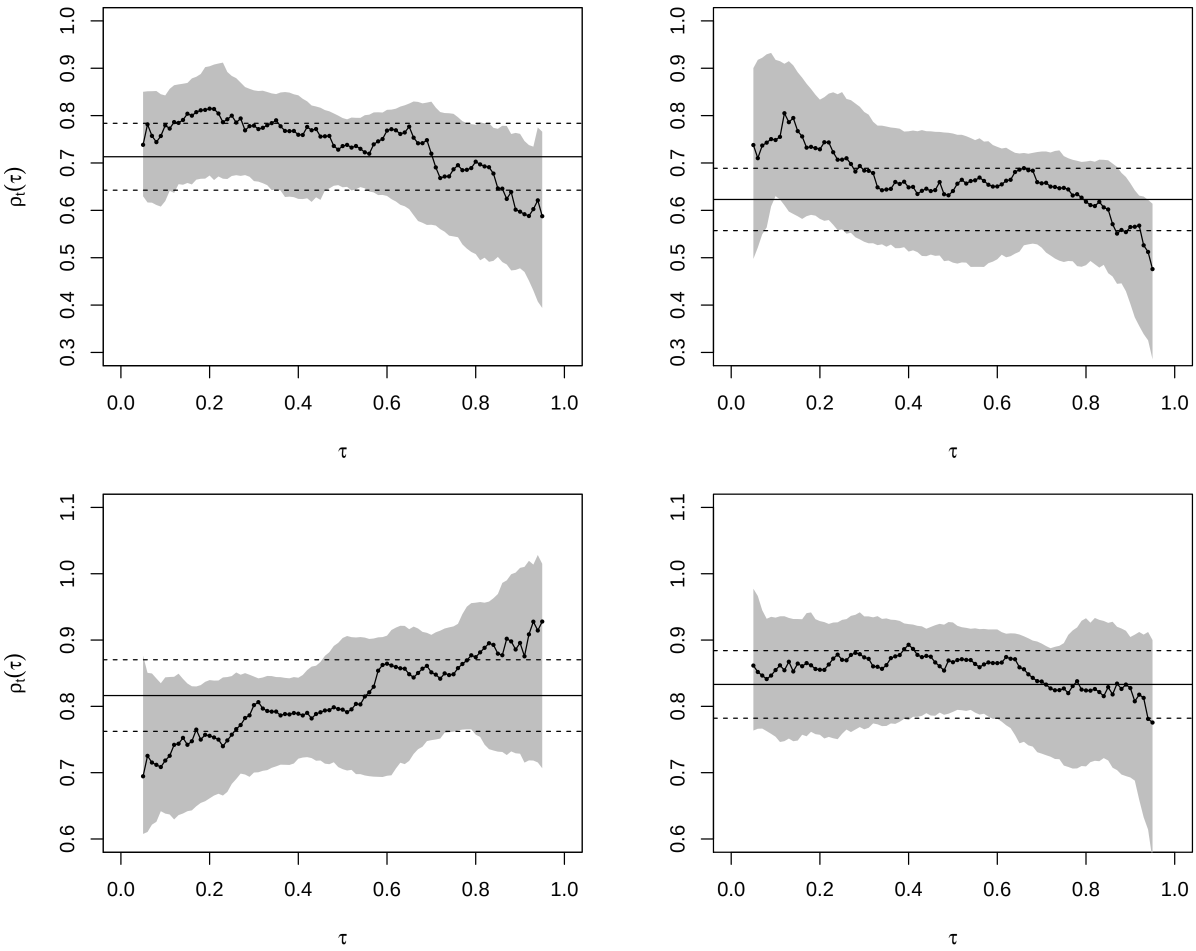


Figure 2.3: Estimation results for the quantile-dependent adjustment coefficient $\rho_t(\tau)$ in the German fuel market (excluding tax and duty). The upper panel and lower panel display the *second stage* and *single stage*, respectively. Diesel prices are on the left and gasoline prices are on the right. Shaded areas correspond to a 95% bootstrap confidence interval.

Table 2.6: Additional estimations and tests for German fuel prices excluding tax and duty.

	Intercept	Slope	EG	QKS	W_n	$W(\tau_{15}=\tau_{85})$	$W(\tau_{10}=\tau_{90})$	$W(\tau_5=\tau_{95})$	$W(R1)$	$W(R2)$
Second stage										
Diesel^{GER}	7.774	1.098	-8.672^{***}	6.970^{**}	11.59^{**}	4.15^{**}	4.19^{**}	4.55^{**}	4.29^{**}	6.13^{**}
Gasoline^{GER}	7.232	0.945	-9.413^{***}	9.252^{***}	5.05	3.58^{*}	4.14^{**}	2.02	3.08^{*}	3.66^{*}
Single stage										
Diesel^{GER}	8.874	1.242	-6.997^{***}	8.365^{**}	2.68	0.18	0.10	0.57	0.29	0.22
Gasoline^{GER}	10.303	1.092	-6.554^{***}	7.577^{**}	6.24	2.53	2.63	3.77^{*}	3.03^{*}	2.93^{*}

Pass-through of long-run equilibrium shocks in weeks

	lower tail					upper tail				
	50%	60%	70%	80%	90%	50%	60%	70%	80%	90%
Second stage										
Diesel^{GER}	3.1	4.1	5.4	7.2	10.3	1.8	2.4	3.2	4.3	6.1
Gasoline^{GER}	2.0	2.6	3.5	4.6	6.6	1.6	2.1	2.8	3.7	5.3
Single stage										
Diesel^{GER}	2.5	3.3	4.3	5.8	8.3	4.2	5.6	7.3	9.8	14.0
Gasoline^{GER}	5.0	6.6	8.7	11.6	16.7	3.6	4.8	6.3	8.5	12.1

Note: EG denotes the Engle-Granger test. The number of lags is based on the Bayesian Information Criterion (BIC). Critical values are taken from MacKinnon (2010), 10%: −3.05, 5%: −3.35, 1%: −3.91. QKS denotes the modified quantile Kolmogorov-Smirnov test with 600 bootstrap replications. W_n denotes the supremum Wald test for equality of mean and quantile effects with null hypothesis $\rho_M = \rho_t(\tau_5) = \rho_t(\tau_6) = \cdots = \rho_t(\tau_{95})$. The Wald tests $W(\tau_{15}=\tau_{85})$, $W(\tau_{10}=\tau_{90})$ and $W(\tau_5=\tau_{95})$ test the null hypothesis $\rho_t(\tau_{15}) = \rho_t(\tau_{85})$, $\rho_t(\tau_{10}) = \rho_t(\tau_{90})$ and $\rho_t(\tau_5) = \rho_t(\tau_{95})$, respectively. $W(R1)$ corresponds to a Wald test under the hypothesis $\rho_t(\tau_5)+\cdots+\rho_t(\tau_9) = \rho_t(\tau_{91})+\cdots+\rho_t(\tau_{95})$ and $W(R2)$ to a Wald test under the hypothesis $\rho_t(\tau_5)+\cdots+\rho_t(\tau_{14}) = \rho_t(\tau_{86})+\cdots+\rho_t(\tau_{95})$. The pass-through durations for the lower tail are based on the 25% quantile, while the upper tail results are estimated based on the 75% quantile. The durations are computed for the hypothetical case that the quantile-dependent adjustment coefficients stay at the 25% (75%) quantile. It needs to be emphasized that this situation is unrealistic since the coefficients are allowed to change every period.
$^{***}\ p < 0.01$, $^{**}\ p < 0.05$, $^{*}\ p < 0.1$

Chapter 3

Are gold and silver cointegrated? New evidence from quantile cointegration

1 Introduction

Gold and silver share a long-standing relationship that goes back to the first issuance of gold and silver coins that were used as currency. The monetary use of gold and silver was facilitated by their unique characteristics. They are rare, easily transportable, malleable and do not corrode so that they serve as a perfect store of value. The monetary system of, for example, Germany was backed by silver until 1873 and the gold-backed Bretton Woods system de facto ended in 1971 with the change to a system of national fiat monies. Subsequently, the relationship between gold and silver changed drastically with the transformation from commodity money to fiat money.

Although precious metals are still seen as stores of value, their commercial uses have gained importance. Gold is used, among others, in restorative dentistry and, since it is highly conductive, for high quality electrical connectors. Silver is the most reflective known metal and therefore used in photography, optics, as well as the solar energy industry. Both metals are also used in jewellery (demand for jewellery accounted for around 50 per cent of world gold demand and 20 percent of global silver demand in 2014[1]).

Gold and silver also play a prominent role as investments. In times of financial turmoil which are characterized by rapidly decreasing values of stock indices, the prices of precious metals tend to move in the opposite direction. Investors are interested in assets which are uncorrelated or ideally negatively correlated with the general market developments to hedge against adverse financial events. Evidence for a safe haven role of gold has recently been found by Baur and Lucey (2010) and Baur and McDermott (2010). Agyei-Ampomah *et al.* (2014) report that other precious metals, including silver, may present even better investment alternatives than gold in financial crises periods.

It is of considerable interest to market participants to know whether a long-run relationship between gold and silver prices exists for the following reasons: First, the knowledge of the dependence between prices may be used for forecasting purposes. Maintaining an equilibrium relationship over an extended period of time implies that at least one variable adjusts to disequilibrium states. The adjustment behaviour can then help to predict future returns of the adjusting variable(s). Second, a cointegrated gold and

[1] The estimates are taken from the World Gold Survey 2016 and World Silver Survey 2016 (GFMS (2016a) and GFMS (2016b)).

silver portfolio would be a suitable long-term hedge and could qualify for a market-neutral pairs trading strategy (Alexander (1999)).[2] Third, as gold and silver are seen as substitutes to reduce similar types of risks in portfolios (Ciner (2001)), finding evidence of cointegration provides statistical support that gold and silver follow a common stochastic trend. Fourth, additional information about the trajectory of gold prices might reduce uncertainty for central banks and other major institutions.

The question of whether gold and silver are cointegrated has already drawn some attention in the literature: Escribano and Granger (1998) investigate the relationship between gold and silver prices after the collapse of the Bretton Woods system. They use monthly data from 1971 to 1990 and investigate a cointegration relationship between gold and silver prices. However, they have to pre-specify regimes in order to find evidence for cointegration and the null hypothesis of no cointegration cannot be rejected for the full sample. They argue that the cointegration relationship only holds for the well-known Hunt brothers episode ('silver bubble') from June 1979 to March 1980 and the post-bubble period in the 1980s, but markets begin to separate at the end of their sample. The authors encourage further research to focus on the potential nonlinearity in the data, particularly on the time-varying dependence between the prices. Ciner (2001) responds to the claim of a long-run relationship between gold and silver and uses daily closing prices of gold and silver futures contracts traded on the Tokyo Commodity Exchange (TOCOM) to verify whether markets indeed became separate. The results do not support a stable long-run relationship between gold and silver futures for the period from 1992 to 1998. Lucey and Tully (2006) use a dynamic cointegration approach which involves a recursive or rolling window estimation and identify periods of weak and strong dependence. They use a sample of Friday closing prices from 1978 to 2002 for their analysis and conclude that overall a cointegration relationship has been maintained. Baur and Tran (2014) revisit the dataset used by Escribano and Granger (1998) and expand the time period to July 2011. They find evidence for cointegration in the full sample but the results suggest that the cointegrating vector changes during bubble and crisis periods. They conclude that the long-run relationship between gold and silver is not stable. The results point to a comovement only in episodes of financial stress in which the store of value aspect of precious metals is particularly important.

Potential nonlinearity in the long-run relationship between gold and silver has so far been treated either as a structural break in the cointegrating vector or as a recursive/rolling window estimation to identify periods of stronger and weaker dependence. On the one hand, a division of the sample period into subperiods requires that dummy variables have to be specified arbitrarily. On the other hand, an application of dynamic cointegration models might identify a number of subperiods with strong dependence but estimation requires specifying the appropriate length of the estimation window which influences the result. Moreover, nonlinearities cannot be quantified.

In this paper, we propose a quantile cointegration approach which enables to model a state-dependent and time-varying cointegrating vector. The values of the cointegrating vector may vary over the innovation quantile. Thereby, the degree of comovement between gold and silver does not depend on prevailing market conditions but rather on the state of the individual prices. Specifically, this allows to measure the

[2] In pairs trading, two or more assets are identified that share similar characteristics and for which prices should be similar, i.e. they hold a long-run relationship. Then if the relative pricing between the assets indicates a mispricing, the trading strategy consists of buying the lower-priced asset and selling the higher-priced asset leading to a statistical arbitrage in the short-run. However, it is assumed that the mispricings will be corrected in the long-run. Prices are usually modelled as a random walk so that a cointegration analysis has to be employed to capture the long-run relationship between prices. If evidence for a cointegration relationship between the assets can be established, the disequilibrium series is mean-reverting and mispricings have to be corrected to maintain the long-run equilibrium.

response of silver prices to gold prices, if silver prices are high and vice versa. The effects of financial turmoil on the prices is implicitly modelled since prices of precious metals tend to increase in financial crisis periods and thereby the state of the prices is altered. To determine whether gold and silver are cointegrated under the quantile cointegration framework, we use a cointegration test developed by Xiao (2009). This test is based on the CUSUM testing principle and, contrary to conventional unit root tests applied in the Engle-Granger framework, tests the null hypothesis of cointegration by measuring the fluctuation in the residuals. If the null hypothesis of the quantile cointegration test is not rejected, it is possible to test for constancy of the cointegrating vector over a range of quantiles to quantify the degree of nonlinearity in the long-run relationship.

This paper contributes to the empirical literature by modelling the state- and time-dependence of the long-run relationship between gold and silver prices and attempts to explain why gold and silver move together in the long-run. First, we revisit an extended gold and silver dataset in a monthly frequency to allow a comparison to the Escribano and Granger (1998) and Baur and Tran (2014) studies. Furthermore, we also conduct the analysis using observations at a daily frequency as well as using prices of futures contracts from 1980 to 2014 to examine the robustness of our results to different frequencies and whether our results are driven by unique characteristics of the spot market. We are able to reveal an asymmetric pattern in the monthly spot prices relationship characterized by a stronger response of silver prices to gold prices when silver prices are high and of gold prices to silver prices when gold prices are high.

The remainder of the paper is organized as follows: Section 2 discusses economic reasons why gold and silver might share a common stochastic trend, Section 3 introduces a CUSUM test for linear cointegration models and describes the quantile cointegration methodology by Xiao (2009). In Section 4 , we apply these techniques to the gold and silver relationship and Section 5 concludes on our results.

2 Why should gold and silver share a common stochastic trend?

Although gold and silver possess similar characteristics, their differing commercial uses suggest that their markets are separated and hence no long-run relationship between them exists. Granger (1986) states that prices generated on a jointly efficient, speculative market cannot be cointegrated since this would violate the efficient market hypothesis. However, the findings on whether gold and silver markets are efficient are mixed. For example, Smith (2002) investigates London gold prices and finds autocorrelated returns of the twice-daily fixing prices, speaking against the random walk hypothesis. The closing prices, by contrast, are generated randomly. Pierdzioch *et al.* (2014) account for transaction costs and show that a trading rule which incorporates publicly available information does not outperform a buy-and-hold strategy, implying that the gold market is informationally efficient. Ntim *et al.* (2015) extend their analysis of gold price efficiency to different markets. They report a higher probability of rejecting the weak-form efficiency in emerging gold markets than developed ones. Charles *et al.* (2015) find that return predictability of precious metals markets has been changing over time. Gold seems to have a higher degree of market efficiency over silver and platinum.

The exact mechanisms of the price formation of precious metals prices is still little understood. Precious metals are seen both as a commodity as well as a financial asset. While financial asset returns are strongly correlated with macroeconomic indicators and each other, commodity returns are typically less correlated with financial assets returns and returns of other commodities (Tang and Xiong (2012)).

As a distinctive feature of precious metals, and in contrast to other commodities like crude oil, the price is largely unaffected by annual production since its life span is practically infinite and stockpile outweighs annual production. The price formation is therefore determined on the demand side. The annual production of gold and silver is depicted in Figure 3.1 and Figure 3.2 shows a decomposition of gold and silver demand. In 2014, around 10% of total gold demand and 50% of total silver demand was

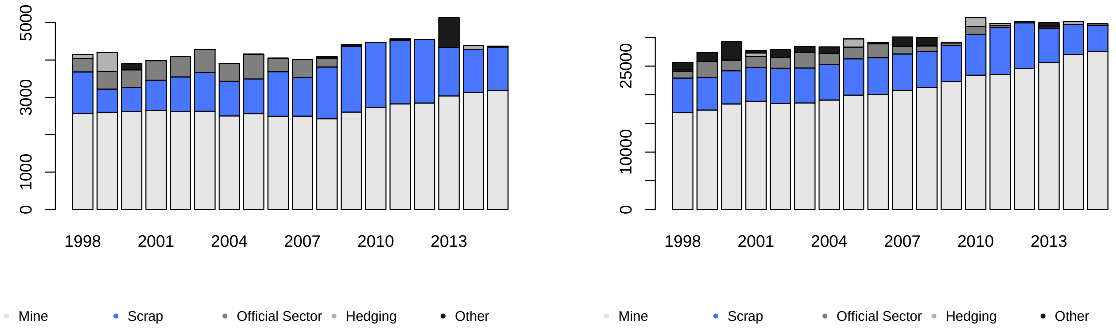


Figure 3.1: Supply of gold (right) and silver (left) in tonnes. Data taken from the GFMS gold and silver surveys (1998 - 2015).

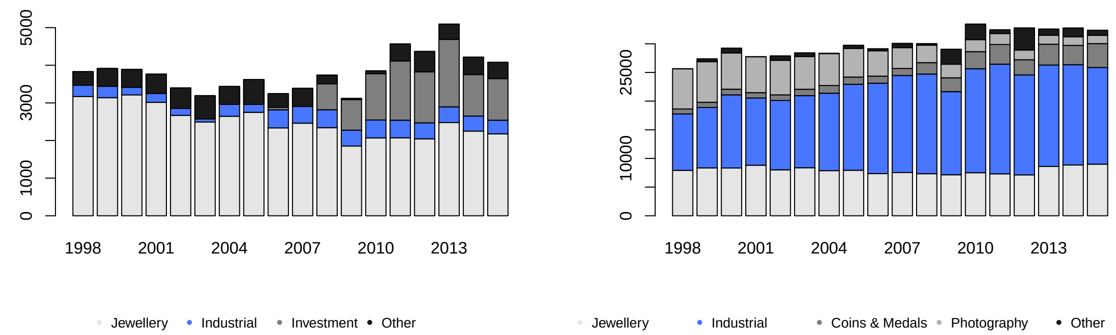


Figure 3.2: Demand for gold (right) and silver (left) in tonnes. Data taken from the GFMS gold and silver surveys (1998 - 2015).

attributed to industrial fabrication. Taking into consideration that jewellery items are often seen as stores of value, gold seems to be mainly used as a cash-like asset, while silver prices are determined largely by industrial demand. Nevertheless, gold and silver show a visible comovement in historical price series (see Figure 3.3).

A closer inspection of the time series plot reveals that gold and silver boom and bust during the same time periods. However, the behaviour in tranquil times is far less synchronized. The long-run relationship, if it exists, might be characterized by episodes of stronger and weaker dependence. Although gold and silver are no industrial substitutes, their use on financial markets, especially as a safe haven asset in crisis periods, could translate to periods in which the store of value aspect of gold and silver is pronounced and might be the reason why the individual prices follow a similar trajectory.

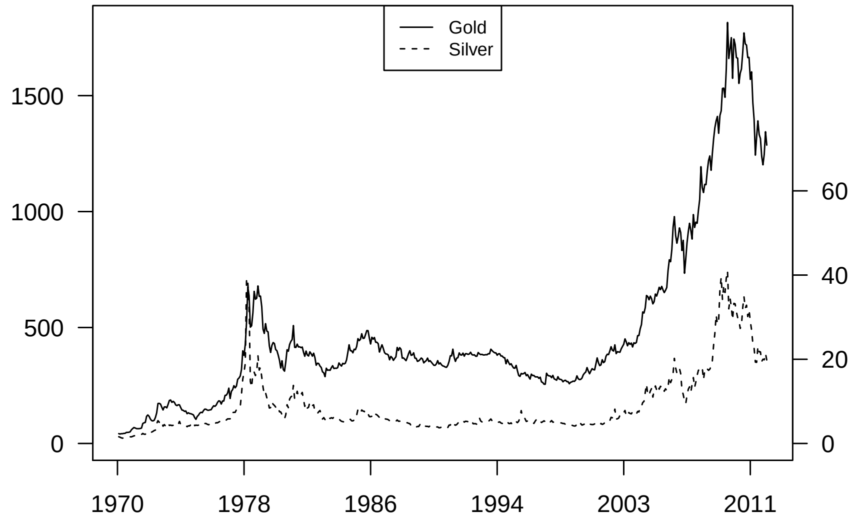


Figure 3.3: Historic gold and silver spot prices. The left (right) axis describes the gold (silver) prices in USD.

3 Econometric framework

The quantile cointegration model builds on the idea of the residual-based cointegration approach proposed in Engle and Granger (1987). The long-run equilibrium equation is specified as

$$y_t = \alpha + \beta' x_t + u_t, \tag{3.1}$$

where x_t is a k-dimensional vector of $I(1)$ variables. For the long-run equilibrium to hold u_t must be mean zero stationary. While the parameters $\theta = (\alpha, \beta)$ are usually estimated using least squares in the linear cointegration model, we estimate θ using quantile regression. Thereby, the coefficients are thought to be varying over time. In particular, the value of the coefficients may vary over the innovation quantile. Hence, the quantile cointegration model may be viewed as a stochastic cointegration model with strongly dependent coefficients (Xiao (2009)). The quantile regression estimator $\widehat{\theta}(\tau) = (\widehat{\alpha}(\tau), \widehat{\beta}(\tau))'$ for each quantile $\tau \in \mathscr{T}$ is obtained by solving

$$\widehat{\theta}(\tau) = \arg \min_{\theta \in \mathbb{R}^{k+1}} \sum_{t=1}^{n} \rho_\tau(y_t - \alpha(\tau) - \beta(\tau)' x_t), \tag{3.2}$$

where $\rho_\tau(u) = u(\tau - \mathbb{1}\{u < 0\})$ is the asymmetric weights function as in Koenker and Bassett (1978), $\mathbb{1}\{\cdot\}$ is a Heaviside indicator function and n is the sample size. In contrast to least squares estimation, where the conditional expected value of y_t is expressed as a function of the variables x_t, in quantile regression the τth quantile of y_t conditional on the information set $\mathscr{F}_t$ in period t is estimated,

$$\widehat{Q}_{y_t}(\tau | \mathscr{F}_t) = \widehat{\alpha}(\tau) + \widehat{\beta}(\tau) x_t + F_u^{-1}(\tau). \tag{3.3}$$

The residual weights are computed as $\psi_\tau(u) = \tau - \mathbb{1}\{u < 0\}$ and the τth residual series as $u_{t\tau} = y_t - \widehat{\alpha}(\tau) - \widehat{\beta}(\tau)'x_t$.

Estimating equation (3.1) using a pair of potentially cointegrated variables $(y_t, x_t)'$ introduces an endogeneity problem if x_t is not weakly exogenous. Although the quantile regression estimator is still super-consistent, it is second order biased due to the dependence of x_t and u_t. The second order bias complicates the development of inference procedures about the cointegration vector. Two modifications are proposed in the literature to restore the asymptotic properties of the quantile regression estimator in cointegrating regressions. The first approach adds leads and lags of x_t to the long-run equation (3.1) so that we arrive at

$$y_t = \alpha + \beta' x_t + \sum_{j=-K}^{K} \Delta x'_{t-j} \Pi_j + \varepsilon_t. \tag{3.4}$$

In this dynamic OLS (DOLS) method, originally proposed by Saikkonen (1991), the error term u_t is decomposed into a component related to Δx_t and a pure innovation term ε_t. The quantile regression estimator applied to (3.4) is then asymptotically unbiased. From a practical perspective, the drawback of this approach is the uncertainty regarding the dynamic specification as the number of leads and lags is generally unknown. However, standard model selection criteria can be used to determine the lag length (Choi and Kurozumi (2012)). The second approach involves a nonparametric correction of the original estimator, known as fully modified OLS (FM-OLS) estimation (for a detailed discussion of fully-modified quantile regression estimators, refer to Xiao (2009)).

Cointegration testing is based on the residuals obtained by estimating the long-run equation (3.1). In contrast to the Engle-Granger procedure with the null hypothesis of no cointegration, we follow Xiao and Phillips (2002) and Xiao (2009) and test the null hypothesis of cointegration directly. If y_t and x_t are cointegrated, the residuals should reflect this by displaying fluctuations that resemble a stationary process. A substantial stochastic trend in the residuals would lead to inflated variation over time and would point to the alternative of no cointegration relationship between y_t and x_t.

We begin with the description of the testing procedure for quantile cointegration regression (Xiao (2009)). To measure the fluctuation in the residuals, a partial sum process (related to the CUSUM test literature (Shin (1994))) is constructed as

$$Y_{n\tau} = \frac{1}{\widehat{\omega}^*_\psi \sqrt{n}} \sum_{j=1}^{n} \psi_\tau(\hat{\varepsilon}_{j\tau}) \tag{3.5}$$

where $\widehat{\omega}^{*2}_\psi$ is a consistent estimate of the long-run variance of $\psi_\tau(\hat{\varepsilon}_{j\tau})$. The CUSUM test is based on the residual weights which are mean-zero instead of the residuals for which the τth quantile is zero. The quantile regression residual $u_{t\tau}$ and residual weights $\psi_\tau(\hat{\varepsilon}_{j\tau})$ are obtained from the lead-lag augmented regression in equation (3.4). We use a Kolmogorov-Smirnov type test and base the cointegration test for the τth quantile regression on the supremum of $Y_{n\tau}$. Under the alternative of no cointegration, $\sup Y_{n\tau}$ diverges to infinity.

As a benchmark, we use a conditional-mean cointegration test which follows the same principle. The cointegration test proposed in Xiao and Phillips (2002) uses the test statistic

$$CS_n = \max_{k=1,\ldots,n} \frac{1}{\widehat{\omega}_{\upsilon u} \sqrt{n}} \left| \sum_{j=1}^{k} \hat{u}^+_j \right| \tag{3.6}$$

where $\widehat{\omega}_{\upsilon u}^2 = \widehat{\omega}_u^2 - \widehat{\Omega}_{u\upsilon}\widehat{\Omega}_{\upsilon\upsilon}^{-1}\widehat{\Omega}_{\upsilon u}$ and $\hat{u}^+$ is the vector of FM-OLS residuals. The test statistic is based on the fully modified estimator for β. We define the FM-OLS estimator of β as

$$\hat{\beta}_{LS}^+ = \left[\sum_t y_t^+ x_t' - n\widehat{\lambda}_{\upsilon u}^+\right]\left[\sum_t x_t x_t'\right]^{-1} \tag{3.7}$$

where $y_t^+ = y_t - \upsilon_t'\widehat{\Omega}_{\upsilon\upsilon}^{-1}\widehat{\Omega}_{\upsilon u}$, $\upsilon_t = \Delta x_t$ and $\widehat{\lambda}_{\upsilon u}^+ = \widehat{\lambda}_{\upsilon u} - \widehat{\lambda}_{\upsilon\upsilon}\widehat{\Omega}_{\upsilon\upsilon}^{-1}\widehat{\Omega}_{\upsilon u}$. The relevant long-run (co-)variances are estimated by applying a kernel estimator to the residuals obtained by estimating the cointegrating regression (3.1) with least squares. We choose a Bartlett kernel $k(\cdot)$ with the plug-in bandwidth $M = 1.1447(\phi(1)n)^{1/3}$ according to Andrews (1991), where

$$\phi(1) = \frac{4\hat{\rho}^2}{(1-\hat{\rho}^2)^2} \tag{3.8}$$

and $\hat{\rho}$ is the estimated first order autocorrelation of the least squares residual $\hat{u}_t$. We arrive at the kernel estimates

$$\begin{aligned}
\widehat{\lambda}_{\upsilon u} &= \sum_{h=0}^{\lfloor M\rfloor} k\left(\tfrac{h}{M}\right) C_{\upsilon u}(h), & \widehat{\lambda}_{\upsilon\upsilon} &= \sum_{h=0}^{\lfloor M\rfloor} k\left(\tfrac{h}{M}\right) C_{\upsilon\upsilon}(h), \\
\widehat{\Omega}_{\upsilon u} &= \sum_{h=-\lfloor M\rfloor}^{\lfloor M\rfloor} k\left(\tfrac{h}{M}\right) C_{\upsilon u}(h), & \widehat{\Omega}_{\upsilon\upsilon} &= \sum_{h=-\lfloor M\rfloor}^{\lfloor M\rfloor} k\left(\tfrac{h}{M}\right) C_{\upsilon\upsilon}(h), \\
\widehat{\omega}_u^2 &= \sum_{h=-\lfloor M\rfloor}^{\lfloor M\rfloor} k\left(\tfrac{h}{M}\right) C_{uu}(h),
\end{aligned} \tag{3.9}$$

where $C_{\upsilon\psi}(h)$, $C_{\upsilon\upsilon}(h)$ and $C_{uu}(h)$ are sample covariances defined by $C_{\upsilon u}(h) = n^{-1}\sum \upsilon_t \hat{u}_{t+h}$, $C_{\upsilon\upsilon}(h) = n^{-1}\sum \upsilon_t \upsilon_{t+h}'$, $C_{uu}(h) = n^{-1}\sum \hat{u}_t\hat{u}_{t+h}$, respectively. For a more comprehensive discussion of fully modified least squares, see Hansen (1992) and Xiao and Phillips (2002).

Quantile cointegration is able to reveal a quantile-dependent structure of the cointegration relationship. If the cointegration relationship is nonlinear and state-dependent, the quantile-dependent coefficients should be different from the constant cointegrating coefficients in at least one quantile. Hence, it is of interest to test the null hypothesis $H_0 : \beta(\tau) = \bar{\beta}$ over a sequence of quantiles $\tau \in \mathscr{T}$. We consider the least squares estimator, $\hat{\beta}_{LS}$, obtained from the linear cointegration model to be a suitable candidate to approximate $\bar{\beta}$. Xiao (2009) proposes the process

$$V_n(\tau) = n(\hat{\beta}(\tau) - \hat{\beta}_{LS}) \tag{3.10}$$

and evaluates $\sup|V_n(\tau)|$ with a bootstrap procedure assuming a constant β. A large $\sup|V_n(\tau)|$ statistic points to overall quantile effects. It turns out to be a computational advantage to use the lead-lag augmentation to obtain $\hat{\beta}(\tau)$ and $\hat{\beta}_{LS}$ for the resampling algorithm. The bootstrap procedure is divided into five steps:

(1) Obtain the estimates $\hat{\beta}(\tau)$ and $\hat{\beta}_{LS}$ from (3.4) by quantile regression and linear regression, respectively. Further, calculate $\sup|V_n(\tau)|$ and the least squares residuals,

$$\hat{u}_t = y_t - \hat{\alpha} - \hat{\beta}_{LS}' x_t \qquad t = 1, \ldots, n.$$

(2) Define $\hat{w}_t = (\upsilon_t, \hat{u}_t)$ with $\upsilon_t = \Delta x_t$, estimate the bivariate VAR model,

$$\hat{w}_t = \sum_{j=1}^{q} B_j \hat{w}_{t-j} + e_t \qquad t = q+1, \ldots, n,$$

and save the fitted residuals $\hat{e}_t = \hat{w}_t - \sum_{j=1}^{q} \hat{B}_j \hat{w}_{t-j}$. The lag length q can be chosen on the basis of the AIC.

(3) Center the residuals

$$\hat{e}_t - \frac{1}{n-q} \sum_{j=q+1}^{n} \hat{e}_t$$

and draw samples e_t^* from the centered residuals. Generate w_t^* using the estimate $\hat{B}_j$ and e_t^* so that

$$w_t^* = \sum_{j=1}^{q} \hat{B}_j w_{t-1}^* + e_t^* \qquad t = q+1, \ldots, n$$

with $w_j^* = \hat{w}_j$ for $j = 1, \ldots, n$.

(4) Generate x_t^* from $x_t^* = x_{t-1}^* + \upsilon_t^*$ with $x_1^* = \upsilon_1^*$ and

$$y_t^* = \hat{\alpha} + \hat{\beta}' x_t^*.$$

(5) Obtain the bootstrap estimates $\hat{\beta}^*(\tau)$ and $\hat{\beta}_{LS}^*$ from

$$y_t^* = \alpha + \beta' x_t^* + \sum_{j=-K}^{K} \Delta x^{*\prime}_{t-j} \Pi_j + \varepsilon_t^*,$$

and calculate $V_n^* = n(\hat{\beta}^*(\tau) - \hat{\beta}_{LS}^*)$. Repeat steps 2-5 sufficiently often to approximate the distribution of $\sup|V_n(\tau)|$.

Should the overall quantile effects test lead to a rejection of the null hypothesis, we can assume an inherent nonlinearity in the cointegration relationship between y_t and x_t.

4 Empirical Analysis

We analyze gold and silver spot prices at a monthly frequency from August 1971 to April 2014 and daily spot and futures prices from March 1980 to April 2014. The London OTC market and New York COMEX are considered major gold and silver markets. We use the morning official fixing price at the London Bullion market for the daily price series and build a monthly price series from the first price reported in each month. The futures prices are obtained for COMEX 100 ounces gold contracts and COMEX 5000 ounces silver contracts. We denote the spot prices of silver and gold as S and G, respectively. The futures prices are denoted as S^F and G^F. Gold prices are denominated in USD per troy ounce whereas silver is denominated in USD cents per troy ounce.

Since the sample of the spot prices includes the Hunt brothers' attempt to corner the silver market in the late 1970s and early 1980s, we have to treat this period as a separate regime. The Hunt brothers and their collaborators tried to restrict the supply of silver on the market so that it became difficult for investors who sold short to deliver at the end of the contract. The price of silver subsequently increased dramatically and this peak appears as a striking anomaly in the data. However, they only acted in the silver market and did not act on the gold market in the same fashion. It has to be assumed that the potential long-run relationship between gold and silver was exogenously altered during the Hunt brothers episode. In contrast, we do not treat the financial crisis in 2008 as a separate regime since gold and silver markets were both affected. Prices of precious metals increased due to a higher demand of investors for safe haven assets without necessarily changing the relationship between them.

We start the analysis by testing all price series for their order of integration. Each series is determined to be integrated of order one. The results of the unit root tests are depicted in Table 3.1.

Table 3.1: Augmented Dickey-Fuller tests for gold and silver prices

	drift	lags	*trend*	lags		*drift*	lags
Gold_m	-0.158	1	-1.083	1	ΔGold_m	-16.01^{***}	1
Silver_m	-2.632^{*}	1	-3.138^{*}	1	ΔSilver_m	-16.58^{***}	1
Gold_d	-0.049	1	-1.452	1	ΔGold_d	-76.73^{***}	1
Silver_d	-1.707	1	-2.666	1	ΔSilver_d	-79.12^{***}	1
Gold_d^F	-0.109	1	-1.475	1	ΔGold_d^F	-67.21^{***}	1
Silver_d^F	-1.674	1	-2.561	1	ΔSilver_d^F	-65.75^{***}	1

The subscript m denotes monthly observations and d denotes daily observations, respectively. Including an intercept in the ADF test equation is indicated with *drift*, including an additional linear trend term with *trend*. The lag selection was achieved via Bayesian Information Criterion (BIC).
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

4 .1 Monthly spot prices

In the first part of the empirical analysis, we revisit a similar data set found in Escribano and Granger (1998). The long-run equilibrium model between gold and silver prices is expressed as

$$S_t = \alpha + \beta\, G_t + \gamma_1\, d_t + \gamma_2\, d_t \cdot G_t + u_t. \tag{3.11}$$

The dummy variable d_t and the interaction term $d_t \cdot G_t$ model a change in the intercept and the slope parameter for the 'silver bubble' period from June 1979 to March 1980.[3] The cointegrating vector is estimated by FM-OLS and the CUSUM cointegration test is applied to the residuals $\hat{u}_t^+$. The $\sup CS_n$ statistic amounts to 1.569 with a p-value of 0.002 for the monthly series such that the null hypothesis of linear cointegration can be rejected. This means we find strong evidence that gold and silver are not cointegrated in the Engle-Granger framework assuming a constant cointegrating vector with a one-time break. The FM-OLS estimator for β amounts to 1.729 and the DOLS estimator takes the value 1.847. This result supports the findings in Escribano and Granger (1998) who report no cointegration relationship for the full sample.

[3]Removing the dummy variable and the interaction term and thereby ignoring the Hunt brothers episode leaves the results virtually unchanged.

We now test for quantile cointegration. The $\sup Y_n$ statistic is computed for each quantile τ and is plotted in the left panel of Figure 3.4. The null hypothesis of quantile cointegration cannot be rejected at the 5% significance level for any quantile τ. The point estimates for the quantile-dependent estimator are depicted in the left panel of Figure 3.5. It can be inferred from the plots that significant asymmetry is present in the quantile regression estimates. The bootstrap test based on $\sup V_n$ with a value of 1224 and p-value less than 0.001 supports this conjecture for the monthly frequency and points to significant overall quantile effects.[4] The slope parameter β largely coincides with the conditional-mean benchmark (DOLS) with the exception that the lower tail estimates are slightly smaller than the DOLS estimate. However, the point estimates for quantiles above the 80% quantile are significantly larger than the benchmark.

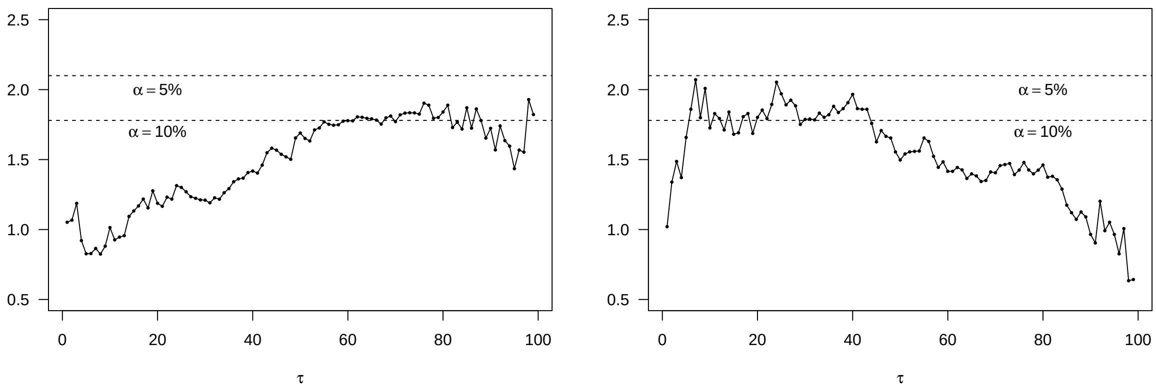


Figure 3.4: The estimated $\sup Y_n$ statistics for the 1% to 99% quantile (monthly series). The left panel corresponds to specification (3.11) and the right panel corresponds to (3.12). The critical value, dashed line, is 1.78 for the 10% significance level (2.1 for the 5% significance level).

The quantile cointegration estimates suggests that silver prices respond stronger to gold prices changes if silver prices are high. A plot of the historic time series (Figure 3.3) shows that high gold and silver prices occurred during the Hunt brothers episode and during economic crisis periods. In general, the quantile cointegration framework is not able to identify periods with stronger responses directly, since conditional quantiles are estimated. However, we are able to indicate the periods in which the residuals were assigned a higher weight. This is depicted in the upper and lower panel of Figure 3.6, where we mark the higher weighted residuals for the lower tail (10% quantile) and the upper tail (90% quantile) with a red rhombus.

The indicated periods of weaker dependence match the results of Escribano and Granger (1998) who claim that the cointegration relationship dissolves towards the end of their sample in 1990. Periods of stronger dependence are found during the 'silver bubble' and during the financial crisis. Lucey and Tully (2006) find a different pattern but their sample period is shorter and excludes the Hunt brothers episode as well as the financial crisis. In general our data-driven framework finds a state-dependence of the long-run relationship between gold and silver that resemble the pre-specified conditional-mean results

[4] We use 600 replications of all variables present in the linear regression for bootstrapping of the $\sup V_n$ test. However, the results of the $\sup V_n$ test have to be interpreted cautiously since we find evidence against cointegration in the conditional-mean benchmark model. The estimate of β under constancy is not well-defined and the bootstrap procedure involves nonstationary variables. In this case the nonlinearity test is potentially oversized.

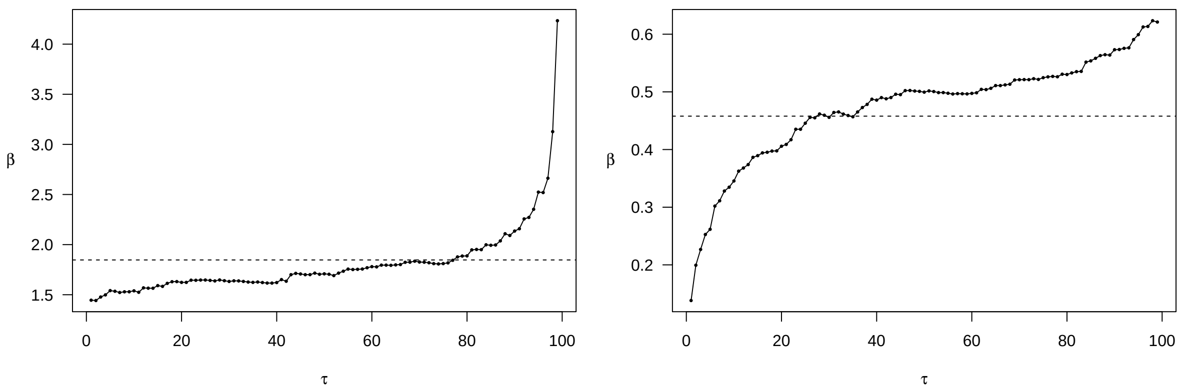


Figure 3.5: Estimation results for the slope coefficient in the lead-lag augemented quantile cointegration regression (monthly series). The left panel corresponds to specification (3.11) and the right panel corresponds to (3.12). The DOLS estimate serves as a benchmark (dashed line).

by Baur and Tran (2014) who also find a different cointegrating vector during times of financial stress.

Switching the roles of gold and silver, the long-run relationship can be respecified in the form of

$$G_t = \alpha + \beta S_t + \gamma_1 d_t + \gamma_2 d_t \cdot S_t + u_t, \tag{3.12}$$

assuming that silver leads the pricing process. It is possible to use this alternative normalization, since the estimator for β is unbiased in either specification after the endogeneity correction through FM-OLS or DOLS estimation is applied. Now, the quantile-dependent coefficients $\beta(\tau)$ measure the response of gold prices to silver prices given the information set in period t. The CUSUM cointegration test for the conditional-mean case results in a $\sup CS_n$ statistic of 1.682, so that the null hypothesis of cointegration is rejected at the 5% significance level. The $\sup Y_n$ statistics for the quantile process is depicted in the right panel of Figure 3.4. The pattern of the quantile-dependent estimates differs compared to the previous specification: The response to silver prices is again stronger for upper conditional quantiles of gold but instead of a slow increase until the 80% quantile, the pattern resembles logarithmic growth in τ. The point estimates of $\beta(\tau)$ can be found in the right panel of Figure 3.5. The conditional-mean estimates are 0.458 (DOLS) and 0.562 (FM-OLS), respectively.

4 .2 Daily spot and futures prices

The historic price series for gold and silver futures contracts starts in March 1980. Since the Hunt Brothers episode is excluded from the sample, we estimate the long-run equation without the need of any dummy variables. The results for the daily spot prices series are largely in accordance with the monthly series, we obtain conditional-mean estimates 1.954 (DOLS) and 1.89 (FM-OLS). The quantile-dependent estimates are depicted in the left panel of Figure 3.7. The response is weaker compared to the benchmark value in lower quantiles and stronger for upper quantiles. The CUSUM test statistic based on the fully modified residuals is 1.3 with a p-value of 0.024. Thus, we find only weak evidence against the null hypothesis of cointegration considering the sample size of 8884 for daily prices compared to the sample size of 513 for the monthly series. The quantile cointegration test statistics are depicted in the left

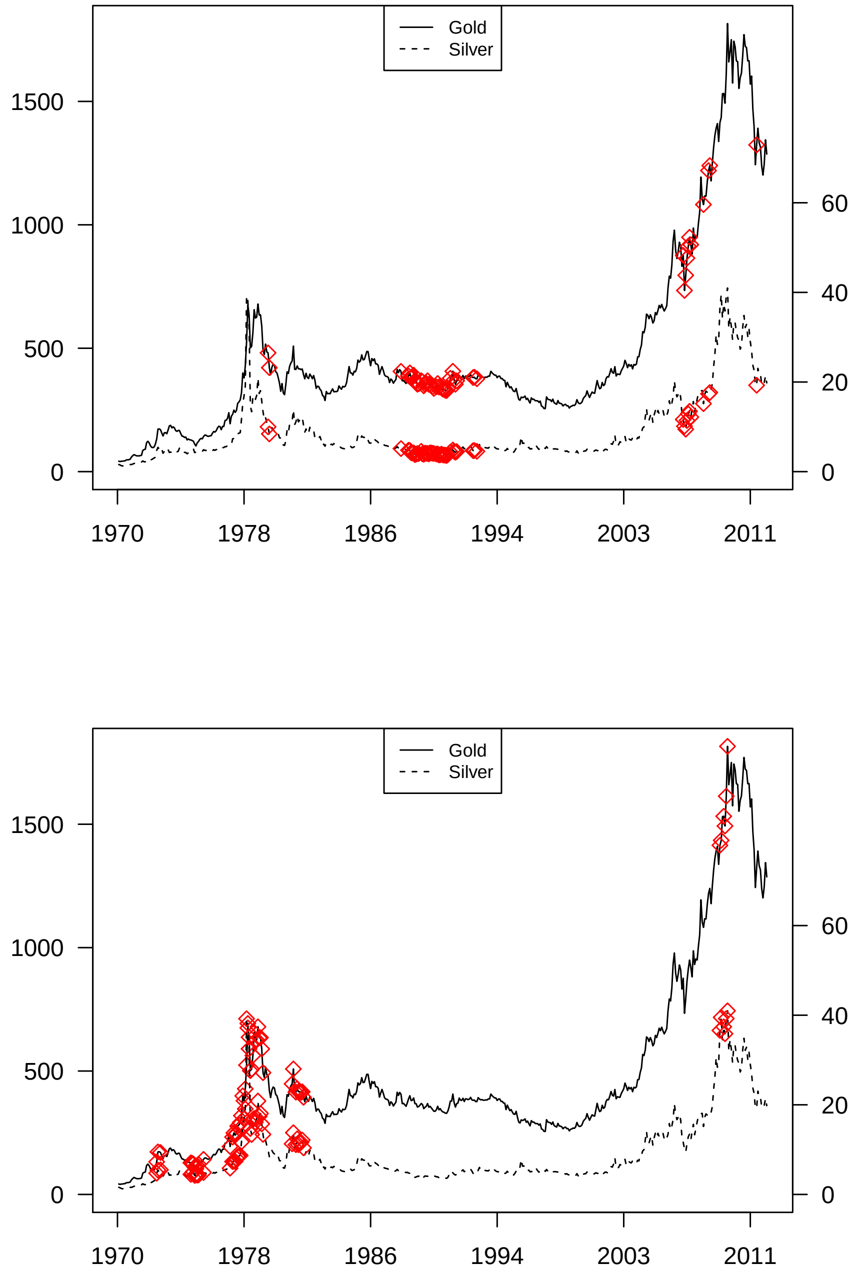


Figure 3.6: Periods of conditional 10% quantile monthly silver prices (upper panel) and conditional 90% quantile monthly silver prices (lower panel). Observations are marked with a red rhombus if they received a higher weight in the loss function of the 10% (90%) quantile regression for silver as the dependent variable.

panel of Figure 3.8. We observe generally larger $\sup Y_n$ statistics for the daily series and have to reject the null hypothesis for upper quantiles. However, the results for the daily series are not unexpected since the power of the $\sup Y_n$ naturally increases with sample size which is represented by the 0.1% significance level. The test of constancy of the cointegrating vector over all quantiles τ gives $\sup V_n = 3320$ and a p-value below 0.001. Hence, we also find a statistically significant nonlinear response of daily silver prices to daily gold prices. Interestingly, the periods of conditional 10% (90%) quantile daily silver prices, depicted in Figure 3.9, are identified slightly different compared to the quantile cointegration model for monthly prices.

Switching to specification (3.12) again yields results similar to the monthly series. The values $\beta(\tau)$, depicted in the right panel of Figure 3.7, can be characterized as a linear function of the quantiles of the conditional distribution of gold. The response to silver prices is weak for lower quantiles and strong for upper quantiles. The null hypothesis of quantile cointegration is only rejected at the 0.1% significance level for intermediate quantiles between 15%-25%.

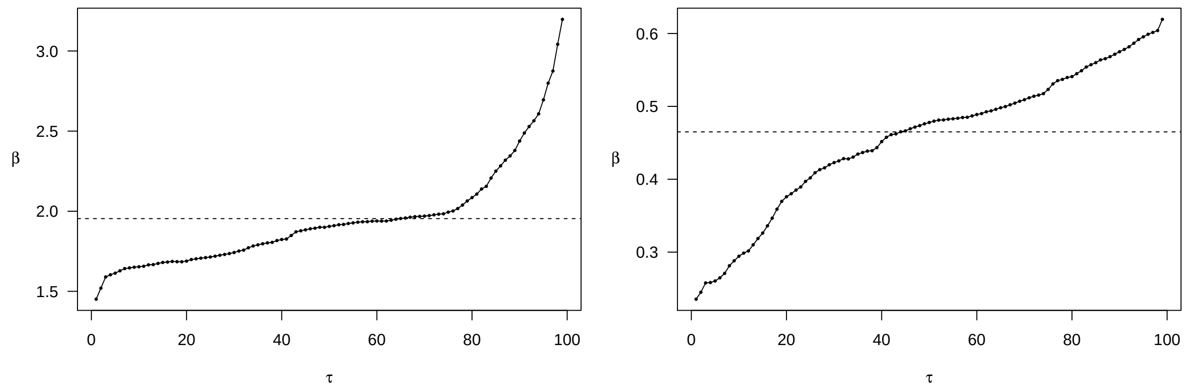


Figure 3.7: Estimation results for the slope coefficient in the lead-lag augemented quantile cointegration regression (daily series). The left panel corresponds to specification (3.11) and the right panel corresponds to (3.12). The DOLS estimate serves as a benchmark (dashed line).

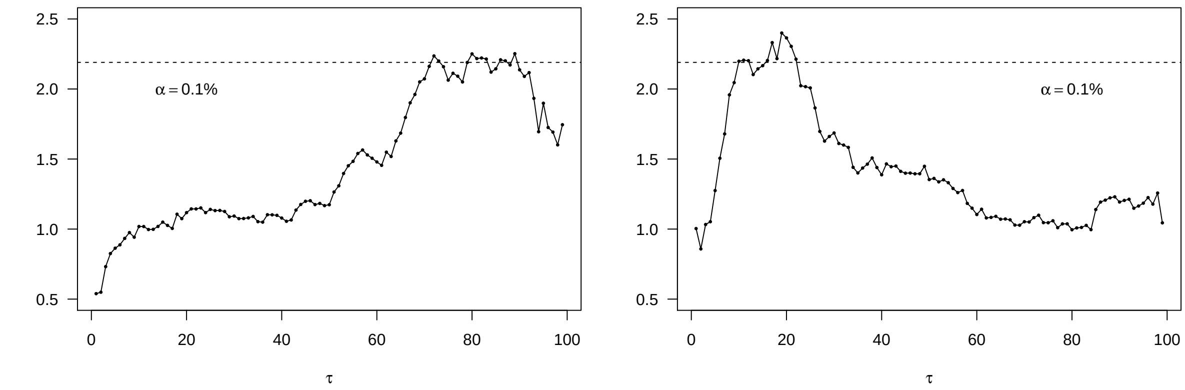


Figure 3.8: The estimated $\sup Y_n$ statistics for the 1% to 99% quantile (daily series). The left panel corresponds to specification (3.11) and the right panel corresponds to (3.12). The critical value for the 0.1% significance level is 2.19.

The long-run equilibrium relationship between the prices of gold and silver futures contracts is expressed as

$$S_t^F = \alpha + \beta G_t^F + u_t. \quad (3.13)$$

The $\sup CS_n$ statistic amounts to 1.201 with a p-value of 0.043 which does not lead to a rejection of the

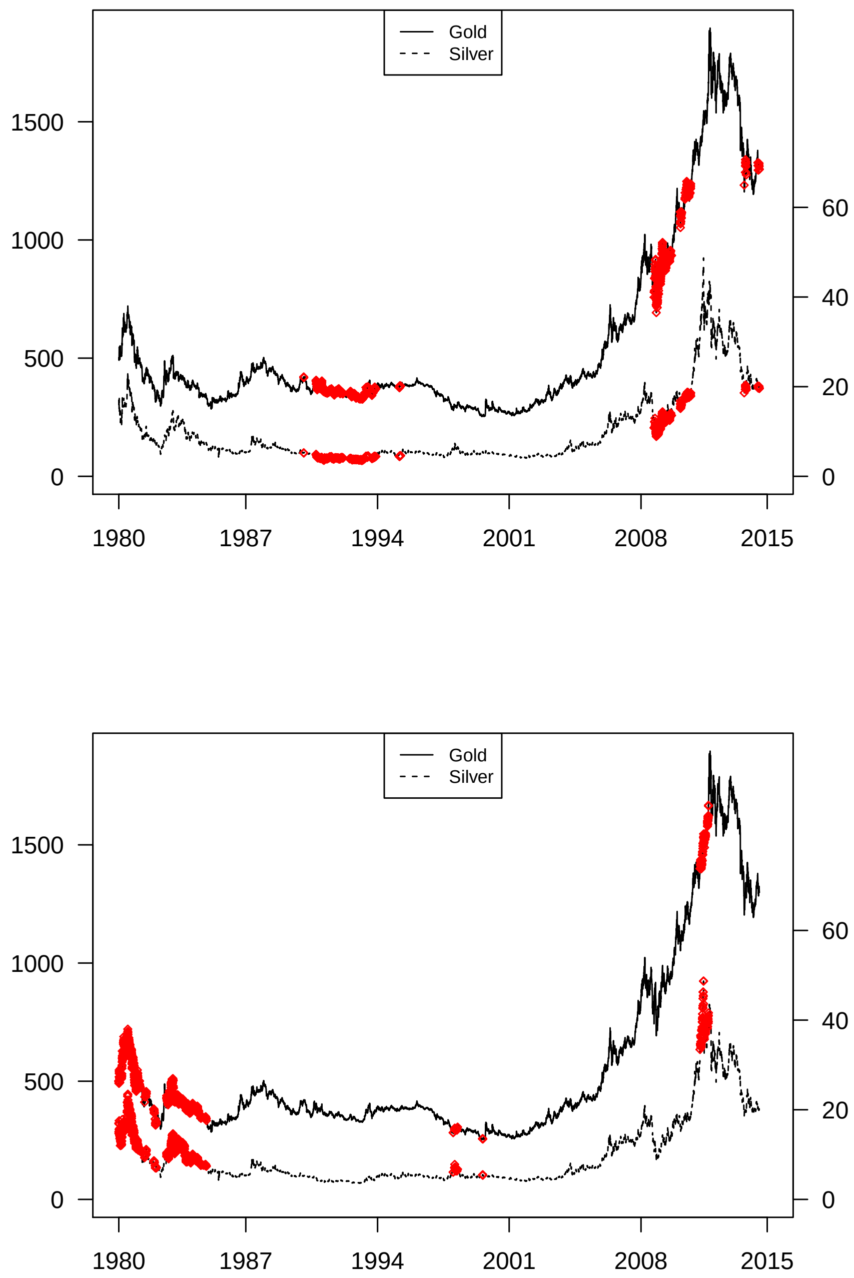


Figure 3.9: Periods of conditional 10% quantile daily silver prices (upper panel) and conditional 90% quantile daily silver prices (lower panel). Observations are marked with a red rhombus if they received a higher weight in the loss function of the 10% (90%) quantile regression for silver as the dependent variable.

null hypothesis of cointegration considering the large sample size. Gold and silver futures prices seem to be cointegrated even with a constant cointegrating vector. However, the $\sup Y_n$ process points to no cointegration for upper quantiles of silver. The quantile-dependent estimates of β are depicted in the left panel of Figure 3.10 and display an increasing response to gold futures prices for upper quantiles of the conditional distribution. The test of overall quantile effects results in the test statistic $\sup V_n = 1517$

with p-value less than 0.001 so that we also find strong evidence for nonlinearity in the cointegration relationship between gold and silver futures prices.

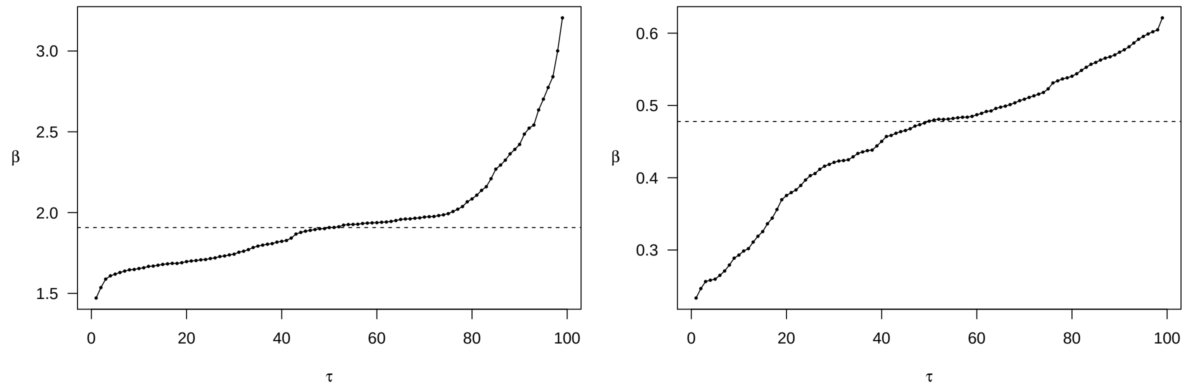


Figure 3.10: Estimation results for the slope coefficient in the lead-lag augemented quantile cointegration regression (futures contract prices). The left panel corresponds to specification (3.13) and the right panel corresponds to (3.14). The DOLS estimate serves as a benchmark (dashed line).

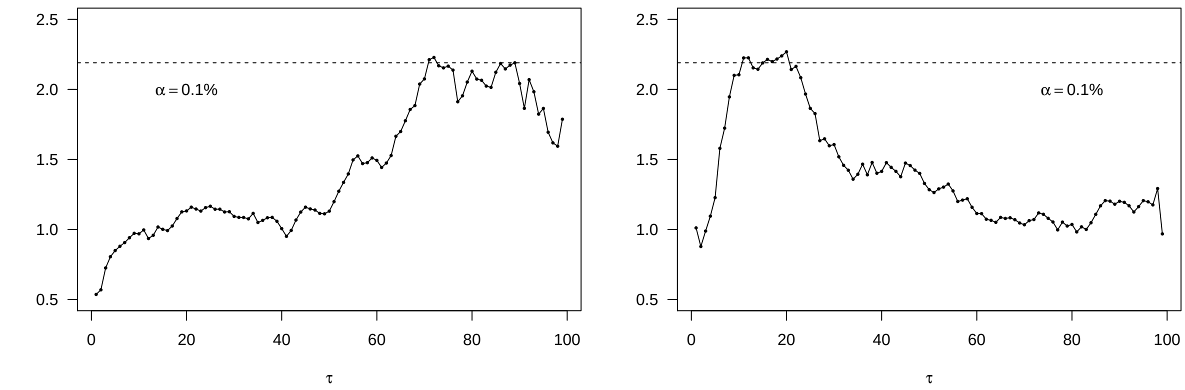


Figure 3.11: The estimated $\sup Y_n$ statistics for the 1% to 99% quantile (futures contract prices). The left panel corresponds to specification (3.13) and the right panel corresponds to (3.14). The critical value for the 0.1% significance level is 2.19.

The respecified long-run equilibrium takes the form of

$$G_t^F = \alpha + \beta S_t^F + u_t. \tag{3.14}$$

The CUSUM cointegration test gives a $\sup CS_n$ statistic of 1.159 with p-value 0.055 and the null hypothesis of cointegration can not be rejected at appropriate significance levels. The $\sup Y_n$ statistic leads to the rejection of the null hypothesis for lower conditional quantiles of the gold futures prices. The quantiles-dependent estimates show a pattern similar to the daily spot price series with an increasing response from lower to upper conditional quantiles of gold. The $\sup V_n$ test supports the graphical illustration and rejects the null hypothesis of constancy of the slope parameter across quantiles.

The results for daily spot and futures prices are very similar which indicates that the asymmetrical

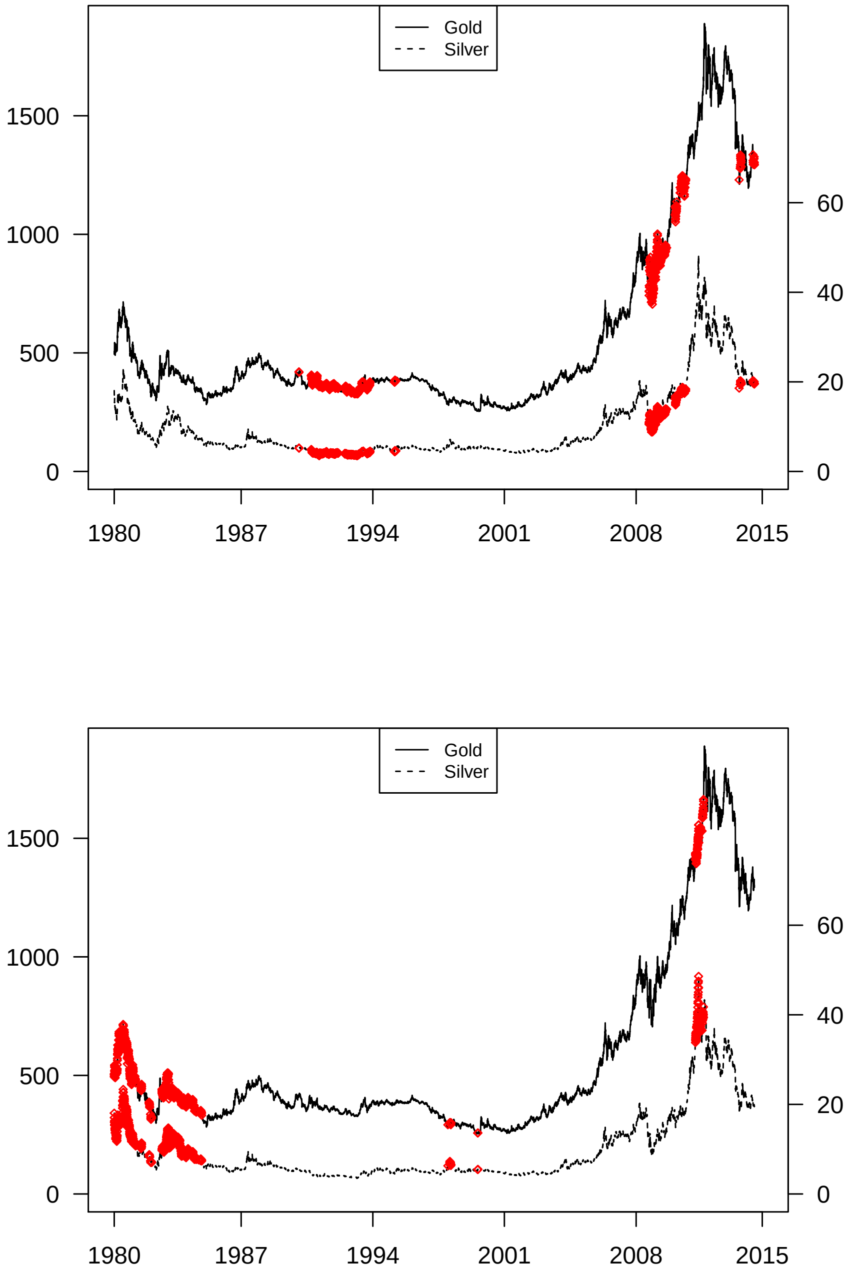


Figure 3.12: Periods of conditional 10% quantile silver futures prices (upper panel) and conditional 90% quantile silver futures prices (lower panel). Observations are marked with a red rhombus if they received a higher weight in the loss function of the 10% (90%) quantile regression for silver as the dependent variable.

pattern is not a unique feature of the price discovery in spot markets, i.e. the comovement in bubble and crisis period is not necessarily created by distinct features of the gold fixing process but rather could be generated by a general need of investors for safe haven assets. Both precious metals share store of value characteristics which are most sought after during times of financial market turbulence. In tranquil times, the individual (industrial) demand for gold or silver seems to drive the individual prices.

5 Conclusion

In this paper, we estimate a time-varying cointegrating vector for the gold and silver long-run relationship depending on the innovation quantile. Our empirical results point to a significantly asymmetric dependence between silver prices and gold prices. We observe a stronger response of silver prices to gold price changes when silver prices are at a relatively high level and a stronger response of gold prices to silver price changes when gold prices are at a relatively high level. The long-run relationship between gold and silver is therefore best characterized by a state-dependence. More specifically, after the prices were deregulated in 1971, high gold and silver prices can generally be found in times of financial stress and only in those periods, we find a strong dependence between prices which results in a visible comovement. It can be suspected that one of the key properties of gold and silver – the store of value aspect – plays a more prominent role in periods of financial turbulence where other assets lose value and the investors' search for safe haven assets increases demand for gold and silver. This in turn increases prices for gold and silver simultaneously. Moreover, the analysis over a post-bubble sample and at a different frequency shows that the asymmetrical pattern is remarkably stable and the results can easily be transferred to the futures market.

In general, we emphasize the abilities of the quantile cointegration framework to detect nonlinearities in a cointegration relationship. Considering our empirical results, it is now possible to understand the difficulties, described in previous studies, to find a stable long-run relationship between the two precious metals. Although we observe a comovement of both prices over decades, we fail to estimate a single constant cointegrating coefficient that connects both prices. Allowing for a more general time-varying cointegrating vector enables us to capture the time- and state-dependence of the long-run relationship. Taking into account that the cointegrating vector is allowed to change in every period, we conclude that a long-run relationship exists but it is not particularly stable. The estimated relationship cannot directly be used for forecasting, since the exact state of the variables is generally unknown. From that perspective, finding evidence for quantile cointegration but not finding evidence for linear cointegration does not contradict the weak form efficiency of gold and silver markets. In fact, given our results, a statistical arbitrage strategy based on the weakly linked gold and silver prices under the assumption of a single constant coefficient would be very risky.

Chapter 4

Testing for cointegration with SETAR adjustment in the presence of structural breaks

1 Introduction

The residual-based threshold cointegration models developed by Enders and Siklos (2001) are a useful addition to the toolbox of researchers working with multivariate time series. They are easy to apply, allow for discontinuous adjustment to a long-run equilibrium and nest linear cointegration in the sense of Engle and Granger (1987) as a special case. The dynamics of the adjustment process are described by a two-regime threshold autoregressive (TAR) model which partitions the residual process according to a threshold value and specifies different coefficients of the leading autoregressive lag for each regime. It can therefore be considered a restricted model under the general class of TAR models described by Tong (1983). A prominent application in the economics literature is the empirical analysis of asymmetric price transmissions in which case non-stationary price series form a cointegrating relationship and may feature asymmetric adjustment to long-run equilibrium. The speed of adjustment is usually assumed to depend on the sign and size of the deviations from the long-run equilibrium. While threshold cointegration tests are suitable to study these cases, they do not account for possible structural change in the long-run relationship. It is well-known that conventional residual-based cointegration tests perform poorly when a cointegration relationship has structural breaks. Maki (2013) found that the power property of threshold cointegration tests is more robust to structural breaks than, for example, the Engle-Granger cointegration tests assuming linear adjustment. Nevertheless, the power of all residual-based cointegration tests is impaired if the tests do not model the structural breaks explicitly. Consequently, it is difficult to provide evidence for the existence of a cointegration relationship. Furthermore, the estimated adjustment to equilibrium is biased if the cointegrating vector does not account for structural change. This is of special concern from a practical perspective since the Financial Crisis of 2008 is suspected to have induced structural change in several economic relationships (see for example Zhou and Kutan (2011) for an examination of the real exchange rate during the Financial Crisis and Lehkonen (2015) for the effects on stock market integration).

An extensive body of literature exists on the problem of structural instability in time series. Based on the seminal work of Perron (1989), several unit root tests accounting for structural change have been developed (see, inter alia, Zivot and Andrews (1992), Lumsdaine and Papell (1997) and Lee and Strazicich (2003)). Structural breaks in linear cointegration models are addressed in Gregory and Hansen (1996a,b), Carrion-i Silvestre and Sanso (2006), Westerlund and Edgerton (2007) and Hatemi-J (2008). Gregory and Hansen (1996a), henceforth GH, propose a residual-based cointegration test with structural break. Their test does not require a pre-specified breakpoint which is rarely known in empirical applications. Instead, a single unknown breakpoint is determined from the data based on one of three structural break models. However, the GH test is only suitable for cointegration models with linear adjustment.[1] We contribute to the literature by extending the GH test to include a form of non-linear adjustment. This new test is residual-based and uses a SETAR model to describe the adjustment toward equilibrium. Thereby, we also provide an extension to the SETAR Enders-Siklos test which is robust to a structural break in the cointegrating vector.

We derive the limiting distributions of the test statistic considered in this paper and provide a formal proof. The properties of the proposed test are investigated by Monte Carlo experiments for a variety of models ranging from linear adjustment with no structural break to non-linear adjustment with structural break in the intercept and slope coefficients. The results suggest that a break in the intercept does not influence the power of the threshold cointegration test enough to justify modelling the structural break. However, a break in the slope coefficients reduces the power of the SETAR Enders-Siklos test substantially such that our proposed test performs clearly better than its benchmark. In addition, we find that the unknown breakpoints are estimated accurately by the new procedure.

The methodology is applied to empirical data in the context of the 'rockets and feathers' hypothesis. We use US gasoline market data covering the Financial Crisis. We illustrate that empirical evidence for the existence of a long-run relationship between neighbouring stages of the gasoline value-chain can only be provided if we control for a structural break in the cointegrating vector. Using either cointegration model, we do not find evidence for asymmetric adjustment toward equilibrium, i.e. we do not find empirical support for the 'rockets and feathers' hypothesis.

The paper is organized as follows. Section 2 describes the models and the cointegration testing procedure, Section 3 presents the asymptotic distribution of the test statistic. Section 4 is devoted to the Monte Carlo study. Section 5 reports the results of the empirical application, and Section 6 summarizes the study.

2 Models and cointegration testing

The long-run equilibrium equation of EG cointegration models is given by

$$\begin{aligned} y_t &= \mu + \alpha_1 x_{1t} + \alpha_2 x_{2t} + \cdots + \alpha_m x_{mt} + e_t \\ &= \mu + \alpha' x_t + e_t \end{aligned} \tag{4.1}$$

where $t = 1,2,\ldots,T$ is the time series index, y_t and $x_t = (x_{it}, x_{2t}, \ldots, x_{mt})'$ are $I(1)$ variables, μ is an intercept, $\alpha' = (\alpha_1, \alpha_2, \ldots, \alpha_m)$ is a vector of slope coefficients and e_t is the equilibrium error. The null hypothesis of no cointegration is rejected if the residuals obtained from least squared estimation of (4.1)

[1]The effects on the power properties of linear cointegration tests, if the equilibrium error follows a nonlinear adjustment process, are reported in Pippenger and Goering (2000).

are mean-zero stationary. Since the parameters μ and α are time-invariant, a residual-based cointegration test based on (4.1) becomes invalid if the long-run equilibrium is subject to structural change.

Following Perron (1989) and Gregory and Hansen (1996a), we consider three forms of structural change.[2] First, in the C model, a break in the intercept μ is considered. This model is named the 'crash' model and relates to events that cause a parallel shift of the equilibrium equation. Second, the C/T model adds an additional trend term to the equilibrium equation. Third, in the C/S model, a break in both the constant and the slope parameter is specified. The C/S model is named the 'regime shift' model. The three models are given as follows,

$$\begin{array}{lrcl} (C) & y_t & = & \mu_1 + \mu_2 \varphi_{t,\tau} + \alpha' x_t + e_t \\ (C/T) & y_t & = & \mu_1 + \mu_2 \varphi_{t,\tau} + \delta t + \alpha' x_t + e_t \\ (C/S) & y_t & = & \mu_1 + \mu_2 \varphi_{t,\tau} + \alpha_1' x_t + \alpha_2' x_t \varphi_{t,\tau} + e_t \end{array} \tag{4.2}$$

where μ_1, μ_2 are constants, $\alpha_1 = (\alpha_{11}, \alpha_{12}, \ldots, \alpha_{1m})'$ and $\alpha_2 = (\alpha_{21}, \alpha_{22}, \ldots, \alpha_{2m})'$ are slope coefficients. The dummy variable $\varphi_{t,\tau}$ is defined as

$$\varphi_{t,\tau} = \begin{cases} 1 & \text{if } t > [T\tau] \\ 0 & \text{if } t \leq [T\tau] \end{cases}, \tag{4.3}$$

where $\tau \in (0,1)$ denotes the relative timing of the breakpoint (break fraction), and $[\cdot]$ denotes integer part. The timing of the breakpoint is rarely known in empirical applications so that the GH test is constructed without the need of pre-specified breakpoints. More specifically, a grid search over all possible breakpoint is employed, i.e. the structural change model is repeatedly estimated for each possible breakpoint $\tau \in \mathscr{T}$. The set $\mathscr{T}$ can be any compact subset of $(0,1)$ which excludes endpoint results. GH suggest to trim the upper and lower 15 percent and, for computational reasons, consider only integer steps, $\mathscr{T} = ([0.15T], [0.85T])$. Estimating one of the structural break models in (4.3) by least squares for each breakpoint yields a sequence of residuals. The GH test applies the ADF test to each sequence and evaluates the null hypothesis of no cointegration based on the smallest values of the t ratios across all $\tau \in \mathscr{T}$. The infimum statistic is chosen since it puts the most weight on the alternative hypothesis. If the null hypothesis is rejected, the break fraction $\hat{\tau}$ that corresponds to the infimum statistic is considered to be the most likely breakpoint.

In order to account for asymmetric adjustment, the two-regime self-exciting threshold autoregressive (SETAR) model is now used to describe the adjustment toward equilibrium. The SETAR model for the equilibrium error process e_t is given by

$$\Delta e_t = \rho_1 e_{t-1} \mathbb{1}\{e_{t-1} \geq \lambda\} + \rho_2 e_{t-1} \mathbb{1}\{e_{t-1} < \lambda\} + \sum_{j=1}^{K} \gamma_j \Delta e_{t-j} + \varepsilon_t, \tag{4.4}$$

where $\mathbb{1}\{\cdot\}$ denotes the Heaviside indicator function, the parameter λ is a possibly non-zero threshold value and ε_t is a stationary mean zero error term. The coefficient ρ_1 measures the mean-reversion toward the cointegrating vector after a shock greater than or equal to λ whereas ρ_2 measures the mean-reversion

[2] We restrict our analysis to these three models. However, our methodology can easily be adapted for other structural break models, as for example given in Gregory and Hansen (1996b) and Hatemi-J (2008).

toward the cointegrating vector after a shock less than λ. The indicator function in this case is set according to the level of e_{t-1}. In an alternative specification, suggested by Enders and Granger (1998) and Caner and Hansen (2001), the indicator function is set depending on Δe_{t-1}. However, the so-called momentum threshold autoregressive (MTAR) model is not covered in this paper.

Under the null hypothesis of no cointegration, $\rho_1 = \rho_2 = 0$, the data-generating process (DGP) of e_t is symmetric and a unit root is present in both regimes. Model (4.4) is a special case of the general class of threshold autoregressive models in that it does not allow for regime-specific deterministic terms and regime-specific dynamics beyond the leading autoregressive lag. This restriction is convenient since it circumvents the problem of having an identified threshold under the null hypothesis resulting in an asymptotic distribution of the test statistic that depends on nuisance parameters (see Caner and Hansen (2001) for a more detailed discussion in the context of MTAR processes with a unit root). Furthermore, the Engle-Granger test for symmetric adjustment ($\rho_1 = \rho_2$) is itself a special case of (4.4). Petruccelli and Woolford (1984) show that the stationarity of the SETAR process $\{e_t\}_1^\infty$ is ensured if $\rho_1 < 0$, $\rho_2 < 0$ and $(1+\rho_1)(1+\rho_2) < 1$ for any value λ. Assuming stationarity, Tong (1983, 1990) demonstrated that least squares estimators of ρ_1 and ρ_2 are asymptotically normally distributed. Enders and Siklos (2001) recommend a Wald-type F-test to test the null hypothesis of no cointegration in their model without structural breaks. However, since the F-test can lead to rejection of the null hypothesis when only one coefficient is negative, the test should only be applied if both point estimates suggest a mean-reversion behaviour. In other words, the one-sided alternative $\rho_1 < 0 \wedge \rho_2 \geq 0$ or $\rho_2 < 0 \wedge \rho_1 \geq 0$ should not lead to rejection of the null hypothesis.

In the case of a cointegration model with potential structural break and SETAR adjustment, we propose the following cointegration test: First, an appropriate structural break model is selected from (4.3) and the cointegrating regression is estimated by least squares for each break fraction $\tau \in \mathscr{T}$. Then, the F-statistic, F_τ, is computed for each sequence of residuals. Since the null hypothesis of no cointegration is naturally rejected for large values of the F-statistic, the supremum statistic,

$$F^* = \sup_{\tau \in \mathscr{T}} F_\tau, \tag{4.5}$$

is used to evaluate the null hypothesis of no cointegration against the alternative of threshold cointegration with possible structural break. The largest value found in this grid search also determines the most likely breakpoint.

3 Asymptotic distribution

The asymptotic theory for SETAR processes with a unit root was developed in Seo (2008). Maki and Kitasaka (2015) derive the asymptotic distribution of Wald statistics in a three-regime threshold cointegration model of which the two-regime threshold cointegration model is a special case. Gregory and Hansen (1996a) provide results for cointegration test statistics which are functions of the break fraction parameter τ. Hence, we follow Gregory and Hansen (1996a) and Maki and Kitasaka (2015) closely in our derivations. For notational convenience we use '$\Rightarrow$' to signify weak convergence of the associated probability measures. Continuous stochastic processes such as the Brownian motion $B(s)$ on [0,1] are simply written as B. We also write integrals with respect to the Lebesgue measure such as $\int_0^1 B(s)ds$

simply as $\int_0^1 B$.

Let $\{z_t\}_0^\infty$ be an $(m+1)$-vector integrated process whose data generating process is

$$z_t = z_{t-1} + \xi_t, \qquad t = 1, 2, \dots \tag{4.6}$$

where it is assumed that $T^{-1/2} z_0 \xrightarrow{p} 0$ so that z_0 can be treated as either fixed or random and the results do not depend on the initial condition. The $(m+1)$-vector random sequence $\{\xi_t\}_1^\infty$ is defined on the probability space $(X, \mathscr{F}, P)$ and is assumed to be strictly stationary and ergodic with zero mean and finite variance. $\{\xi_t\}_1^\infty$ satisfies the following regularity conditions:

Assumption 1. *ξ_t is a stationary ARMA process with $\xi_t = \sum_{j=0}^{\infty} C_j \nu_{t-j}$, $C_0 = I_n$, $\sum_{j=0}^{\infty} j \left\| C_j \right\| < \infty$ and $\nu_t \sim iid(0, \Sigma)$, where Σ is a positive definite variance matrix and ν_t have absolutely continuous distribution*[3]. *Further, $E|\nu_t|^r < \infty$ for some $r \geq 4$.*

The partial sum process constructed from $\{\xi_t\}$ satisfies the functional central limit theorem (FCLT) for Reyni-mixing processes, described in Hall and Heyde (1980). For $s \in [0,1]$ and as $T \to \infty$, it holds that

$$X_T(s) = T^{-1/2} \sum_{t=1}^{[Ts]} \xi_t \Rightarrow B(s), \tag{4.7}$$

where $B(s)$ is $(m+1)$-vector Brownian motion with covariance matrix

$$\Omega = \lim_{T \to \infty} T^{-1} E \left(\left(\sum_{t=1}^{T} \xi_t \right) \left(\sum_{t=1}^{T} \xi_t' \right) \right). \tag{4.8}$$

We partition $z_t = (y_t, x_t')'$ into the scalar variate y_t and the m-vector x_t with conformable partitions of Ω and B:

$$B = \begin{bmatrix} B_y \\ B_x \end{bmatrix} \qquad \Omega = \begin{bmatrix} \omega_{11} & \omega_{21}' \\ \omega_{21} & \Omega_{22} \end{bmatrix}. \tag{4.9}$$

We assume $\Omega_{22} > 0$ and decompose Ω as $\Omega = L'L$, where L is given by

$$L = \begin{bmatrix} l_{11} & 0 \\ l_{21} & L_{22} \end{bmatrix}, \tag{4.10}$$

with $l_{11} = (\omega_{11} - \omega_{21}' \Omega_{22}^{-1} \omega_{21})^{1/2}$, $l_{21} = \Omega_{22}^{-1/2} \omega_{21}$, and $L_{22} = \Omega_{22}^{1/2}$. Further, we define $W(s)$ to be $(m+1)$-vector standard Brownian motion and from Lemma 2.2 of Phillips and Ouliaris (1990) it follows that $B = L'W$.

Residual-based cointegration tests seek to test the null hypothesis of no cointegration using unit root tests applied to the residuals of the cointegrating regression. Hence, we estimate the cointegrating regression according to one of the structural break models (4.3) using least squares and apply the SETAR model (4.4) to the residuals $\hat{e}_{t\tau}$,

$$\Delta \hat{e}_{t\tau} = \rho_1 \hat{e}_{t-1\tau} \mathbb{1}\{\hat{e}_{t-1\tau} \geq \lambda\} + \rho_2 \hat{e}_{t-1\tau} \mathbb{1}\{\hat{e}_{t-1\tau} < \lambda\} + \sum_{j=1}^{K} \gamma_j \Delta \hat{e}_{t-j\tau} + \varepsilon_{t\tau}, \tag{4.11}$$

[3] A stationary ARMA process is not necessarily strong-mixing. But if the innovations have absolutely continuous distribution, the strong-mixing condition is ensured (see, for example Andrews (1984) and Mokkadem (1988)).

given that the threshold parameter λ is known, i.e. a fixed value. The residual regression (4.11) depends on the relative timing of the breakpoint parameter τ. We assume the lag order K in (4.11) to be large enough to capture the correlation structure of the errors. Since the error term $\varepsilon_{t\tau}$ might have a nonzero MA component, it is necessary to increase K with the sample size ($K \to \infty$ as $T \to \infty$). We follow Said and Dickey (1984) and state:

Assumption 2. *K increases with T in such a way that $K = o(T^{1/3})$.*

Since $\mathbb{1}\{\hat{e}_{t-1\tau} \geq \lambda\}$ and $\mathbb{1}\{\hat{e}_{t-1\tau} < \lambda\}$ are orthogonal, the test statistic is given by

$$F_\tau = \frac{t_1^2 + t_2^2}{2}, \tag{4.12}$$

where t_1 and t_2 are the t ratios for $\hat{\rho}_1$ and $\hat{\rho}_2$ from regression (4.11). F_τ is computed for each possible break fraction $\tau \in \mathscr{T}$ and the sup$F$-statistic, F^*, is computed to evaluate the null hypothesis of no cointegration against the alternative of threshold cointegration with possible structural break. The following theorem presents the asymptotic distributions of F^* for model specifications C, C/T and C/S:

Theorem 1. *If $\{z_t\}_0^\infty$ is generated by (4.6), Assumptions (1) and (2) hold, the threshold parameter λ is fixed and τ belongs to a compact subset of $(0,1)$, then as $T \to \infty$*

$$F^* \Rightarrow \frac{1}{2}\sup_{\tau\in\mathscr{T}}\left\{\frac{\left(\int_0^1 \mathbb{1}\{Q_{\kappa\tau} \geq 0\}Q_{\kappa\tau}dQ_{\kappa\tau}\right)^2}{\kappa_\tau' D_\tau \kappa_\tau \int_0^1 \mathbb{1}\{Q_{\kappa\tau} \geq 0\}Q_{\kappa\tau}^2} + \frac{\left(\int_0^1 \mathbb{1}\{Q_{\kappa\tau} < 0\}Q_{\kappa\tau}dQ_{\kappa\tau}\right)^2}{\kappa_\tau' D_\tau \kappa_\tau \int_0^1 \mathbb{1}\{Q_{\kappa\tau} < 0\}Q_{\kappa\tau}^2}\right\}$$

where

$$\begin{aligned} Q_{\kappa\tau} &= W_y - \left(\int_0^1 W_{x\tau}W_{x\tau}'\right)^{-1}\left(\int_0^1 W_y W_{x\tau}'\right)W_{x\tau} \\ \kappa_\tau &= \left(1, -\left(\int_0^1 W_{x\tau}W_{x\tau}'\right)^{-1}\left(\int_0^1 W_y W_{x\tau}'\right)\right) \end{aligned}$$

Under the alternative of cointegration with two-regime SETAR adjustment, $F^* \to \infty$ as $T \to \infty$. $Q_{\kappa\tau}$ depends on the model:

a) If the residuals are obtained from least squares estimation of model C, then

$$W_{x\tau} = (W_x', 1, \varphi_\tau)'$$

$$D_\tau = \begin{bmatrix} I_{m+1} & 0 \\ 0 & 0 \end{bmatrix}.$$

b) If the residuals are obtained from least squares estimation of model C/T, then

$$W_{x\tau} = (W_x', 1, s, \varphi_\tau)'$$

$$D_\tau = \begin{bmatrix} I_{m+1} & 0 \\ 0 & 0 \end{bmatrix}.$$

c) If the residuals are obtained from least squares estimation of model C/S, then

$$W_{x\tau} = (W_x', 1, W_x'\varphi_\tau, \varphi_\tau)'$$

$$D_\tau = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & I_m & 0 & (1-\tau)I_m & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & (1-\tau)I_m & 0 & (1-\tau)I_m & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

A formal proof of the theorem is provided in the Appendix.

4 Simulation results

Critical values and finite sample properties of the sup$F$ test are examined by Monte Carlo experiments. In the absence of a structural break, we use a DGP according to Engle and Granger (1987) and Banerjee *et al.* (1986) which is given for one regressor ($m = 1$) in the form of

$$\begin{array}{lllllll} y_t & = & \mu + \alpha x_{1,t} + e_t & \Delta e_t & = & \rho e_{t-1} + \vartheta_t & \vartheta_t \sim N(0,1) \\ y_t & = & x_{1,t} + \eta_t & \eta_t & = & \eta_{t-1} + \omega_t & \omega_t \sim N(0,1) \end{array} \tag{4.13}$$

where the parameters of the equilibrium equation are $\mu = 1$ and $\alpha = 2$. First, the null hypothesis of no cointegration is simulated with $\rho = 0$. This enables us to obtain quantiles of the sup$F$ distribution for different sample sizes. Critical values are computed for 10,000 draws for each sample size. The results are reported in Table 4.1.

Table 4.1: Approximate critical values of F^*

		C			C/T			C/S		
	T	90%	95%	99%	90%	95%	99%	90%	95%	99%
$K=0$										
	50	10.55	12.12	15.28	12.85	14.37	18.11	14.76	16.70	20.98
	100	10.21	11.44	14.18	12.26	13.63	16.65	13.90	15.51	18.94
	250	9.94	11.04	14.11	11.84	13.11	15.78	13.57	15.05	18.10
	500	9.81	10.98	13.64	11.74	12.98	15.59	13.59	14.93	17.90
$K=1$										
	50	10.09	11.50	14.98	12.00	13.68	17.47	11.27	12.84	16.33
	100	10.07	11.24	14.26	11.88	13.24	16.24	11.13	12.65	15.68
	250	10.00	11.10	13.72	11.68	12.93	15.65	11.18	12.47	15.36
	500	9.85	11.00	13.56	11.74	12.91	15.60	11.28	12.67	15.24
$K=4$										
	50	8.58	9.84	12.84	9.71	11.14	14.24	9.16	10.40	13.43
	100	9.09	10.26	12.68	10.53	11.68	14.51	9.87	11.06	13.55
	250	9.42	10.68	13.05	11.16	12.33	14.62	10.58	11.90	14.33
	500	9.64	10.83	13.15	11.42	13.00	15.18	10.90	12.18	14.72

Note: C, C/T and C/S denote the structural break models in (4.3). K refers to the number of lags in (4.4).

Table 4.2: Approximate critical values of F^* for more than one regressor

	T	C 90%	C 95%	C 99%	C/T 90%	C/T 95%	C/T 99%	C/S 90%	C/S 95%	C/S 99%
$m=2$										
	50	12.53	14.25	18.00	14.66	16.50	20.41	15.68	17.69	22.33
	100	12.05	13.60	16.69	13.91	15.54	18.97	14.88	16.61	20.38
	250	11.68	13.00	15.69	13.39	14.91	17.89	14.40	16.09	19.36
	500	11.54	12.76	15.52	13.25	14.62	17.48	14.21	15.67	18.89
$m=3$										
	50	14.63	16.53	20.74	16.63	18.70	22.91	19.30	21.65	27.31
	100	13.74	15.21	18.63	15.52	17.20	20.40	17.94	19.74	23.83
	250	13.29	14.63	17.88	14.99	16.58	19.71	17.31	19.05	22.59
	500	13.12	14.41	17.37	14.70	16.04	18.85	17.06	18.60	22.20
$m=4$										
	50	16.35	18.46	22.91	18.44	20.70	25.49	22.91	25.60	32.01
	100	15.38	17.14	20.71	17.22	19.13	23.20	21.04	23.24	27.78
	250	14.97	16.45	19.62	16.62	18.10	21.35	20.18	21.96	25.64
	500	14.64	16.05	19.16	16.30	17.65	20.78	19.80	21.66	25.32

Note: C, C/T and C/S denote the structural break models in (4.3). m refers to the number of columns of the regressor matrix x_t.

The power of the sup$F$ test under structural change is evaluated with a DGP designed in line with Gregory and Hansen (1996a). A slight modification was, however, necessary to allow for SETAR adjustment to the long-run equilibrium. The following DGP is employed for a bivariate cointegrated system,

$$\begin{aligned} y_t &= \mu_t + \alpha_t x_{1,t} + e_t \qquad & \Delta e_t &= \begin{cases} \rho_1 e_{t-1} + \vartheta_t & \text{if } e_{t-1} \geq 0 \\ \rho_2 e_{t-1} + \vartheta_t & \text{if } e_{t-1} < 0 \end{cases} \qquad & \vartheta_t \sim N(0,1) \\ y_t &= x_{1,t} + \eta_t & \eta_t &= \eta_{t-1} + \omega_t & \omega_t \sim N(0,1) \end{aligned} \tag{4.14}$$

$$\begin{bmatrix} \mu_t = \mu_1, & \alpha_t = \alpha_1, & t \leq [T\tau] \\ \mu_t = \mu_2, & \alpha_t = \alpha_2, & t > [T\tau] \end{bmatrix}$$

in which symmetric adjustment is nested as $\rho_1 = \rho_2$. A change in the intercept is modelled by means of an increase from $\mu_1 = 1$ to $\mu_2 = 4$ at the breakpoint $[T\tau]$, whereas a change in the slope is modelled as an increase from $\alpha_1 = 2$ to $\alpha_2 = 4$. The simulation set-up used for cointegrated systems with symmetric adjustment directly follows Gregory and Hansen (1996a) so that the results for the sup$F$ test can be compared with the results for the GH test.

Table 4.4 displays the frequency of rejection under structural stability and asymmetric adjustment, i.e. how often the sup$F$ test rejects the null hypothesis of no cointegration for a given combination of autoregressive coefficients. For each pair of coefficients the series was generated with sample size $T = 100$. The process was replicated 2,500 times for every specification. If the series are generated under asymmetric adjustment with a stable cointegrating vector, we find that the sup F test operates with less power than the threshold cointegration test by Enders and Siklos (2001). Falsely incorporating breaks in form of additional dummy variable in the equilibrium equation thus reduces the power against the null hypothesis. Accordingly, the most parsimonious model C performs best among the three structural break models.

Table 4.5 presents the power performance under cointegration with symmetric adjustment and break in either the intercept or slope. We can see that the $\sup F$ test has generally higher rejection frequencies than either the Engle-Granger test using the ADF test statistic or the threshold cointegration test without breakpoint estimation. The simulation reveals that the $\sup F$ test is as powerful as the GH test. The SETAR Enders-Siklos test seems to be rather robust to a break in the intercept but suffers from a drastic reduction in power if a break in the slope is considered. The $\sup F$ test shows sufficient power at sample sizes above $T = 100$ and moderate adjustment rate $\rho = -0.5$. As expected, the model C outperforms model C/T and C/S if a break in the intercept is considered, while C/S performs best if the slope changes at one point in the sample.

The simulation results under symmetric adjustment can also be used to analyze the estimation accuracy of the pre-specified breakpoint in the DGP. The timing of the break is varied and takes place either at the beginning, the middle or near the end of the series. The results are summarized in Table 4.6 and reveal that breakpoint estimates are in large parts very accurate. In general, it seems that a break at the beginning ($\lambda = 0.25$) is the most difficult to detect and the $\sup F$ test often indicates a later breakpoint. Breaks in the intercept and the slope are estimated with equal accuracy as long as the correct structural break model is applied.

Finally, the behaviour of the $\sup F$ test is evaluated under parameter instability and asymmetric adjustment. For that matter, we draw from the DGP in (4.14) using a subset of the parameter combinations displayed in Table 4.4. In the first panel of Table 4.7, we consider a break in the intercept. The $\sup F$ test shows dismal power properties and is outperformed by the SETAR Enders-Siklos test in each parameter combination. The loss in power of the original threshold cointegration test due to a break in the intercept does not justify the additional parameter estimation and grid search of the C model. The C/T and C/S models involve an additional parameter and, as expected, have lower rejection frequencies. With a break in the slope (second panel of Table 4.7), we find the picture to be quite different. All structural break models have more power against the null hypothesis than the SETAR Enders-Siklos test. While the C models performs slightly better than the correctly specified C/S model for weak adjustment, the power of the C/S model exceeds all others under moderate adjustment. In the third panel of Table 4.7, we display the results for a break in the intercept and the slope. Again, the C/S model performs best among the structural break models and far exceeds the benchmark. In general, we find a break in the slope to have a more substantial impact on the power function than a break in the intercept. Since structural change most likely involves all parameters of the equilibrium equation and the $\sup F$ test based on the C/S model performs best in those situations, it has to be considered the preferred model for cointegration relationships with asymmetric adjustment which are subject to parameter instability.

5 Empirical application

In this section, we apply the $\sup F$ test methodology to study the 'rockets and feathers' hypothesis[4] in the US gasoline market. The 'rockets and feathers' hypothesis describes the adjustment behaviour of prices faced with input price shocks. More precisely, the hypothesis states that prices adjust faster to input

[4]The name originates from the Bacon (1991) paper entitled: 'Rockets and feathers: the asymmetric speed of adjustment of UK retail gasoline prices to cost changes'

price increases than to input price decreases. In the terms of Bacon (1991)'s seminal paper, the price goes up like a rocket, but falls down like a feather. While early studies on the matter (Bacon (1991), Manning (1991), Borenstein *et al.* (1997)) focused on the short-run asymmetry in the pricing process, the focus quickly shifted to the economically meaningful long-run asymmetry estimated by asymmetric error correction models (Bachmeier and Griffin (2003)). We demonstrate the capabilities of the $\sup F$ test using US gasoline data over a span that covers the Financial Crisis from 2008.

Crude oil passes different stages of processing and distribution until it reaches the end customer. For the analysis, we examine the prices transmission at two points of the production chain. First, we analyze the speed of adjustment for deviations from the long-run relationship between crude oil prices and gasoline spot prices (*first stage*), i.e. the relationship between pre- and post-refinement prices. Second, we analyze the pass-through from gasoline spot prices to retail prices (*second stage*). Finally, the direct link between crude oil prices and retail prices is analyzed (*single stage*). Naturally, we expect the speed of adjustment at the first and second stage to be faster than at the single stage transmission. Asymmetry in the sense of the 'rockets and feathers' hypothesis is found if negative deviations from the long-run equilibrium are adjusted faster than positive deviations, i.e. $\rho_1 = \rho^- < \rho^+ = \rho_2$.

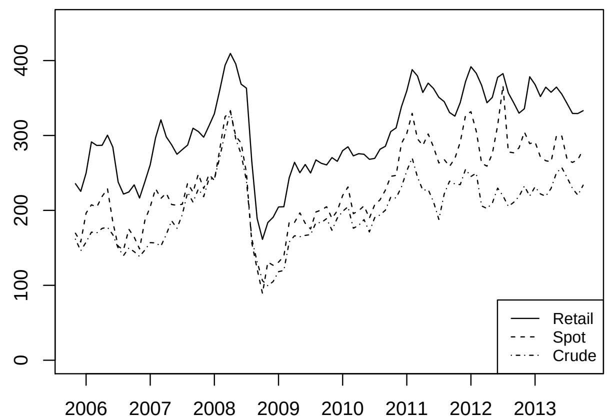


Figure 4.1: WTI crude oil prices, spot gasoline prices and retail gasoline prices from January 2006 to December 2013

Our sample reaches from January 2006 to December 2013 to include the collapse of commodity prices in 2009 and their subsequent recovery. We observe prices at a monthly frequency yielding a total of 96 observations. The West Texas Intermediate prices (crude), regular gasoline spot prices (spot) and regular gasoline retail prices (retail) are all obtained from the U.S. Energy Information Administration (EIA). Figure 4.1 depicts the trajectory of the prices and shows volatile behaviour of prices for petroleum products during the Financial Crisis. Although the times series are affected by global events, it does not immediately follow that the long-run relationship between them changes. However, from our simulation study, we know that an existing instability of the cointegrating vector can severely decrease the power of a threshold cointegration test. A closer inspection of the trajectories reveals a larger margin between crude oil and gasoline spot prices in the later part of the sample.

First, we estimate a threshold cointegration model according to Enders and Siklos (2001). We specify the long-run equilibrium equations

$$\begin{array}{lccl} (I) & \text{spot}_t & = & \mu + \alpha\, \text{crude}_t + e_t \\ (II) & \text{retail}_t & = & \mu + \alpha\, \text{spot}_t + e_t \\ (S) & \text{retail}_t & = & \mu + \alpha\, \text{crude}_t + e_t \end{array} \tag{4.15}$$

where the (I), (II), (S) denote *first stage*, *second stage* and *single stage*, respectively. The coefficients of the cointegrating vector are estimated using least squares and the SETAR model is applied to the residuals. The results are reported in panel (a) of Table 4.3. The adjustment coefficients show the expected signs but do not reveal significant asymmetry in the adjustment process. Surprisingly, we do not find sufficient evidence for a long-run relationship between crude oil prices and gasoline spot prices. In contrast, retail gasoline prices and crude oil prices seem to maintain a long-run equilibrium which is a less likely result from an economic perspective than the existence of a crude/spot relationship.

Table 4.3: Long-run adjustment along the gasoline value-chain

	Panel (a): No structural break									
	μ		α			ρ^+	ρ^-	Φ_{SETAR}	$\rho^+ = \rho^-$	
(I)	5.492		1.145			-0.242	-0.181	4.922	-	
(II)	79.499		0.960			-0.567	-0.887	17.830***	2.581	
(S)	76.376		1.141			-0.250	-0.326	6.727**	0.263	
	Panel (b): Structural break model C									
	μ_1	μ_2	α			ρ^+	ρ^-	$\sup F$	$\rho^+ = \rho^-$	break
(I)	35.385	41.263	0.916			-0.516	-0.490	13.281**	0.021	12/10
(II)	79.438	-6.778	0.978			-0.621	-0.964	20.561***	2.921*	01/09
(S)	103.287	37.904	1.028			-0.445	-0.497	12.627**	0.095	02/11
	Panel (c): Structural break model C/T									
	μ_1	μ_2	α		δ	ρ^+	ρ^-	$\sup F$	$\rho^+ = \rho^-$	break
(I)	40.094	50.459	0.923		-0.199	-0.561	-0.522	14.821**	0.049	12/10
(II)	75.814	-3.736	0.982		-0.001	-0.584	-0.948	19.567***	3.317*	09/10
(S)	110.132	51.776	0.951		-0.302	-0.527	-0.527	15.087**	0.000	02/11
	Panel (d): Structural break model C/S									
	μ_1	μ_2	α_1	α_2		ρ^+	ρ^-	$\sup F$	$\rho^+ = \rho^-$	break
(I)	32.176	100.412	0.935	-0.256		-0.561	-0.474	14.553**	0.234	02/11
(II)	60.491	22.381	1.062	-0.123		-0.630	-1.018	21.404***	3.696*	10/08
(S)	93.796	195.438	0.989	-0.698		-0.453	-0.588	14.880**	0.621	02/11

Note: μ (α) denotes the intercept (slope coefficient) of the long-run equilibrium equation without structural break. μ_1 (α_1) and μ_2 (α_2) denote the intercept (slope coefficient) of the long-run equilibrium equation before the break and after the break, respectively. δ is the linear trend coefficient. Φ_{SETAR} denotes the F-statistic based on the null hypothesis $H_0 : \rho^+ = \rho^- = 0$. We conduct F tests to test the null hypothesis $\rho^+ = \rho^-$.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Second, we estimate the long-run equilibrium equations again with each of the three structural break models. We put emphasis on the results of the C/S specification since this specification of the $\sup F$ test performed best in the simulation study if the slope coefficient changed at one point in time and is best-suited for modelling unspecific regime shift events. The results are reported in panel (b)-(d) of Table 4.3. The null hypothesis of no cointegration can now be rejected at all stages along the gasoline value-chain.

The breakpoint is located either at the beginning of the crisis, i.e. at the peak crude oil prices, or after the prices had recovered in 2011. We do not find statistical evidence for asymmetric adjustment processes in the gasoline value-chain whether we model a structural break or not.

6 Summary

This paper proposed an extension to the GH test to include SETAR adjustment. Thereby, we constructed a threshold cointegration test which endogenously determines the location of a structural break in the cointegrating vector and tests the null of no cointegration. We derived the limiting distribution for the structural break models C, C/T and C/S and tabulated their critical values which were obtained by Monte Carlo simulations. Analysis of the finite sample properties under the alternative of linear and threshold cointegration revealed that the test exhibits considerable power gains over the conventional SETAR Enders-Siklos test if a break in the slope coefficient is present. We applied the $\sup F$ test to US gasoline market data and found evidence for a long-run relationship between prices along the value-chain after we accounted for structural breaks. None of the models we estimated provided sufficient evidence for asymmetric price transmissions.

B Appendices

B .1 Mathematical proofs

Proof of Theorem 1. The asymptotic distribution is derived by adapting the results of Gregory and Hansen (1992) to match the F-statistic process involving a threshold indicator function using results in Maki and Kitasaka (2015). However, Maki and Kitasaka (2015) use a different definition of the threshold parameter space. The threshold parameter in our model is fixed, i.e. belongs to a trivial compact subset of $\mathbb{R}$ whereas the parameter space in Maki and Kitasaka (2015) is data dependent (see the discussion on threshold parameter space in Section 2.2 of their paper). Indicator functions with threshold parameters defined on compact sets are treated in Seo (2008). The proof only refers to model C/S while the results for the remaining models can be deduced from the results obtained for this model. Hence, we consider the cointegrating regression,

$$y_t = \hat{\alpha}_1' x_t + \hat{\mu}_1 + \hat{\alpha}_2' x_t \varphi_{t,\tau} + \hat{\mu}_2 \varphi_{t,\tau} + \hat{e}_{t\tau}, \tag{4.16}$$

where $\hat{e}_{t\tau}$ is an integrated process under the null hypothesis of no cointegration and $z_t = (y_t, x_t')'$ is generated according to (4.6).

Define the $(2m+3)$-vector $X_{t\tau} = (y_t, x_t, 1, x_t\varphi_{t,\tau}, \varphi_{t,\tau})'$ and partition $X_{t\tau} = (X_{1t\tau}, X_{2t\tau})'$ where $X_{1t\tau} = y_t$ and $X_{2t\tau}$ contains all regressors of (4.16). Define $\delta_T = diag(T^{-1/2}I_{m+1}, 1, T^{-1/2}I_m, 1)$, $\varphi_\tau(s) = \mathbb{1}\{s > \tau\}$ and $X_\tau(s) = (B(s)', 1, B_x(s)'\varphi_\tau(s), \varphi_\tau(s))'$. Partition $\delta_T = (\delta_{1T}, \delta_{2T})$ in conformity to $X_{t\tau}$.

Next, we partition the $(m+1)$-vector standard Brownian Motion W as $W = (W_y, W_x')'$ where

$$\begin{aligned} W_y &= l_{11}^{-1}\left(B_y - \omega_{21}'\Omega_{22}^{-1}B_x\right) \\ W_x &= \Omega_{22}^{-1/2}B_x. \end{aligned} \tag{4.17}$$

Furthermore, we define

$$W_{x\tau} = (W_x', 1, W_x'\varphi_\tau, \varphi_\tau)' \tag{4.18}$$

and $W_\tau = (W_y, W_{x\tau}')'$.

First, we consider the least squares estimator of the parameters of the cointegrating regression. It is shown in Gregory and Hansen (1992) using the FCLT for vector processes in Phillips and Durlauf (1986) and the continuous mapping theorem (CMT, see Billingsley (1999), Theorem 2.7) that

$$T^{-1}\delta_T \sum_{t=1}^{T} X_{t\tau}X_{t\tau}'\delta_T \Rightarrow \int_0^1 X_\tau X_\tau' \tag{4.19}$$

uniformly over τ.

We define the vector $\hat{\theta}_\tau = (\hat{\alpha}_1', \hat{\mu}_1, \hat{\alpha}_2', \hat{\mu}_2)$ as the least squares estimator of (4.16) for each τ. It follows from (4.19) and the CMT that

$$\begin{aligned} T^{-1/2}\delta_{2T}{}^{-1}\hat{\theta}_\tau &= \left(T^{-1}\delta_{2T}\sum_{t=1}^{T}X_{2t\tau}X_{2t\tau}{}'\delta_{2T}\right)^{-1}\left(T^{-1}\delta_{2T}\sum_{t=1}^{T}X_{2t\tau}X_{1t\tau}{}'\delta_{1T}\right) \\ &\Rightarrow \left(\int_0^1 X_{2\tau}X_{2\tau}{}'\right)^{-1}\left(\int_0^1 X_{2\tau}X_{1\tau}{}'\right). \end{aligned} \tag{4.20}$$

When we set $\hat{\eta}_\tau = T^{-1/2}\delta_T{}^{-1}(1,-\hat{\theta}_\tau')' = (1,-\delta_{2T}{}^{-1}\hat{\theta}_\tau')'$, it follows that

$$\hat{\eta}_\tau \Rightarrow \left(1,-\left(\int_0^1 X_{2\tau}X_{2\tau}{}'\right)^{-1}\left(\int_0^1 X_{1\tau}X_{2\tau}{}'\right)\right) = \eta_\tau. \tag{4.21}$$

Next, we state some useful convergence results for the residuals of the cointegrating regression. We define the residual series $\hat{e}_{t\tau} = y_t - \hat{\alpha}_1'x_t - \hat{\mu}_1 - \hat{\alpha}_2'x_t\varphi_{t,\tau} - \hat{\mu}_2\varphi_{t,\tau}$ which is dependent on τ. Note that $\hat{e}_{t\tau}$ can be expressed as

$$\hat{e}_{t\tau} = T^{1/2}\hat{\eta}_\tau'\delta_T X_{t\tau}. \tag{4.22}$$

Using Lemma 2.2 of Phillips and Ouliaris (1990) yields

$$T^{-1/2}\hat{e}_{t\tau} \Rightarrow \eta_\tau' X_\tau = l_{11}\kappa_\tau W_\tau = l_{11}Q_{\kappa\tau}, \tag{4.23}$$

where

$$\begin{aligned} \kappa_\tau &= \left(1,-\left(\int_0^1 W_{x\tau}W_{x\tau}'\right)^{-1}\left(\int_0^1 W_y W_{x\tau}'\right)\right) \\ L\eta_\tau &= l_{11}\kappa_\tau \\ Q_{\kappa\tau} &= W_y - \left(\int_0^1 W_{x\tau}W_{x\tau}'\right)^{-1}\left(\int_0^1 W_y W_{x\tau}'\right)W_{x\tau}. \end{aligned} \tag{4.24}$$

The first-differenced residuals are expressed as $\Delta\hat{e}_{t\tau} = \hat{\eta}_\tau'\Delta X_{t\tau}$, where

$$\begin{aligned} \Delta X_{t\tau} &= \Delta(y_t, x_t, 1, x_t\varphi_{t,\tau}, \varphi_{t,\tau})' \\ &= (\xi_{1t}, \xi_{2t}, 0, x_{t-1}\Delta\varphi_{t,\tau} + \Delta x_t\varphi_{t,\tau}, \Delta\varphi_{t,\tau})' \\ &= (\xi_{1t}, \xi_{2t}, 0, x_{t-1}\Delta\varphi_{t,\tau} + \xi_{2t}\varphi_{t,\tau}, \Delta\varphi_{t,\tau})' \end{aligned} \tag{4.25}$$

and

$$\Delta\varphi_{t,\tau} = \begin{cases} 1 & \text{if } t = [T\tau] \\ 0 & \text{if } t \neq [T\tau] \end{cases}. \tag{4.26}$$

The asymptotic counterpart to $\Delta\varphi_{t,\tau}$ is the differential $d\varphi_\tau$, a Dirac function concentrating the unit mass at the point $t = \tau$ so that

$$\int_0^1 f d\varphi_\tau = \lim_{z \uparrow \tau} f(z)$$

for all functions with left-limits. Then, it holds that $\Delta \hat{e}_{t\tau} \Rightarrow \eta_\tau' \Delta X_\tau$, where

$$\Delta X_\tau(s) = (dB(s)', 0, B_x(s)' d\varphi_\tau(s) + dB_x(s)' \varphi_\tau(s), d\varphi_\tau(s))'. \tag{4.27}$$

Under Assumption 1, ξ_t is a stationary VARMA process and consequently, the scalar process $\eta_\tau' \Delta X_{t\tau}$ is also a stationary ARMA process except for a point mass at $t = [T\tau]$. Following Phillips and Ouliaris (1990) we write the AR representation of the SETAR error term process as $\varepsilon_{t\tau} = \sum_{j=0}^{\infty} D_j (\Delta X_{t-j\tau})' \eta_\tau = D(L)(\Delta X_{t\tau})' \eta_\tau$. Under Assumption 2, the lag structure is chosen in a way that $\varepsilon_{t\tau}$ is an orthogonal $(0, \sigma^2(\eta, \tau))$ sequence with $\sigma^2(\eta, \tau) = D(1)^2 \eta_\tau' \Omega_\tau \eta_\tau$. From Lemma 2.1 of Phillips and Ouliaris (1990), it follows that

$$T^{-1/2} \sum_{t=1}^{[Ts]} \varepsilon_{t\tau} = D(L) \eta_\tau' \left(\sum_{t=1}^{[Ts]} \delta_T \Delta X_{t\tau} \right) + o_p(1) \Rightarrow D(1) \eta_\tau' X_\tau(s) \tag{4.28}$$

for each τ, where $D(1) = \sum_{j=0}^{\infty} D_j$.

Now, we consider the auxiliary regression. We apply the SETAR model to the residuals according to (4.11) and compute the test statistics F_τ. Note that the estimated adjustment coefficients might be correlated with the estimated coefficients of the additional lagged differences. Therefore, we write the least squares estimator of $\rho = (\rho_1, \rho_2)'$ in the breakpoint-specific notation under the null hypothesis $\rho_1 = \rho_2 = 0$ as $\hat{\rho} = (U_\tau' Q_K U_\tau)^{-1} U_\tau' Q_K \varepsilon_\tau$, where

$$U_\tau = \begin{bmatrix} \hat{e}_{0\tau} \mathbb{1}\{\hat{e}_{0\tau} \geq \lambda\} & \hat{e}_{0\tau} \mathbb{1}\{\hat{e}_{0\tau} < \lambda\} \\ \hat{e}_{1\tau} \mathbb{1}\{\hat{e}_{1\tau} \geq \lambda\} & \hat{e}_{1\tau} \mathbb{1}\{\hat{e}_{1\tau} < \lambda\} \\ \vdots & \vdots \\ \hat{e}_{T-1\tau} \mathbb{1}\{\hat{e}_{T-1\tau} \geq \lambda\} & \hat{e}_{T-1\tau} \mathbb{1}\{\hat{e}_{T-1\tau} < \lambda\} \end{bmatrix}, \tag{4.29}$$

$\varepsilon_\tau = (\varepsilon_{1\tau}, \varepsilon_{2\tau}, \ldots, \varepsilon_{T\tau})'$ and $Q_K = I - M_K(M_K' M_K)^{-1} M_K'$ is the projection matrix onto the space orthogonal to $M_K = (\Delta \hat{e}_{t-1\tau}, \ldots, \Delta \hat{e}_{t-K\tau})$.

Partition the matrix U_τ as $U_\tau = (U_{1\tau}, U_{2\tau})$, then the t ratio of $\hat{\rho}_1$ can be expressed as

$$t_1 = \frac{\hat{\rho}_1}{se(\hat{\rho}_1)} = \frac{\hat{\rho}_1}{(\hat{\sigma}^2 (U_{1\tau}' Q_K U_{1\tau})^{-1})^{1/2}} = \frac{U_{1\tau}' Q_K \varepsilon_\tau}{\hat{\sigma} (U_{1\tau}' Q_K U_{1\tau})^{1/2}} \tag{4.30}$$

and similarly the t ratio of $\hat{\rho}_2$ can be expressed as

$$t_2 = \frac{U_{2\tau}' Q_K \varepsilon_\tau}{\hat{\sigma} (U_{2\tau}' Q_K U_{2\tau})^{1/2}}. \tag{4.31}$$

In the remainder of the proof, we focus on t_1. Scaling the t ratio appropriately yields the numerator

$$\begin{aligned} T^{-1} U_{1\tau}' Q_K \varepsilon_\tau &= T^{-1} U_{1\tau}' \varepsilon_\tau - T^{-1/2} \cdot T^{-1} U_{1\tau}' M_K (T^{-1} M_K' M_K)^{-1} T^{-1/2} M_K' \varepsilon_\tau \\ &= T^{-1} U_{1\tau}' \varepsilon_\tau + o_p(1) = N_T(\lambda, \tau) + o_p(1) \end{aligned} \tag{4.32}$$

and the term

$$\begin{aligned} T^{-2}U_{1\tau}'Q_K U_{1\tau} &= T^{-2}U_{1\tau}'U_{1\tau} - T^{-1}\cdot T^{-1}U_{1\tau}'M_K(T^{-1}M_K'M_K)^{-1}T^{-1}M_K'U_{1\tau} \\ &= T^{-2}U_{1\tau}'U_{1\tau} + o_p(1) = D_T(\lambda,\tau) + o_p(1). \end{aligned} \tag{4.33}$$

Finally, we need convergence results for $N_T(\lambda,\tau)$, $D_T(\lambda,\tau)$ and $\hat{\sigma}^2$. From (4.23) and since $x \mapsto x\mathbb{1}\{x \geq \lambda\}$ is a regular function, it follows from Theorem 3.1 of Park and Phillips (2001) that

$$\begin{aligned} T^{-1/2}\hat{e}_{t-1\tau}\mathbb{1}\{\hat{e}_{t-1\tau} \geq \lambda\} &= \hat{\eta}_\tau'\delta_T X_{t-1\tau}\mathbb{1}\{T^{1/2}\hat{\eta}_\tau'\delta_T X_{t-1\tau} \geq \lambda\} \\ &= \hat{\eta}_\tau'\delta_T X_{t-1\tau}\mathbb{1}\{\hat{\eta}_\tau'\delta_T X_{t-1\tau} \geq T^{-1/2}\lambda\} \\ &\Rightarrow \eta_\tau' X_\tau \mathbb{1}\{\eta_\tau' X_\tau \geq 0\} = l_{11}Q_{\kappa\tau}\mathbb{1}\{Q_{\kappa\tau} \geq 0\}. \end{aligned} \tag{4.34}$$

Thus, Theorem 2.2 of Kurtz and Protter (1991) combined with results (4.28) and (4.34) yields

$$\begin{aligned} N_T(\lambda,\tau) &= T^{-1}\sum_{t=1}^{T}\mathbb{1}\{\hat{e}_{t-1\tau} \geq \lambda\}\hat{e}_{t-1\tau}\varepsilon_{t\tau} \\ &= \hat{\eta}_\tau'\delta_T\sum_{t=1}^{T}\mathbb{1}\{\delta_T\hat{\eta}_\tau' X_{t-1\tau} \geq T^{-1/2}\lambda\}X_{t-1\tau}D(L)(\Delta X_{t\tau})'\delta_T\eta_\tau \\ &\Rightarrow D(1)\eta_\tau'\int_0^1 \mathbb{1}\{\eta_\tau' X_\tau \geq 0\}X_\tau dX_\tau'\eta_\tau \\ &= D(1)l_{11}^2\int_0^1 \mathbb{1}\{Q_{\kappa\tau} \geq 0\}Q_{\kappa\tau}dQ_{\kappa\tau}, \end{aligned} \tag{4.35}$$

while (4.28), (4.34) and the CMT yield

$$\begin{aligned} D_T(\lambda,\tau) &= T^{-2}\sum_{t=1}^{T}\mathbb{1}\{\hat{e}_{t-1\tau} \geq \lambda\}\hat{e}_{t-1\tau}^2 \\ &= \hat{\eta}_\tau'\delta_T T^{-1}\sum_{t=1}^{T}\mathbb{1}\{\delta_T\hat{\eta}_\tau' X_{t-1\tau} \geq T^{-1/2}\lambda\}X_{t-1\tau}X_{t-1\tau}{}'\delta_T\hat{\eta}_\tau \\ &\Rightarrow \eta_\tau'\int_0^1 \mathbb{1}\{\eta_\tau' X_\tau \geq 0\}X_\tau X_\tau'\eta_\tau \\ &= l_{11}^2\int_0^1 \mathbb{1}\{Q_{\kappa\tau} \geq 0\}Q_{\kappa\tau}^2. \end{aligned} \tag{4.36}$$

For the variance estimate, $\hat{\sigma}^2$, we note that $\hat{\rho}_1 = O_p(T^{-1})$ and $\hat{\rho}_2 = O_p(T^{-1})$, but $(\hat{\gamma}_j - \gamma_j) = O_p(T^{-1/2})$.

Using Lemma 2.2 of Phillips and Ouliaris (1990) yields

$$\begin{aligned}
\hat{\sigma}^2 &= T^{-1}\sum_{t=1}^{T}\left(\Delta\hat{e}_{t\tau}-\hat{\rho}_1\hat{e}_{t-1\tau}\mathbb{1}\{\hat{e}_{t-1\tau}\geq\lambda\}-\hat{\rho}_2\hat{e}_{t-1\tau}\mathbb{1}\{\hat{e}_{t-1\tau}<\lambda\}-\sum_{j=1}^{K}\hat{\gamma}_j\Delta\hat{e}_{t-j\tau}\right)^2 \\
&= T^{-1}\sum_{t=1}^{T}\varepsilon_{t\tau}^2+o_p(1)=T^{-1}\sum_{t=1}^{T}D(L)^2\eta_\tau'\Delta X_{t\tau}(\Delta X_{t\tau})'\eta_\tau \qquad (4.37)\\
&\Rightarrow D(1)^2\eta_\tau'\Omega_\tau\eta_\tau=D(1)^2 l_{11}^2\kappa_\tau' D_\tau\kappa_\tau,
\end{aligned}$$

where the long-run covariance matrix is given by

$$\Omega_\tau=\begin{bmatrix}\omega_{11} & \omega_{21}' & 0 & (1-\tau)\omega_{21}' & 0\\ \omega_{21} & \Omega_{22} & 0 & (1-\tau)\Omega_{22} & 0\\ 0 & 0 & 0 & 0 & 0\\ (1-\tau)\omega_{21} & (1-\tau)\Omega_{22} & 0 & (1-\tau)\Omega_{22} & 0\\ 0 & 0 & 0 & 0 & 0\end{bmatrix} \qquad (4.38)$$

and

$$D_\tau=\begin{bmatrix}1 & 0 & 0 & 0 & 0\\ 0 & I_m & 0 & (1-\tau)I_m & 0\\ 0 & 0 & 0 & 0 & 0\\ 0 & (1-\tau)I_m & 0 & (1-\tau)I_m & 0\\ 0 & 0 & 0 & 0 & 0\end{bmatrix}. \qquad (4.39)$$

Similar results can be obtained for t_2 so that the results (4.35), (4.36), (4.37) combine with the CMT to proof the theorem under the null hypothesis.

Under the alternative, the system is cointegrated so that we have $\hat{\eta}_\tau \xrightarrow{p} \eta_\tau$ and

$$\hat{\eta}_\tau=\eta_\tau+O_p(T^{-1}) \qquad (4.40)$$

from Phillips and Durlauf (1986), Theorem 4.1. Thus, for the residual series it holds that

$$\hat{e}_{t\tau}=\hat{\eta}_\tau' z_t=\eta_\tau' z_t+O_p(T^{-1/2})=q_{t\eta\tau}+O_p(T^{-1/2}). \qquad (4.41)$$

By assumption a stationary SETAR representation of $q_{t\eta\tau}$ exists and is given by

$$q_{t\eta\tau}=a_{11}q_{t-1\eta\tau}\mathbb{1}\{q_{t-1\eta\tau}\geq\lambda\}+a_{12}q_{t-1\eta\tau}\mathbb{1}\{q_{t-1\eta\tau}<\lambda\}+\sum_{j=2}^{\infty}a_j q_{t-j\eta\tau}+\varepsilon_{t\eta\tau}^*, \qquad (4.42)$$

where $\varepsilon_{t\eta\tau}^*$ is an orthogonal $(0,\sigma_{\varepsilon_{\eta\tau}^*})$ sequence. This can alternatively be written as

$$\Delta q_{t\eta\tau}=\psi_{11}q_{t-1\eta\tau}\mathbb{1}\{q_{t-1\eta\tau}\geq\lambda\}+\psi_{12}q_{t-1\eta\tau}\mathbb{1}\{q_{t-1\eta\tau}<\lambda\}+\sum_{j=2}^{\infty}\psi_j\Delta q_{t-j\eta\tau}+\varepsilon_{t\eta\tau}^*. \qquad (4.43)$$

If we consider the t ratio of $\hat{\rho}_1$ and use the expression

$$t_1=\frac{1}{\hat{\sigma}}\left(\hat{\rho}_1\left(U_{1\tau}'Q_K U_{1\tau}\right)^{1/2}\right), \qquad (4.44)$$

we find that $\hat{\rho}_1 \xrightarrow{p} \psi_{11} \neq 0$ and $\hat{\sigma}^2 \xrightarrow{p} \sigma^2_{\varepsilon^*_{\eta\tau}}$. Further, we observe

$$U'_{1\tau} Q_K U_{1\tau} = U'_{1\tau} U_{1\tau} - U'_{1\tau} M_K (M'_K M_K)^{-1} M'_K U_{1\tau} = O_p(T) \tag{4.45}$$

which yields $t_1 = O_p(T^{1/2})$ and similarly $t_2 = O_p(T^{1/2})$. Hence, we immediately see that $F^* \to \infty$ as $T \to \infty$. $\square$

B .2 Tables

Table 4.4: Rejection frequencies of the $\sup F$ test under asymmetric adjustment

ρ_1	ρ_2	C			C/T			C/S			Φ		
		10%	5%	1%	10%	5%	1%	10%	5%	1%	10%	5%	1%
−0.025	−0.05	0.113	0.066	0.014	0.114	0.061	0.012	0.037	0.018	0.002	0.123	0.070	0.019
	−0.10	0.121	0.070	0.014	0.104	0.059	0.018	0.037	0.016	0.002	0.141	0.078	0.020
	−0.15	0.139	0.071	0.020	0.137	0.075	0.023	0.049	0.022	0.005	0.160	0.086	0.023
	−0.20	0.146	0.076	0.016	0.142	0.080	0.018	0.046	0.022	0.003	0.174	0.102	0.026
	−0.25	0.157	0.091	0.024	0.149	0.080	0.023	0.050	0.025	0.005	0.185	0.113	0.031
−0.05	−0.10	0.132	0.078	0.019	0.152	0.080	0.017	0.044	0.019	0.003	0.174	0.093	0.024
	−0.15	0.152	0.094	0.022	0.137	0.079	0.021	0.050	0.024	0.004	0.211	0.121	0.033
	−0.20	0.189	0.111	0.028	0.174	0.097	0.028	0.060	0.024	0.005	0.248	0.151	0.042
	−0.25	0.196	0.115	0.031	0.176	0.105	0.028	0.073	0.038	0.004	0.276	0.170	0.052
−0.10	−0.15	0.211	0.122	0.030	0.205	0.119	0.030	0.082	0.034	0.009	0.338	0.205	0.060
	−0.20	0.263	0.166	0.046	0.229	0.138	0.042	0.110	0.049	0.007	0.411	0.262	0.085
	−0.25	0.291	0.174	0.058	0.264	0.163	0.048	0.118	0.063	0.012	0.477	0.313	0.114

Note: C, C/T and C/S denote the structural break models in (4.3). Φ denotes the Enders-Siklos cointegration test with SETAR adjustment. The table is based on 2,500 replications of the DGP described in (4.14) for sample size $T = 100$.

Table 4.5: Rejection frequencies of the sup F test under structural change and symmetric adjustment

	$\mu_1 = 1, \mu_2 = 4, \alpha_1 = 2, \alpha_2 = 2$						$\mu_1 = 1, \mu_2 = 1, \alpha_1 = 2, \alpha_2 = 4$					
	T = 50			T = 100			T = 50			T = 100		
τ	0.25	0.50	0.75	0.25	0.50	0.75	0.25	0.50	0.75	0.25	0.50	0.75
C	0.535	0.547	0.545	0.993	0.994	0.993	0.787	0.613	0.696	0.936	0.863	0.955
C/T	0.415	0.405	0.417	0.972	0.968	0.974	0.652	0.476	0.485	0.897	0.839	0.924
C/S	0.296	0.310	0.319	0.939	0.933	0.934	0.693	0.489	0.341	0.995	0.982	0.975
ADF (c)	0.139	0.096	0.096	0.391	0.274	0.277	0.089	0.060	0.086	0.126	0.100	0.145
ADF (c + t)	0.124	0.125	0.116	0.397	0.481	0.434	0.076	0.058	0.096	0.109	0.122	0.187
GH (C)	0.486	0.486	0.508	0.977	0.976	0.979	0.259	0.265	0.492	0.643	0.587	0.898
GH (C/T)	0.387	0.377	0.390	0.929	0.930	0.920	0.203	0.220	0.277	0.508	0.546	0.761
GH (C/S)	0.376	0.377	0.387	0.940	0.935	0.939	0.399	0.378	0.408	0.969	0.968	0.968
Φ	0.297	0.261	0.263	0.883	0.733	0.863	0.194	0.144	0.194	0.331	0.279	0.354

Note: C, C/T and C/S denote the structural break models in (4.3). ADF (c) and ADF (c + t) refer to the Engle-Granger test with intercept and intercept plus trend, respectively. GH denotes the Gregory-Hansen test. Φ denotes the Enders-Siklos cointegration test with SETAR adjustment. The table is based on 2,500 replications of the DGP described in (4.14). The autoregressive coefficients are $\rho_1 = \rho_2 = -0.5$, i.e. the adjustment is constant and symmetric.

Table 4.6: Estimates of the breakpoint under symmetric adjustment

	$\mu_1 = 1, \mu_2 = 4, \alpha_1 = 2, \alpha_2 = 2$						$\mu_1 = 1, \mu_2 = 1, \alpha_1 = 2, \alpha_2 = 4$					
	T = 50			T = 100			T = 50			T = 100		
τ	0.25	0.50	0.75	0.25	0.50	0.75	0.25	0.50	0.75	0.25	0.50	0.75
C	0.32(0.15)	0.53(0.11)	0.70(0.15)	0.28(0.10)	0.51(0.08)	0.74(0.11)	0.34(0.18)	0.55(0.13)	0.72(0.13)	0.28(0.12)	0.54(0.11)	0.75(0.10)
	0.28(0.04)	0.52(0.04)	0.74(0.04)	0.26(0.02)	0.51(0.02)	0.76(0.02)	0.28(0.05)	0.54(0.04)	0.76(0.04)	0.26(0.02)	0.52(0.02)	0.77(0.02)
C/T	0.38(0.19)	0.50(0.16)	0.66(0.22)	0.33(0.16)	0.51(0.11)	0.69(0.16)	0.39(0.19)	0.53(0.15)	0.65(0.20)	0.31(0.15)	0.53(0.11)	0.73(0.13)
	0.28(0.26)	0.50(0.08)	0.74(0.34)	0.27(0.03)	0.51(0.02)	0.75(0.03)	0.28(0.26)	0.52(0.10)	0.74(0.22)	0.27(0.02)	0.52(0.02)	0.75(0.02)
C/S	0.35(0.16)	0.53(0.12)	0.68(0.16)	0.30(0.11)	0.51(0.07)	0.72(0.12)	0.33(0.14)	0.54(0.09)	0.71(0.13)	0.27(0.07)	0.51(0.05)	0.75(0.07)
	0.28(0.18)	0.54(0.04)	0.76(0.12)	0.25(0.02)	0.51(0.02)	0.76(0.03)	0.26(0.14)	0.54(0.04)	0.78(0.08)	0.25(0.02)	0.51(0.02)	0.77(0.01)

Note: C, C/T and C/S denote the structural break models in (4.3). The left panel and right panel report the estimates of the break fraction following a shift in the intercept and a shift in the slope, respectively. Upper rows contain the mean breakpoint estimate and the empirical standard deviation. Lower row contain the median breakpoint and the interquartile range. The autoregressive coefficients are $\rho_1 = \rho_2 = -0.5$, i.e. the adjustment is constant and symmetric.

Table 4.7: Rejection frequencies of the sup$F$ test under structural change and asymmetric adjustment

	$\mu_1 = 1, \mu_2 = 4, \alpha_1 = 2, \alpha_2 = 2$												
		C			C/T			C/S			Φ		
ρ_1	ρ_2	10%	5%	1%	10%	5%	1%	10%	5%	1%	10%	5%	1%
−0.025	−0.05	0.096	0.050	0.013	0.098	0.050	0.009	0.037	0.018	0.004	**0.119**	**0.062**	**0.014**
	−0.15	0.111	0.059	0.014	0.109	0.058	0.012	0.043	0.018	0.003	**0.143**	**0.067**	**0.018**
	−0.25	0.128	0.070	0.016	0.120	0.062	0.018	0.048	0.022	0.004	**0.153**	**0.084**	**0.022**
−0.05	−0.10	0.120	0.061	0.013	0.112	0.064	0.013	0.043	0.020	0.003	**0.147**	**0.080**	**0.019**
	−0.25	0.160	0.089	0.021	0.149	0.082	0.024	0.056	0.026	0.003	**0.182**	**0.108**	**0.032**
−0.10	−0.15	0.173	0.106	0.023	0.157	0.086	0.020	0.058	0.028	0.004	**0.210**	**0.118**	**0.037**
	−0.25	0.226	0.140	0.041	0.196	0.114	0.030	0.081	0.040	0.008	**0.258**	**0.158**	**0.053**
	$\mu_1 = 1, \mu_2 = 1, \alpha_1 = 2, \alpha_2 = 4$												
		C			C/T			C/S					
ρ_1	ρ_2	10%	5%	1%	10%	5%	1%	10%	5%	1%	10%	5%	1%
−0.025	−0.05	**0.241**	0.165	**0.097**	0.196	0.137	0.065	0.214	**0.168**	0.093	0.114	0.057	0.012
	−0.15	**0.263**	0.190	**0.109**	0.215	0.157	0.076	0.255	**0.201**	0.107	0.117	0.059	0.014
	−0.25	**0.274**	0.200	0.114	0.226	0.164	0.081	0.271	**0.212**	**0.118**	0.119	0.064	0.016
−0.05	−0.10	0.268	0.198	0.114	0.219	0.164	0.073	**0.269**	**0.207**	**0.113**	0.114	0.059	0.013
	−0.25	0.298	0.221	0.132	0.243	0.182	0.088	**0.325**	**0.256**	**0.140**	0.126	0.067	0.017
−0.10	−0.15	0.319	0.246	0.138	0.255	0.179	0.092	**0.340**	**0.271**	**0.155**	0.136	0.072	0.018
	−0.25	0.353	0.277	0.160	0.286	0.206	0.109	**0.394**	**0.311**	**0.189**	0.193	0.110	0.037
	$\mu_1 = 1, \mu_2 = 4, \alpha_1 = 2, \alpha_2 = 4$												
		C			C/T			C/S			Φ		
ρ_1	ρ_2	10%	5%	1%	10%	5%	1%	10%	5%	1%	10%	5%	1%
−0.025	−0.05	0.080	0.060	0.037	0.228	0.163	0.080	**0.235**	**0.185**	**0.111**	0.111	0.057	0.011
	−0.15	0.171	0.136	0.080	0.246	0.180	0.093	**0.268**	**0.215**	**0.123**	0.112	0.058	0.013
	−0.25	0.274	0.195	0.110	0.263	0.185	0.100	**0.278**	**0.231**	**0.136**	0.118	0.062	0.015
−0.05	−0.10	**0.269**	**0.198**	**0.116**	0.224	0.155	0.074	0.265	0.198	0.111	0.116	0.058	0.011
	−0.25	0.308	0.225	0.129	0.248	0.175	0.086	**0.314**	**0.250**	**0.138**	0.130	0.068	0.017
−0.10	−0.15	0.327	0.243	0.135	0.250	0.184	0.093	**0.343**	**0.270**	**0.153**	0.134	0.069	0.017
	−0.25	0.355	0.277	0.160	0.284	0.204	0.109	**0.388**	**0.315**	**0.184**	0.143	0.081	0.022

Note: C, C/T and C/S denote the structural break models in (4.3). Φ denotes the Enders-Siklos cointegration test with SETAR adjustment. The table is based on 2,500 replications of the DGP described in (4.14). The breakpoint occurs mid-sample, i.e. $\tau = 0.5$. The test with the highest rejection rates is highlighted in boldface.

Chapter 5

A Markov regime-switching model of crude oil market integration

1 Introduction

The discussion on whether world crude oil markets are globalized or regionalized has received a great deal of attention in recent years. Adelman (1984) described the world crude oil market as 'one great pool'. Changes in market conditions in one region are then expected to affect other geographical regions immediately. An existing price differential in local oil markets that exceeds the transportation costs of third party exporters gives rise to arbitrage opportunities. The subsequent supply pressure is expected to close the difference in prices. The idea of 'one great pool' was challenged by Weiner (1991) who finds empirical support for a high degree of regionalization. His findings imply that the world crude oil market is fragmented and the effects of price shocks to regional crude oil prices are restricted to this specific regional market.

This initial discussion has triggered numerous empirical studies, among them Guelen (1999), Fattouh (2010), Reboredo (2011) and Ji and Fan (2015), that tackle the 'globalization-regionalization' hypothesis from different angles. The majority of recent studies finds evidence for a globalized crude oil market. However, the structure of the market does not seem to be stable over time.

Our paper contributes to the literature by proposing a regime-switching model for the long-run relationships among benchmark crude oil prices. This allows us to relax the assumption of constant dynamics over the sample period which has to hold for linear cointegration models. More specifically, we apply a Markov-switching vector error correction model (MSVECM) to capture changing roles of crudes in the world crude oil market and a changing degree of market integration. This enables us to identify regime-shifts from the data without the need to pre-specify structural breaks. We aim to account for increasingly volatile crude oil prices and changing economic and geopolitical conditions over a sample reaching from 1987 to 2015. Our data-set consists of five major crude oil benchmark prices – WTI, Brent, Bonny Light, Dubai and Tapis – representative of five crude oil producing regions.

The question whether the crude oil market is globalized or regionalized has important policy implications. Developed countries hold strategic petroleum reserves to provide emergency crude oil in times of disruptive supply shocks. Members of the International Energy Agency are required to stockpile crude oil equal to 90 days of prior year's net oil imports[1]. If effects of supply shocks were restricted to one region, higher reserves would have to be stockpiled than in a globalized market where arbitrage opportunities lead to supply of cheaper oil from other production sites.

Furthermore, a precise assessment of the market behaviour is needed to anticipate the scope of new energy policies. Energy markets are currently experiencing fundamental changes since production of giant oil fields declines (Höök *et al.* (2009)) whereas new technologies, like hydraulic fracturing, are used to revitalize existing oil fields. Also, the interest in renewable energy has recently increased as might be reflected by the renewable energy directive of the European Union (European Commission (2016)). The decision to invest in the energy sector requires an accurate prediction of future crude oil prices. Focussing on the classical benchmarks (WTI and Brent) or only on local benchmark prices might prevent assessing the correct market behaviour if they do not reflect global supply and demand.

Moreover, a precise assessment of crude oil prices is needed for hedging purposes and the pricing of other derivatives related to crude oil prices. It is therefore of interest which benchmark price reflects crude oil market developments first and leads the pricing process. This may become even more important since activity in commodity exchange contracts has risen in recent years which is discussed under the term 'financialization' of commodity markets in the literature (see, for example, Buyuksahin and Harris (2011) and Tang and Xiong (2012)). Although activity in crude oil exchange trading has increased accordingly, trading physical oil is still carried out in large quantities and is non-transparent to the public. In practice, price reporting agencies, like Platts, provide assessments of benchmark crude oil prices. The prices in the physical oil market are collected by a window or market-on-close process in which bids, offers and the trade volume are assessed and prices are published as an end-of-day value. This leads to price-discovery which rests on voluntary and selective disclosure by market participants as well as subjective judgement of the price reporting agency. Although WTI, Brent and Dubai are considered to be the most important crude oil benchmarks, there is no universally recognized *global* crude oil spot price. Market agents exposed to crude oil price risks, therefore, are particularly interested in how different crude oil benchmarks interact and which of them responds fastest to changing conditions on the crude oil market.

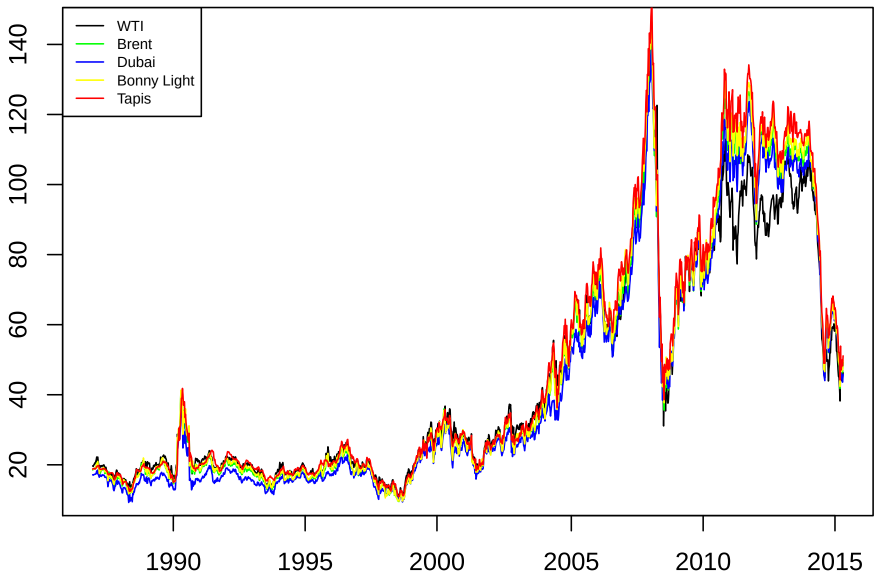
The remainder of the paper is organized as follows. Section 2 describes the structure of the world crude oil market and the role of benchmark prices. In Section 3 , we review the literature on crude oil market integration, Section 4 outlines the econometric framework used in the empirical part of the paper, Section 5 reports the results of the empirical application, Section 6 relates our findings to previous studies and Section 7 concludes.

2 Market structure and the role of benchmark prices

Internationally traded crude oil comes in different qualities and characteristics. Lighter crude oils yield a higher percentage of gasoline and diesel fuel than heavier crudes (usually measured in American Petroleum Institute (API) gravity). Since sulphur is an undesirable component, 'sour' crudes with a higher sulphur level are less sought after than 'sweet' crudes. Generally, light and sweet crudes are priced at a premium relative to heavy and sour crudes. Buyers and sellers of crude oil rely on the use

[1] The International Energy Agency (IEA) was founded in the wake of the first oil crisis. Historically, the majority of member states were net oil importers. Net exporters are exempt from this requirement. Although the role of the US as a net importer has to be reconsidered, following the resurgence of shale oil fields, the largest crude oil stockpiles are concentrated in the US.

Figure 5.1: Time series plots for regional crude oil price series (WTI, Brent, Dubai, Bonny Light, Tapis)



Source: DATASTREAM (see Section 5 .1).

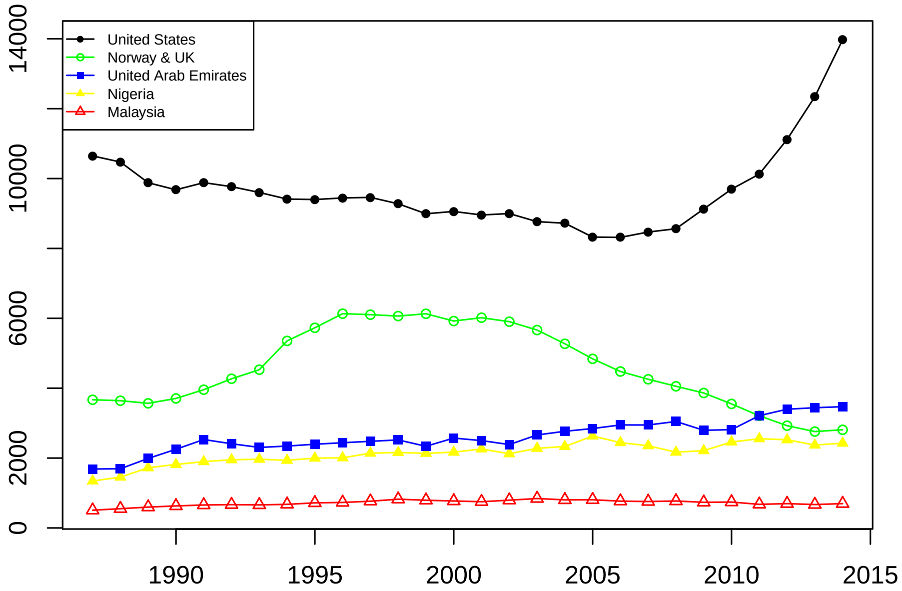
of benchmark crude oils (price markers) to price the different types of crude oil. These benchmarks typically exhibit the following properties: First, the volume of production must be sufficiently large to ensure physical liquidity. Second, the oilfield has to be located in a geopolitically and financially stable region to encourage market interactions. Third, delivery points have to be provided at locations suitable for trade with other market hubs to enable arbitrage. Finally, a diverse ownership of production should be present to prevent market interference and price manipulation. In practice, however, major crude oil benchmarks do not fulfil all the requirements equally. Non-benchmark crudes are priced relative to the benchmark crude at a premium or discount depending on their quality. This is known as formula pricing.

Brent is the reference for about 65% of crude oil traded around the globe according to the Intercontinental Exchange, whereas WTI is the dominant benchmark in the US (Intercontinental Exchange (2016)). Dubai is the main reference for Persian Gulf oil delivered to the Asian market. Bonny Light is a benchmark for West African oil fields and Tapis serves as a benchmark crude for the Asian Pacific region. Figure 5.1 shows the trajectories of the five benchmark prices from 1987 to 2015. The amount of oil production over time is depicted in Figure 5.2.

Originally, crude oil extracted from the Brent oilfield, which was discovered in 1971, formed the Brent benchmark (API gravity of 38.3° and 0.37% sulphur). Production from the Brent oilfield started to decline in the mid-1980s which led to volatile prices. Commingling Brent with oil produced in the Ninian oil field, also located in the North Sea, alleviated this problem temporarily. A further decline in production led to the inclusion of oil from the Forties, Oseberg and Ekofisk fields (Fattouh (2006)). Today, the production is still declining (see Figure 5.2) and a substantial share of Europe's crude oil supply comes from Russia, which raises the question whether Brent has retained its role as a benchmark price.

The North American crude oil West Texas Intermediate (WTI), which has an API of 39.6° and contains 0.24% sulphur, making it a light and sweet crude, is transported from the extraction sites via pipelines to Cushing, Oklahoma. In 1983, NYMEX chose Cushing as the official delivery point for

Figure 5.2: Temporal evolution of crude oil production in five production sites denoted in thousand barrels per day



Source: US Energy Information Agency (2017).

its light sweet crude futures contract which in turn connects the oil fields to refineries and ports. Following the explosive growth in production from shale oil fields, the Cushing pipeline nexus has turned out to be a bottleneck. Oil is transported to Cushing in large quantities but the ill-equipped infrastructure delayed the distribution of oil. Consequently, the build-up in inventory caused WTI to trade at a discount compared to other benchmark crude oils and to decouple from the world crude oil market. This phenomenon is known in the literature as the 'broken benchmark' (Fattouh (2007), Fattouh (2010) and Ji and Fan (2015)). If WTI was considered the global price setter, a decoupling effect would severely impair effective hedging against risks related to energy prices and would lead to incorrect pricing of other derivatives based on crude oil.

WTI and Brent held a constant price differential until around 2010. Historically, WTI traded at a premium compared to Brent, attributed to the fact that WTI is the lighter and sweeter crude oil. Beginning in 2010, the spread has been reversed. The hydraulic fracturing boom in the US helped to increase the US crude oil production by 75% from 2008 to 2014 according to the US Energy Information Agency (US Energy Information Agency (2016)) and subsequently ensured full inventories. Hydraulic fracturing is not utilized with the same intensity in the oil fields of the North Sea. A significant widening of the price differential can be observed after the shale oil boom in the US picked up speed. Moreover, the US ban on crude oil exports during our observational period may have prohibited the reduction of overcapacities through international trade[2].

Dubai is of slightly lower grade than WTI or Brent. An API gravity of 31° and 2% sulphur makes Dubai a medium heavy and sour crude. It comprises of crudes from different oil fields in Dubai, Oman and Abu Dhabi. Despite the existence of other regional crudes with a larger physical base, Dubai serves as a benchmark price for oil extracted in the Gulf region.

Bonny Light is a sweet but medium heavy crude oil (API 33.4°, 0.16% sulphur). The Bonny Light

[2]The US have lifted the crude oil export ban in January 2016.

production is concentrated in the onshore and offshore areas of the Niger Delta of Nigeria. West African crude oil is mostly refined outside the region, in Asia, Europe and the US. Violent conflicts in the Niger region led to temporary disruption of the oil production in September 2004.

Tapis is produced offshore in the South China Sea (the Seligi, Guntong, Tapis, Semangkok, Irong Barat, Tebu, and Palas fields). It is of the highest quality with an API gravity of 45.2° and low sulphur content (0.03%).

Historically, none of the five benchmark prices in our study has emerged as a universally recognized global price setter. A price setter is defined as a price that influences other prices in the same category directly or indirectly without being influenced itself. In terms of our empirical application which focuses on a cointegrated system, a price setter can be identified as a variable which does not adjust to deviations from the long-run equilibrium which is instead maintained by the remaining variables. The price setter takes the role of a *lead* variable whereas the remaining variables act as *lag* variables.

We believe that focussing on benchmark prices reduces the problem encountered by studies involving both benchmark and non-benchmark prices (Wlazlowski *et al.* (2011) and Candelon *et al.* (2013)): Non-benchmark prices are priced in relation to the regional benchmark with price adjustments made depending on quality and transportation costs (formula pricing). While we expect the benchmark/non-benchmark relation to be strong, we are primarily interested in the relationship between geographically separated markets. Only if we find long-run co-movement and short-run adjustments among prices without a formula pricing relationship, we can argue in favour of a globalized crude oil market.

3 Literature

After Adelman (1984) and Weiner (1991) initiated the discussion on the integration of international crude oil markets, a substantial body of literature on the subject has emerged. Empirical studies mostly employ cointegration models to assess the relations among crude oil prices. For instance, Rodriguez and Williams (1993) aim to test the ‘one great pool’ hypothesis using a cointegration analysis for monthly data from 1982 to 1992. They claim to find evidence for integrated crude oil markets by rejecting the hypothesis of no cointegration which implies the presence of a long-run stable relationship among regional crude oil prices. However, Weiner (1993) emphasizes that, although prices follow a common trend, the short-run dynamics are important to characterize the relationship among regional prices. More precisely, Weiner (1993) argues that only price reactions to changes in other crude oil prices in the short-run should lead to a rejection of the ‘regionalization’ hypothesis. He criticizes the use of linear cointegration models which are not able to capture the true dynamics of a changing world crude oil market.

Guelen (1999) tries to account for structural change by applying cointegration models to subsamples of falling and rising crude oil prices. He finds evidence for stronger co-movement in periods of increasing prices, implying that linear cointegration models indeed are not well-suited for the analysis of price dynamics in global crude oil markets. Further, he finds that WTI and Brent take the role of global benchmark prices. Bentzen (2007) specifies a vector error correction model for daily crude oil prices from the Middle East, North America and the North Sea. Using data from January 1988 to December 2004, evidence is found for a globalized market with an increasing role of OPEC prices, thereby reducing the strength of WTI and Brent as global benchmarks.

Hammoudeh *et al.* (2008) and Fattouh (2010) use threshold cointegration models to capture a potentially non-linear relationship among crude oil prices. More specifically, Hammoudeh *et al.* (2008) examine the relationship among four benchmark prices (WTI, Brent, Dubai, Maya) based on daily data from 1990 to 2006. They use momentum threshold autoregressive (MTAR) models which allow for different adjustment depending on whether the spread between crudes is widening or narrowing. While all price pairs are cointegrated, Brent and WTI are found to be leading the pricing process in the long-run. Instead, Fattouh (2010) analyzes crude oil price differentials at a weekly frequency from 1997 to 2008 using threshold autoregressive (TAR) models. Prices of crude oils with a similar quality show a strong comovement over the sample whereas divergence of prices for crudes of different qualities can be observed.

Liu *et al.* (2013) investigate the role of China in the world crude oil market. Since China is one of the major oil importers with increasing demand in recent years, China's energy policy has an important influence on regional crude oil prices. If price changes of the regional benchmark, Daqing, were transmitted to world crude oil prices, indications of market integration would be found. However, the results of a threshold VECM reveal only a one-directional effect from world crude oil markets to the regional Daqing benchmark. Wilmot (2013) focusses on the Canadian-US market integration. He argues that the 'globalization' hypothesis also requires that a long-run relationship among secondary 'non-benchmark' crudes exists. Evidence from a cointegration analysis of Edmonton Par, a light crude, and Western Canadian Select, a heavy crude, and its US (Mexican) analogues, confirm a long-run relationship. However, the analysis reveals a structural break in the cointegrating vector and the breakpoint is determined to coincide with the Financial Crisis.

More recently, Ji and Fan (2015) investigate the long-run equilibrium relationships among the five major regional crude oil benchmarks (WTI, Brent, Dubai, Bonny Light, Tapis) by using a VECM combined with a directed acyclic graph technique. Based on tests for the presence of structural breaks, they split their sample at the break point in October 2010. They find that WTI was a price setter before 2010 while Brent is in a leading role since 2011. Tapis has always been a price taker whereas Dubai and Bonny Light have taken both roles at times. Mann and Sephton (2016) use band-TAR threshold cointegration models to examine the long-run relationships between WTI and Brent and WTI and Oman. They find these crude oil price pairs to be tied together in the long-run. Since each price adjusts to the long-run equilibrium at some point, they conclude that a unique global benchmark prices does not exist.

Additionally, there are further studies that focus on the changing conditions on the crude oil market. Reboredo (2011) models the dependence structure between crude oil benchmark prices using a copula approach. Upper and lower tail dependence is found, suggesting that benchmark crude oils boom and crash simultaneously. This is considered evidence for a globalized world crude oil market. Candelon *et al.* (2013) examine causal linkages at regional oil markets when prices are on average extremely high or low. The study reveals benchmark prices besides WTI and Brent. Moreover, market integration is found to be weaker during extreme times. Instead of Candelon *et al.* (2013)'s set of 32 different crudes, Lu *et al.* (2014) restrict their analysis to four benchmark prices (WTI, Brent, Dubai, Tapis) and find a stronger market integration after disruptive events take place. Zhang and Zhang (2015) employ a Markov-switching autoregressive model to investigate the short-run dynamics between Brent and WTI. They find three price regimes which are characterized by different dynamics.

In all, evidence is mounting that crude oil markets are 'globalized'. Crude oil prices seem to hold long-run equilibrium relationships. However, the degree of market integration does not seem to be stable over time.

4 Econometric methodology

The long-run and short-run dynamics of the crude oil prices, collected in a vector y_t, are modelled using a vector error correction model (VECM). The model assumes that the prices are linked by stable long-run relationships. However, the variables deviate from these equilibrium relationships in the short-run due to random shocks. Maintaining the long-run relationships requires that deviations are corrected by the variables in the short-run. Put differently, the variables are said to adjust to equilibrium errors. Following Johansen (1988)'s notation, the linear VECM is given as

$$\Delta y_t = \mu + \Pi y_{t-1} + \sum_{i=1}^{p-1} \Gamma^i \Delta y_{t-i} + u_t, \tag{5.1}$$

where y_t is a $N \times 1$ vector of $I(1)$ variables, μ is a vector of drift parameters and u_t is a vector of white noise error terms. The $k \times k$ parameter matrix $\Pi = \alpha\beta'$ captures both the long-run equilibrium relations and the adjustment behaviour. The matrix β contains r cointegrating vectors and α carries the loadings in each of the r vectors.

A particular feature of the linear VECM is that it assumes constancy of all parameters in its data generating process. Certainly, this assumption appears to be restrictive in the context of a volatile crude oil market. Previous studies described relevant disruptive events concerning the energy market (see Lu *et al.* (2014) for a list of events from 2002 to 2011), and specific issues on the crude oil market, for example WTI, as a 'broken benchmark'. These events are likely to induce structural changes in the relations among crude oil prices. Although we expect the crude oil prices to maintain constant long-run equilibria since crude oils are close substitutes[3], the roles of crude oils in the market, for example, switching from price takers to prices setters and vice versa, might change over time. Particularly, a decoupling of WTI from the world crude oil market might have led to exogeneity of WTI for this period. We therefore study the evolution of the adjustment coefficients while the long-run equilibrium relationships are assumed to stay constant over time.

To account for potential time-varying adjustment, we apply a Markov-switching VECM (MSVECM) to the data. Markov-switching models in a time series econometrics framework were introduced by Hamilton (1989) and the MSVECM used in this paper was proposed by Krolzig (1997). We consider a q-regime VECM which allows the parameters to be state-dependent. The MS(q)-VECM takes the form of

$$\Delta y_t = \mu_{s_t} + \Pi_{s_t} y_{t-1} + \sum_{i=1}^{p-1} \Gamma^i_{s_t} \Delta y_{t-i} + u_t, \qquad u_t | s_t \sim N(0, \Sigma_{s_t}), \tag{5.2}$$

where μ_{s_t} are state-dependent drift terms, Π_{s_t} is the state-dependent long-run impact matrix, $\Gamma^i_{s_t}$ are state-dependent short-run dynamics and the error terms have a normal distribution conditional on the state s_t. A Cholesky decomposition of the error term variance-covariance matrix gives $\Sigma = LS^2L'$ where L is a normalized lower triangular matrix and S is diagonal. The error term variance can either be restricted to stay fixed over all states, $\Sigma_{s_t} = \Sigma$ for all $s_t = 1, 2, \ldots, q$, or change over states. We distinguish between

[3]Differences in quality (density and sulphur content) are reflected in discount or premium prices.

a switching scale, $\Sigma_{s_t} = LS^2_{s_t}L'$, and a fully switching variance, where each element of Σ_{s_t} is switching according to s_t, $\Sigma_{s_t} = L_{s_t}S^2_{s_t}L'_{s_t}$. A fully switching variance-covariance matrix comes at the cost of an increasing number of parameters that have to be estimated.

The state of the data-generating process is governed by a latent integer state variable s_t. The probability that s_t attains some particular value $j \in \{1,2,\ldots,q\}$ depends only on the most recent value s_{t-1}:

$$P(s_t = j|s_{t-1} = i, s_{t-2} = k, \ldots) = P(s_t = j|s_{t-1} = i) = p_{ij} \quad \forall i, j = 1,2,\ldots,q. \tag{5.3}$$

Such a process is described as a q-state Markov chain with constant transition probabilities $p_{ij} > 0$, $\sum_{j=1}^{q} p_{ij} = 1$ (Hamilton (1994)). We assume the Markov chain to be irreducible and ergodic, which means that each regime can be reached from any previous regime (absence of absorbing states) and no regime has a periodic occurrence.

The state-dependent long-run impact matrix Π_{s_t} is decomposed in the constant cointegrating vectors and the state-dependent weighting matrix α_{s_t},

$$\Pi_{s_t} = \alpha_{s_t}\beta', \tag{5.4}$$

where α_{s_t} contains the state-dependent adjustment coefficients which measure the reaction to deviations from the long-run equilibria for each regime. In our application, we are particularly concerned with the evolution of the adjustment coefficients over time and regimes. The adjustment coefficients can be interpreted in the context of a lead-lag relationship among the crude oil prices. If one of our crudes was a global price setter, it would not adjust to deviations from the long-run equilibrium induced by random shocks. The price setting crude thus takes the role of a *lead* variable. Analyzing the long-run relationships among crude oil prices via a MSVECM provides further insights in the structure of the world crude oil market since it enables us to identify exogenous benchmark prices under particular regimes of the process.

The dynamic properties are further investigated by observing the behaviour of the system after shocks to variables of the system using regime-specific orthogonalized impulse response functions. For this matter, we need to transform the VECM representation given in (5.2) to a vector moving average (VMA) representation,

$$y_t = u_t + \Psi^1_{s_t}u_{t-1} + \Psi^2_{s_t}u_{t-2} + \Psi^3_{s_t}u_{t-3} + \ldots \tag{5.5}$$

Since the error terms u_t are correlated with each other, we use the Cholesky decomposition of the regime-specific error term variance-covariance matrix again and construct orthogonalized impulse response functions,

$$IRF^1_{s_t}(\hat{\theta}) = \widehat{L}_{s_t}, \quad IRF^2_{s_t}(\hat{\theta}) = \widehat{\Psi}^1_{s_t}\widehat{L}_{s_t}, \quad \ldots, \quad IRF^h_{s_t}(\hat{\theta}) = \widehat{\Psi}^{h-1}_{s_t}\widehat{L}_{s_t}, \tag{5.6}$$

where $\hat{\theta}$ denotes the entirety of all estimated parameters.

Naturally, the number of parameters to estimate increases with the number of states which are specified in the MSVECM, so that a parsimonious model specification leads to a maximum of two or three states. However, the exact number of states is usually not known a priori and has to be jointly selected with additional variables, that is, further lags to capture short-run dynamics. Psaradakis and Spagnolo (2006) found that information criteria can accurately identify the appropriate number of states for a Markov-switching model. Awirothananon and Cheung (2009) argued for the use of the BIC to select the number of states based on results of Monte Carlo experiments. In the following application, we follow Awirothananon and Cheung (2009) and use the BIC for model selection with respect to the number of states, the lag length and switching behaviour of the drift terms as well as elements of the variance-covariance matrix.

5 Empirical analysis

5 .1 Data

For this study, we observe crude oil price data at weekly frequency from May 1987 until October 2015. All crude oil prices are free on board (FOB) spot prices[4], observed at each Monday and denominated in US dollars per barrel. The time series are obtained from DATASTREAM[5] and the original observations were transformed by taking natural logarithms.

First, the time series are tested for their order of integration. The results of ADF and KPSS unit root tests are reported in Table 5.1. Furthermore, we apply the Lee-Strazicich (LS) unit root test which accounts for two structural breaks in the null and alternative (Lee and Strazicich (2003)). The null hypothesis of the ADF and LS tests cannot be rejected at the 1% significance level for all prices while the null hypothesis of the KPSS test is rejected at all conventional significance levels. We obtain opposite results for the returns. The tests support the hypothesis that all prices follow a unit root process and are integrated of order one.

Table 5.1: Unit root tests of the logarithmized crude oil prices.

Variables	ADF	LS	KPSS	Variables	ADF	LS	KPSS
WTI	-2.635	-2.846^{*}	0.668^{***}	Δ WTI	-22.153^{***}	-37.433^{***}	0.064
Brent	-2.901	-3.087^{**}	0.738^{***}	Δ Brent	-20.234^{***}	-34.436^{***}	0.068
Dubai	-2.794	-3.459^{**}	0.742^{***}	Δ Dubai	-19.773^{***}	-41.847^{***}	0.071
Bonny Light	-2.520	-3.014^{*}	0.742^{***}	Δ Bonny Light	-20.167^{***}	-31.848^{***}	0.069
Tapis	-2.575	-3.485^{**}	0.725^{***}	Δ Tapis	-18.630^{***}	-42.166^{***}	0.072

Note: The ADF, LS and KPSS test equations are estimated including an intercept and trend for the variables in levels. The test equations for the first differences include an intercept. Lag selection is based on the Bayesian Information Criterion (BIC).
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

5 .2 Linear cointegration analysis

To test for cointegration, we rely on the Johansen rank test which is based on the VECM specified in Equation (5.1). The cointegrating rank r is determined by the number of estimated eigenvalues of the estimated adjustment coefficient matrix Π that are significantly greater than zero. Johansen (1988, 1991) proposed likelihood ratio type tests of which we use the trace test variant[6]. The trace test examines the null hypothesis, $\text{rank}(\Pi) = r_0$, against the alternative hypothesis, $r_0 < \text{rank}(\Pi) \leq k-1$.

The results of the cointegration test are presented in panel (a) of Table 5.2. Since the null hypothesis $r_0 = 3$ can soundly be rejected, we assume the maximum number of cointegrating vectors of four. The

[4]Pertains to a transaction whereby the seller makes the product available within an agreed on period at a given port at a given price; it is the responsibility of the buyer to arrange for the transportation and insurance. (US Energy Information Administration)

[5]The data can be found using Mnemonic (Code): OILTPMY (S214WT), OILDUBI (T15609), OILBRNP (S04107), CRUDWTC (S369VW), OILAFRB (S00112).

[6]The maximum eigenvalue test reaches the same conclusion: The null hypothesis of at most three cointegration vectors is rejected.

normalized cointegrating vectors are displayed in Panel (b) of Table 5.2. We find that the price differentials between WTI and the four remaining crudes are relevant long-run equilibria. The trade-off between a parsimonious specification and sufficiently capturing the short-run dynamics of the system leads to two additional lagged differences ($K = 2$).

Table 5.2: Cointegration tests and linear VECM

	$N-r$	r	Eig.value	Trace	5% Crit. val.	p-Value
Panel (a): I(1)-analysis						
	5	0	.1084	361.75	76.07	.000
	4	1	.0651	192.16	53.12	.000
	3	2	.0374	92.61	34.91	.000
	2	3	.0292	36.22	19.96	.000
	1	4	.0013	1.95	9.24	.783

	WTI	Brent	Bonny	Dubai	Tapis	μ
Panel (b): Cointegration vectors						
β_1	-1.087				1	.276
β_2	-1.136			1		.584
β_3	-1.097		1			.355
β_4	-1.094	1				.363
Panel (c): Adjustment coefficients						
α_1	$.066^{*}$	$.104^{***}$	$.110^{***}$	.049	$-.162^{***}$	
	(1.879)	(2.974)	(3.159)	(1.514)	(−6.198)	
α_2	.028	$.064^{**}$	$.063^{**}$	−.016	$.061^{***}$	
	(1.005)	(2.270)	(2.238)	(−.606)	(2.877)	
α_3	$-.214^{*}$	$-.229^{**}$	$-.432^{***}$	$-.272^{***}$	−.001	
	(−1.933)	(−2.062)	(−3.915)	(−2.638)	(−.018)	
α_4	$.198^{*}$	.053	$.257^{**}$	$.242^{**}$	.090	
	(1.701)	(.451)	(2.216)	(2.237)	(1.028)	
Panel (d): Weak exogeneity						
LR(4)	16.47^{***}	22.87^{***}	33.38^{***}	7.54	43.07^{***}	

Lag	1	2	3	4	5
Panel (e): Test for residual autocorrelation					
	3.398	9.366	66.174^{***}	148.79^{***}	196.38^{***}
Panel (f): Test for ARCH effects					
	2081.5^{***}	2937.7^{***}	3790.5^{***}	4971.8^{***}	5529.4^{***}

Note: Panel (a) reports Johansen (1988) cointegration tests. The critical values are taken from Osterwald-Lenum (1992). p-values are computed using a simulation study with 10,000 replications. Panel (b) displays the estimates of the cointegrating vectors. Insignificant variables have been excluded from the cointegrating vector. Panel (c) reports the estimates of the adjustment coefficients with t-statistics in parentheses. Estimates of the short-run dynamics, drift terms and variance-covariance matrix are not shown to conserve space. Panel (d) reports weak exogeneity tests. The likelihood ratio (LR) statistics are χ^2 distributed with degrees of freedom in parentheses. Panel (e) shows the results of vector portmanteau tests of the residuals. Panel (f) shows the results of tests for ARCH effects.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

We now briefly turn to the results of the linear VECM to obtain a useful summary of the 'average' adjustment dynamics provided by a linear specification. The adjustment coefficients of the linear VECM are reported in panel (c) of Table 5.2. A surprising feature of the results is the adjustment of the cointegrated system to the WTI-Brent price differential. Neither WTI, nor Brent adjust strongly to the deviations from their long-run equilibrium. By contrast, Bonny Light and Dubai react to deviations from the WTI-Brent price differential in the previous period. Tests for weak exogeneity of particular crude oil prices are presented in panel (d). The tests suggest weak exogeneity of Dubai, although it adjusts significantly to the WTI-Brent and WTI-Bonny Light price differentials. This discrepancy can be attributed to a generally lower power of Wald-type statistics. WTI is found to adjust to all price differentials except WTI-Dubai. Hence, WTI does not seem to be an exogenous price setter although it is the most closely watched benchmark crude oil price in the US.

5 .3 Markov-switching error correction models

Given the evolution of the market conditions, described in Section 2 , we suspect that the adjustment coefficients among crude oil prices do not remain constant over time and therefore consider a MSVECM which allows the model parameters to change between different regimes. As noted previously, the model specification of the MSVECM in terms of number of states is typically not clear a priori. Therefore, we consider both a two-state and a three-state specification and choose the final model specification based on the BIC[7]. Further, in line with the principle of parsimony, we reduce the number of parameters to estimate by testing whether allowing a switching behavior in a parameter matrix improves the model with regard to the BIC. More specifically, in the two-state MSVECM with two lags, henceforth MS(2)VECM(2), the vector of drift terms is restricted to be constant over both states and the variance-covariance matrix Σ is allowed to switch over states. In the three-state MSVECM with two additional lags, henceforth MS(3)VECM(2), we impose constancy of the drift terms and allow for a switching scale of the variance-covariance matrix. A comparison between the MS(2)VECM(2)[8] and MS(3)VECM(2) based on the BIC suggest that the increased goodness-of-fit of a three-state MSVECM indeed outweighs the increasing number of parameters. The regime-specific adjustment parameters for the MS(3)VECM(2) are reported in Table 5.3. We have excluded the short-run dynamics to conserve space and focus on the adjustment to the long-run equilibria. We find evidence for distinct regime-switching, reflected by non-zero transition probabilities and a state variable that assumes state 1 in 17%, state 2 in 15% and state 3 in 68% of the sample period. We refer to those points in time in which the model is confident of being in state 1 as regime 1 (R1), in state 2 as regime 2 (R2) and in state 3 as regime 3 (R3). Smoothed probabilities reflect the estimated probabilities of occurrence of each state at each point in time. This allows us to gain insights into the evolution of the adjustment dynamics over time. The smoothed probabilites are depicted in Figure 5.3.

The cointegrated system seems to be predominantly in state 1 at the beginning of the observational period. The first regime, thus, comprises almost exclusively of the first part of the sample, reaching from 1987 to 1994 and we refer to this as the ‘early regime’.[9] High probabilities of state 2 can be linked to exogenous global events and volatile economic environments. Probabilities close to one coincide with, among others, the period around the events of September 11, 2001, the period after the invasion of Iraq in 2003, and the Financial Crisis beginning in 2008. The second regime can therefore be associated with volatile economic and geopolitical times, hence we call it the ‘crisis regime’. The remaining regime associated with state 3 is referred to as the ‘tranquil regime’ and reflects behavior of the system in periods of relative calm.

We investigate the role of each crude oil price in all three regimes. The regime-specific dynamics help us to obtain new insights regarding the changing roles of regional crudes in the world crude oil market.

We report the results of regime-specific and overall weak exogeneity test in panel (c) of Table 5.3. We find no evidence against the null hypothesis of weak exogeneity of WTI in the ‘early regime’ and in the ‘tranquil regime’ during the later parts of the sample period. However, WTI adjusts significantly to

[7]Higher order MSVECM ($q > 3$) are not in line with a parsimonious model specification

[8]The results for the MS(2)VECM(2) specification are reported in Table 5.4 in the appendix.

[9]Please note that the labelling of the regimes primarily serves the purpose of illustration. The transition probabilities are estimated to be nonzero. Hence, it is, for example, possible that the state variable takes value one at a later point in time and the system switches to the ‘early regime’ again.

Table 5.3: Markov-switching error correction model for major crude oil prices (three-state model).

	WTI			Brent			Bonny Light			Dubai			Tapis		
	R1	R2	R3	R1	R2	R3	R1	R2	R3	R1	R2	R3	R1	R2	R3
Panel (a): Switching adjustment coefficients															
$\alpha_1(s_t)$	−.030	.083	.073*	−.059	.136	.115***	.016	.090	.127***	−.067	.065	.057	−.099***	−.312***	−.108***
	(−.620)	(.673)	(1.740)	(−1.080)	(1.080)	(2.630)	(.314)	(.725)	(2.920)	(−1.300)	(.562)	(1.520)	(−4.660)	(−3.430)	(−3.240)
$\alpha_2(s_t)$	.019	.031	.039	.004	.052	.095**	−.037	.073	.096**	−.076	−.007	.002	−.009	.137**	.039
	(.368)	(.383)	(1.020)	(.077)	(.641)	(2.390)	(−.727)	(.930)	(2.390)	(−1.460)	(−.087)	(.072)	(−.369)	(2.330)	(1.300)
$\alpha_3(s_t)$	.094	−.944***	.052	.114	−.740**	.006	−.301*	−.953***	−.060	−.053	−.735**	−.012	.119*	−.402	.197*
	(.615)	(2.780)	(.347)	(.686)	(−2.190)	(.039)	(−1.950)	(−2.880)	(−.393)	(−.335)	(−2.320)	(−.091)	(1.780)	(−1.630)	(1.660)
$\alpha_4(s_t)$	.045	1.236***	−.150	−.060	.714*	−.254	.321**	.969**	−.200	.255	.821**	−.073	−.000	.661	−.162
	(.287)	(3.180)	(.951)	(−.350)	(1.830)	(−1.560)	(2.000)	(2.510)	(−1.230)	(1.570)	(2.260)	(−.515)	(−.003)	(2.370)	(−1.290)
Panel (b): Weak exogeneity															
LR(4)	5.770	17.128***	5.858	1.251	6.980	19.261***	4.428	10.661**	20.992***	7.088	5.970	5.029	23.571***	21.208***	12.919**
LR(12)		29.252***			30.275***			38.908***			17.759			58.702***	

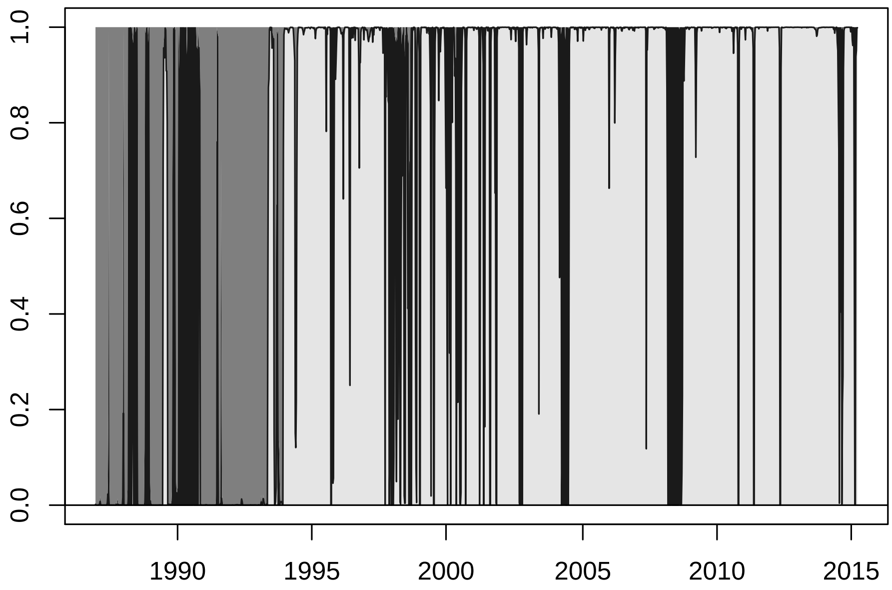
Lag	1	2	3	4	5	6	7	8	9	10
Panel (c): Test for residual autocorrelation										
	4.615	13.313	51.284	89.917	121.29	142.67	177.06	212.82	250.97	275.12
	(.999)	(.999)	(.984)	(.755)	(.577)	(.652)	(.442)	(.254)	(.113)	(.132)
Panel (d): Test for ARCH effects										
	1.631	1.352	1.314	1.316	1.303					
	(.025)	(.050)	(.037)	(.020)	(.014)					

	R1	R2	R3
Panel (e): Transition probabilities			
R1	.952	.200	.021
R2	.042	.761	.039
R3	.005	.043	.940

Note: R1 refers to the 'early regime', R2 to the 'crisis regime' and R3 to the 'tranquil regime', respectively. Panel (a) reports the estimates of the adjustment coefficients for three regimes with t-statistics in parentheses. The estimated cointegrating vectors are identical to panel (a) in Table 5.2. Estimates of the short-run dynamics, drift terms and variance-covariance matrix are not shown to conserve space. Panel (b) reports weak exogeneity tests for each regime (first row) and over all three regimes (second row). The likelihood ratio (LR) statistics are χ^2 distributed with degrees of freedom in parentheses. Panel (c) shows the results of vector portmanteau tests of the residuals with p-values in parentheses. Panel (d) shows the results of tests for ARCH effects with p-values in parentheses. Panel (e) displays the estimated transition probabilities.

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Figure 5.3: Smoothed probabilities MS(3)VECM(2).



This figure shows the probabilities for the cointegrated system being in the 'early regime' (grey), probabilities of being in the 'crisis regime' (black) and probabilities of being in the 'tranquil regime' (light-grey). The probabilities sum up to one in each period.

the WTI/Bonny Light and WTI/Brent price differential in the 'crisis regime'. The hypothesis of overall weak exogeneity is rejected which can be attributed to the significant adjustment in the 'crisis regime'. In other words, WTI seems to react to other crude oil prices primarily in times of uncertainty about future supply and demand. Brent is a weakly exogenous variable in the 'early regime' and the 'crisis regime'. However, Brent adjusts to the WTI/Tapis and WTI/Dubai price differentials in the 'tranquil regime'. Bonny Light is weakly exogenous in the 'early regime', adjusts to WTI/Bonny Light and WTI/Brent price differentials in the 'crisis regime' and to the WTI/Tapis and WTI/Dubai price differentials in the 'tranquil regime'. These findings suggest that WTI and Brent are important signals of world crude oil market news for Bonny Light in crisis periods whereas the price differentials with the Arabian Dubai and the Asian Pacific Tapis are constant factors in the price determination of Bonny Light. This can in parts be explained by the fact that Dubai is a close regionally competitor to the Nigerian Bonny Light. A reaction to its WTI price differential is attributed to the fact that the US is the largest importer of Nigerian crude oil so that US crude oil demand shocks are transmitted to the price of Bonny Light.

Dubai is the only weakly exogenous variable in all regimes. The results of the overall weak exogeneity test for Dubai in the three-state model is in line with the findings for the two-state MSVECM and the linear model (see panel (d) in Table 5.2 and panel (c) in Table 5.4). Also, an alternative normalization in which Dubai is allowed to be an exogenous variable in each equation left the results virtually unchanged. The results of this model are reported in Table 5.6 in the appendix. Economically, the result implies that Dubai acts as a price setter in this set of benchmark crude oil prices. Finally, Tapis is a price taker in all three states.

The orthogonalized impulse response functions[10] are displayed in Figure 5.4a and Figure 5.4b. We find that shocks to one variable in the 'early regime' do not evoke strong responses from the other variables. In contrast, shocks in the 'crisis regime' lead to visible reactions of the system. Adjustment to shocks is relatively fast whereas it takes the system more time to adjust to shocks in the 'tranquil regime'. These findings are in line with Ji and Fan (2015) who document stronger market integration if global exogenous shocks occur.

6 Discussion

Overall, the results are in line with the findings of Lu *et al.* (2014) and Ji and Fan (2015), indicating a stronger market integration in turbulent times. While a globally stable oil market promotes the use of nearby oil fields with lower transportation costs, extreme economic conditions create incentives to re-evaluate the attractiveness of different crude oil sources. Therefore, crude oil prices have to incorporate global information beyond the regional supply and demand changes.

Furthermore, the allocation of regime 1 to the earlier part of our sample, helps to emphasize the evolution of the world crude oil market. With the exception of Tapis, we do not reject weak exogeneity for any crude oil in the 'early regime'. The later part of the sample is partitioned into the 'tranquil regime' and the 'crisis regime', so that either Brent and Bonny Light adjust to long-term equilibria in tranquil times or WTI adjusts to its WTI/Brent and WTI/Bonny Light price differentials to maintain a long-run equilibrium relationship under extreme economic conditions. Dubai's price setting role supports the hypothesis in Bentzen (2007) which states that OPEC prices are gaining influence in the world crude oil market.

Similar to our results, Guelen (1999) finds that crude oil market integration is not stable and is especially strengthened during tight market conditions. His results, however, rely on a pre-specified structural break (the full sample is divided into two subperiods 1991-1993 and 1994-1996). Our study, following a more flexible approach, reveals that focusing only on the magnitude of prices does not seem to provide a more comprehensive picture of the crude oil market dynamics. Specifically, the application of a Markov-switching model to a longer and more varied sample period shows that crude oil market integration is strengthened in periods following geopolitical and economic events. The prices of benchmark crude oil reflect changing market conditions and, for example, tend to increase if supply is uncertain, but we document faster adjustment primarily in high volatility periods.

Moreover, the extent of market integration seems to coincide with the level of macroeconomic and financial uncertainty. To illustrate our notion, we compare the occurrence of the 'crisis regime' with two measures for financial and economic uncertainty. First, we contrast the evolution of the state indicator variable with the CBOE Volatility Index (VXO) which is based on 30-day S&P 100 index at-the-money options. It is a widely used measure for uncertainty in the financial market and has the advantage over other uncertainty measures that it spans the full sample period and is available at weekly frequency. The VXO, however, primarily measures uncertainty in the financial markets while *economic* uncertainty may also be influenced by fluctuations in real activity.

[10]The ordering of the variables which is used for the Cholesky decomposition is given as follows: Dubai $\rightarrow$ WTI $\rightarrow$ Brent $\rightarrow$ Bonny Light $\rightarrow$ Tapis.

Figure 5.4a: Regime-specific orthogonalized impulse response functions for one standard deviation shock in Dubai, WTI, Brent, Bonny Light and Tapis. The dotted, dashed and solid lines represent the OIRF in the ‘early regime’, the ‘crisis regime’ and the ‘tranquil regime’, respectively.

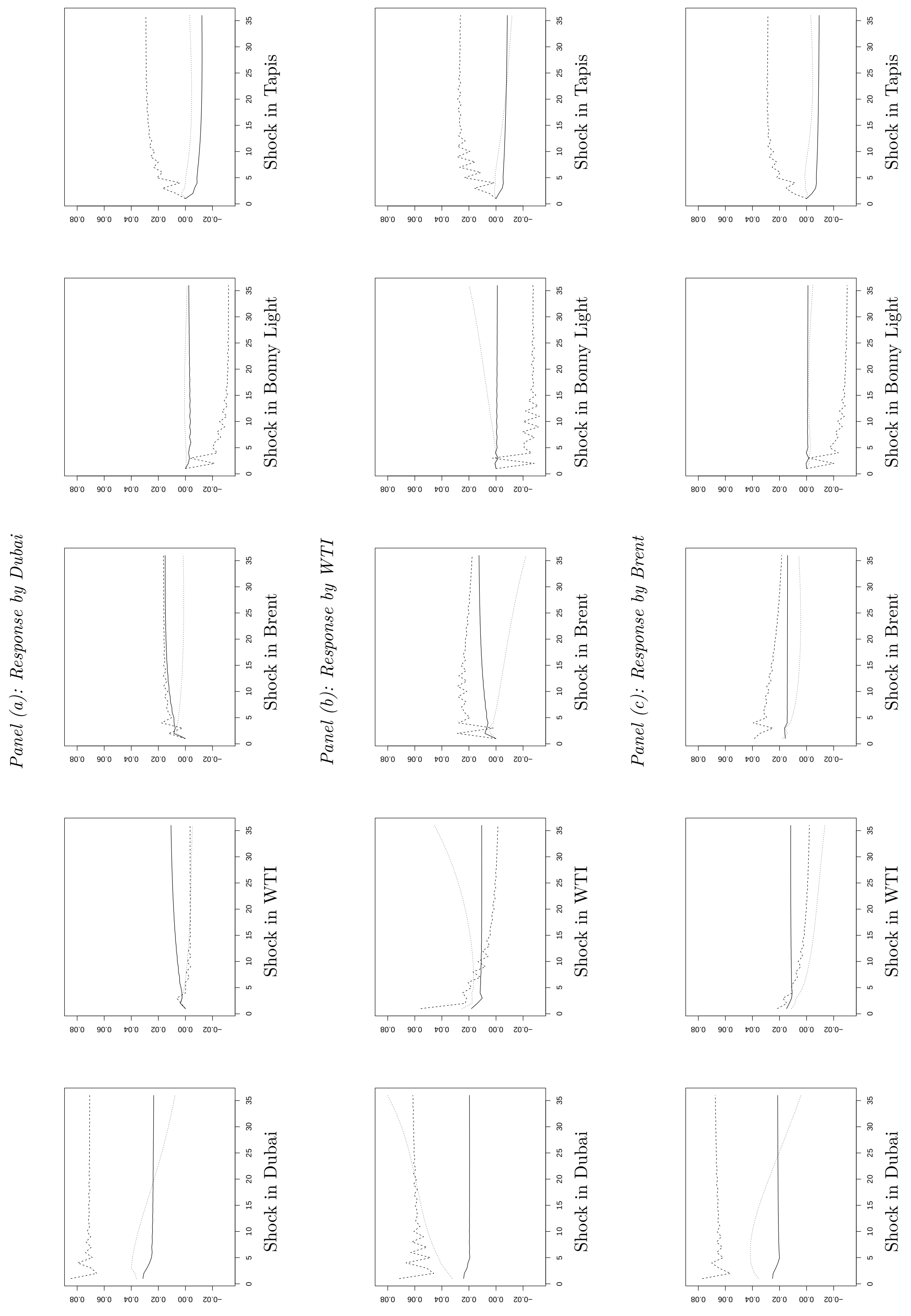
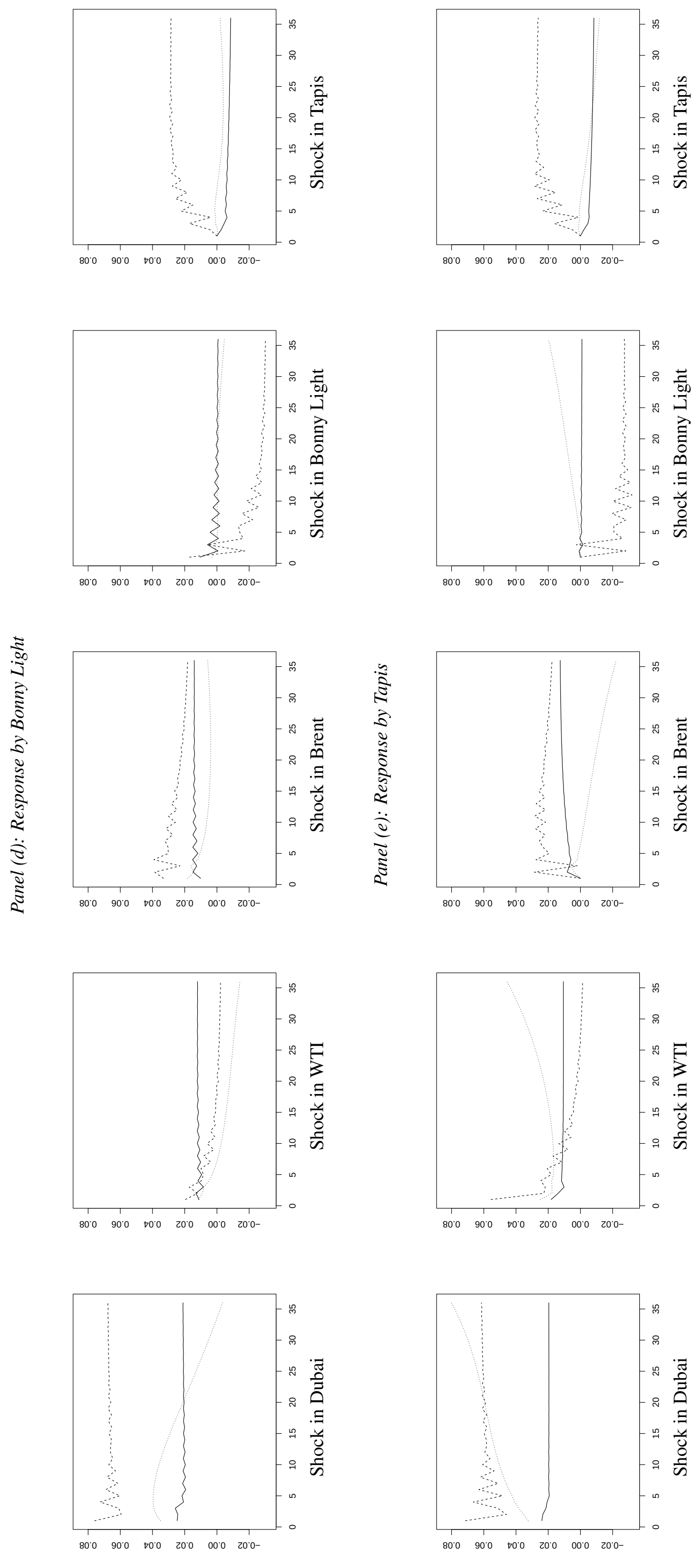
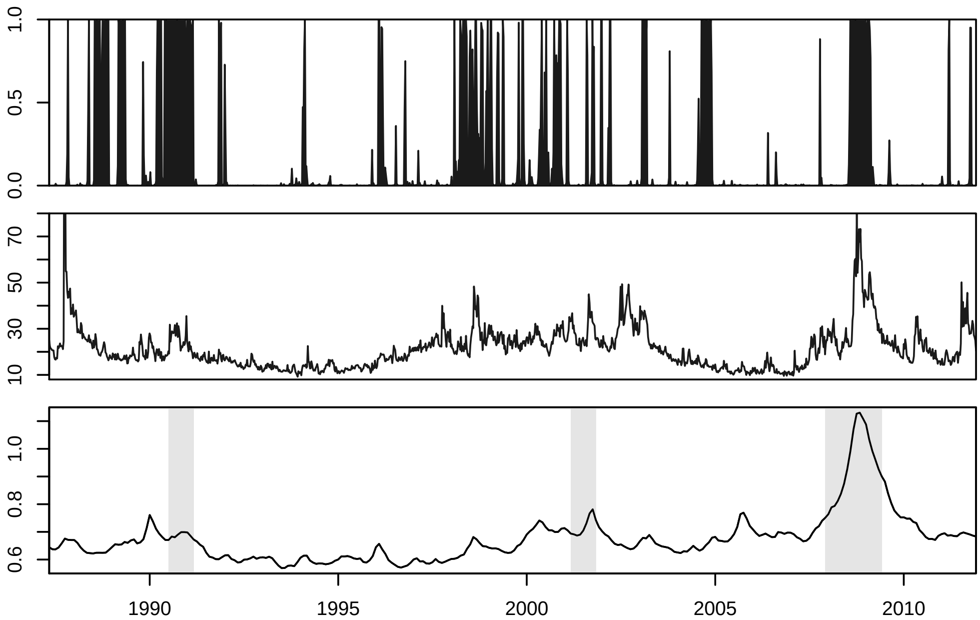


Figure 5.4b: Regime-specific orthogonalized impulse response functions for one standard deviation shock in Dubai, WTI, Brent, Bonny Light and Tapis (continued). The dotted, dashed and solid lines represent the OIRF in the 'early regime', the 'crisis regime' and the 'tranquil regime', respectively.



Second, we therefore also compare the occurrence of 'crisis' episodes in the crude oil market with a measure for macroeconomic uncertainty, recently developed by Jurado *et al.* (2015). This new measure for macroeconomic uncertainty essentially is an index based on various indicators including real output and income, unemployment, consumer spending

Figure 5.5: Smoothed probabilities of the 'crisis regime' and uncertainty measures.



Source: Jurado *et al.* (2015), NBER (2017)

This figure compares the smoothed probabilities of the cointegrated system being in the 'crisis regime' (row one) with the CBOE Volatility Index (row two) and the measure for macroeconomic uncertainty (grey shaded area: NBER recession dates) by Jurado *et al.* (2015) (row three).

and foreign exchange measures. The smoothed probabilities for the 'crisis regime' and our uncertainty measures are depicted graphically in Figure 5.5. It is obvious that the occurrence of the 'crisis regime' matches various peaks in the VXO, particularly, after the stock market crash in 1987, during the Persian Gulf crisis 1990-1991, the September 11, 2001 attack in the US, the 2003 Iraq war and the Financial Crisis starting late 2007. Likewise, peaks in macroeconomic uncertainty match 'crisis' episodes in the crude oil market. Compared to the VXO, Jurado *et al.* (2015)'s measure for macroeconomic uncertainty, however, is much smoother and its relation with the 'crisis regime' appears to be generally less pronounced. Finally, we consider the linear relation between the VXO and the 'crisis regime' indicator.[11] The contemporary correlation of the two time series is 0.277.

In essence, these findings provide descriptive evidence for a link between global economic uncertainty and world crude oil market integration. While they support our notion they do not enable an inferential analysis which we leave for future research.

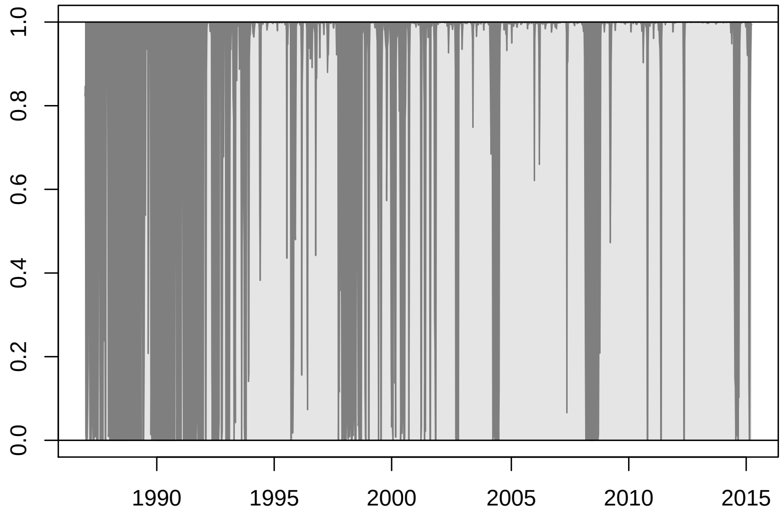
[11] Computing correlations between our state indicator variables and the measure for macroeconomic uncertainty is not possible due to different data frequencies.

7 Conclusion

This study provides a dynamic perspective on crude oil market integration. We employ a Markov regime-switching model based on the vector error correction model to study regime-switching adjustment behavior to constant long-run equilibria. Thereby, we identify three regimes to describe the adjustment behavior in different market conditions. The results highlight the changing landscape of the world crude oil markets. While the crude oil prices did not seem to maintain a long-run equilibrium from 1987 to 1994, the degree of crude oil market integration has strengthened in the later part of the sample. However, the roles of price setter and price taker can change drastically depending on the state of the global economy. Moreover, the results reveal the important role of Dubai as a price setter. Understanding crude oil market dynamics should therefore not be confined to a precise monitoring of WTI and Brent prices but should include Dubai as a third important benchmark price. Although the relationship between crude oil benchmark prices is changing over time, we do not find evidence for a decoupling of the WTI benchmark after the introduction of hydraulic fracturing to the shale oil fields of the US. It seems, that instead global events trigger adjustment to other regional benchmarks, thereby increasing world crude oil market integration.

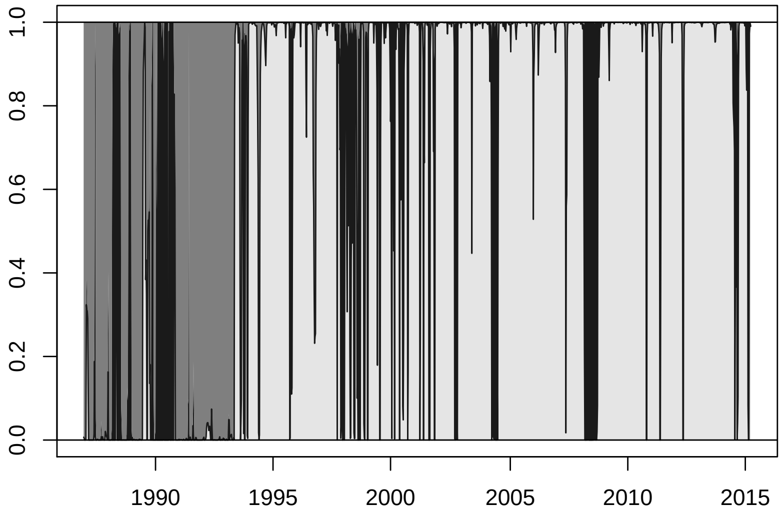
C Appendix

Figure 5.6: Smoothed probabilities MS(2)VECM(2).



This figures depicts the probabilities for the cointegrated system being in regime 1 (grey) and probabilities of being in regime 2 (light-grey). The probabilities sum up to one in each period.

Figure 5.7: Smoothed probabilities MS(3)VECM(2) (Dubai normalization).



Probabilities for the cointegrated system being in the ’early regime’ (medium-grey), probabilities of being in the ’crisis regime’ (dark-grey) and probabilities of being in the ’tranquil regime’ (light-grey). The probabilities sum up to one in each period.

Table 5.4: Markov-switching error correction model for major crude oil prices (two-state model).

	WTI		Brent		Bonny Light		Dubai		Tapis	
	R1	R2	R1	R2	R1	R2	R1	R2	R1	R2
Panel (a): Switching adjustment coefficients										
$\alpha_1(s_t)$	.039	.092**	.064	.142***	.052	.155***	.031	.063*	−.227***	−.097***
	(.553)	(2.340)	(.884)	(3.550)	(.731)	(3.870)	(.454)	(1.780)	(−4.560)	(−3.110)
$\alpha_2(s_t)$	.022	.055	.032	.122***	.036	.124***	−.024	.007	.093**	.038
	(.422)	(1.550)	(.600)	(3.350)	(.683)	(3.400)	(−.468)	(.224)	(2.500)	(1.330)
$\alpha_3(s_t)$	−.538***	.044	−.431**	−.015	−.709***	−.082	−.497**	−.040	−.154	.174
	(−2.630)	(.319)	(−2.070)	(−.108)	(−3.470)	(−.598)	(−2.480)	(−.326)	(−1.070)	(1.580)
$\alpha_4(s_t)$	.772***	−.178	.436*	−.289**	.735***	−.235	.589***	−.053	.321**	−.146
	(3.410)	(−1.230)	(1.890)	(−1.970)	(3.240)	(−1.610)	(2.670)	(−.414)	(2.010)	(−1.260)
Panel (b): Weak exogeneity										
LR(4)	22.131***	10.941**	6.328	34.432***	14.024***	37.712***	7.319	5.571	28.873***	11.757**
LR(8)	35.555***		42.029***		52.948***		12.586		41.605***	

Lag	1	2	3	4	5	6	7	8	9	10
Panel (c): Test for residual autocorrelation										
	3.090	7.793	44.582	94.171	118.78	148.50	189.55	228.23	275.35**	297.66**
Panel (d): Test for ARCH effects										
	4.028***	3.241***	2.704***	3.616***	3.413***					

	R1	R2
Panel (e): Transition probabilities		
R1	.938	.140
R2	.062	.860

Note: Panel (a) reports the estimates of the adjustment coefficients for two regimes (R1 and R2) with t-statistics in parentheses. The estimated cointegrating vectors are identical to panel (a) in Table 5.2. Estimates of the short-run dynamics, drift terms and variance-covariance matrix are not shown to conserve space. Panel (b) reports weak exogeneity tests for each regime (first row) and over both regimes (second row). The likelihood ratio (LR) statistics are χ^2 distributed with degrees of freedom in parentheses. Panel (c) shows the results of vector portmanteau tests of the residuals. Panel (d) shows the results of tests for ARCH effects. Panel (e) displays the estimated transition probabilities.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 5.5: Cointegration tests and linear VECM (Dubai normalization).

	$N-r$	r	Eig.value	Trace	5% Crit. val.	p-Value
Panel (a): I(1)-analysis						
	5	0	.1084	361.75	76.07	.000
	4	1	.0651	192.16	53.12	.000
	3	2	.0374	92.61	34.91	.000
	2	3	.0292	36.22	19.96	.000
	1	4	.0013	1.95	9.24	.783

	WTI	Brent	Bonny	Dubai	Tapis	μ
Panel (b): Cointegration vectors						
β_1				−.958	1	−.284
β_2			1	−.966		−.209
β_3		1		−.963		−.200
β_4	1			−.881		−.515
Panel (c): Adjustment coefficients						
α_1	.066*	.104***	.110***	.049	−.162***	
	(1.879)	(2.974)	(3.159)	(1.514)	(−6.198)	
α_2	−.214*	−.229**	.432***	−.272***	−.001	
	(−1.933)	(−2.062)	(−3.915)	(−2.638)	(−.018)	
α_3	.198*	.053	.257**	.242**	.090	
	(1.701)	(.451)	(2.216)	(2.237)	(1.028)	
α_4	−.086***	.007	.001	−.003	.011	
	(−3.254)	(.277)	(.050)	(−.110)	(0.570)	
Panel (d): Weak exogeneity						
LR(4)	16.47***	22.87***	33.38***	7.54	43.07***	

Lag	1	2	3	4	5
Panel (e): Test for residual autocorrelation					
	3.398	9.366	66.174***	148.79***	196.38***
Panel (f): Test for ARCH effects					
	2081.5***	2937.7***	3790.5***	4971.8***	5529.4***

Note: Panel (a) reports Johansen (1988) cointegration tests. The critical values are taken from Osterwald-Lenum (1992). p-values are computed using a simulation study with 10,000 replications. Panel (b) displays the estimates of the cointegrating vectors. Insignificant variables have been excluded from the cointegrating vector. Panel (c) reports the estimates of the adjustment coefficients with t-statistics in parentheses. Estimates of the short-run dynamics, drift terms and variance-covariance matrix are not shown to conserve space. Panel (d) reports weak exogeneity tests. The likelihood ratio (LR) statistics are χ^2 distributed with degrees of freedom in parentheses. Panel (e) shows the results of vector portmanteau tests of the residuals. Panel (f) shows the results of tests for ARCH effects.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 5.6: Markov-switching error correction model for major crude oil prices (three-state model, Dubai normalization).

	WTI			Brent			Bonny Light			Dubai			Tapis		
Panel (a): Switching adjustment coefficients															
	R1	R2	R3	R1	R2	R3	R1	R2	R3	R1	R2	R3	R1	R2	R3
$\alpha_1(s_t)$	−.018	.080	.068	−.036	.102	.122***	.016	.063	.134***	−.030	.025	.055	−.144***	−.312***	−.105***
	(−.393)	(.535)	(1.590)	(−.719)	(.666)	(2.650)	(.305)	(.405)	(2.910)	(−.658)	(.175)	(1.440)	(−3.660)	(−2.700)	(−3.280)
$\alpha_2(s_t)$	−.073	−.826**	.072	−.053	−.645*	.009	−.479***	−.835**	−.055	−.217	−.615*	−.015	.035	−.360	.202*
	(−.453)	(−2.220)	(.495)	(−.313)	(−1.740)	(.062)	(−2.660)	(−2.220)	(−.369)	(−1.390)	(−1.750)	(−.115)	(.256)	(−1.270)	(1.860)
$\alpha_3(s_t)$	.166	1.124***	−.176	.094	.627	−.284*	.513***	.853**	−.232	.350**	.720*	−.079	.074	.611*	−.172
	(.968)	(2.670)	(−1.140)	(.524)	(1.470)	(−1.750)	(2.680)	(1.970)	(−1.430)	(2.160)	(1.800)	(−.552)	(.511)	(1.890)	(−1.480)
$\alpha_4(s_t)$	−.098	−.426***	−.018	.018	−.142	.036	.023	−.160	.035	−.022	−.127	.027	.027	−.084	.033*
	(−1.450)	(−3.320)	(−.738)	(.256)	(−1.120)	(1.450)	(.315)	(−1.240)	(1.420)	(−.338)	(−1.050)	(1.210)	(.488)	(−.852)	(1.740)
Panel (b): Weak exogeneity															
LR(4)	3.277	13.250**	5.962	.962	4.173	18.255***	7.744	6.271	19.733***	6.013	3.491	4.816	14.596***	12.992**	13.814***
LR(12)		23.029**			26.316***			36.293***			13.450			42.970***	

Lag	1	2	3	4	5	6	7	8	9	10
Panel (c): Test for residual autocorrelation										
	5.615	12.342	43.175	82.409	103.99	126.05	138.06	196.39	229.13	250.39
	(.999)	(.999)	(.999)	(.999)	(.914)	(.923)	(.816)	(.559)	(.411)	(.481)
Panel (d): Test for ARCH effects										
	2.583	2.382	2.317	2.048	1.964					
	(.000)	(.000)	(.000)	(.000)	(.000)					

	R1	R2	R3
Panel (e): Transition probabilities			
R1	.952	.182	.047
R2	.037	.770	.033
R3	.012	.048	.919

Note: R1 refers to the 'early regime', R2 to the 'crisis regime' and R3 to the 'tranquil regime', respectively. Panel (a) reports the estimates of the adjustment coefficients for three regimes with t-statistics in parentheses. The estimated cointegrating vectors are identical to panel (a) in Table 5.5. Estimates of the short-run dynamics, drift terms and variance-covariance matrix are not shown to conserve space. Panel (b) reports weak exogeneity tests for each regime (first row) and over all three regimes (second row). The likelihood ratio (LR) statistics are χ^2 distributed with degrees of freedom in parentheses. Panel (c) shows the results of vector portmanteau tests of the residuals with p-values are given in brackets. Panel (d) shows the results of tests for ARCH effects with p-values are given in brackets. Panel (e) displays the estimated transition probabilities.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Chapter 6

Critical assessment and conclusion

This thesis demonstrates the usefulness of modelling nonlinear dynamics in the context of cointegration relationships. Each of the main chapters provides an extension of the literature either in terms of a theoretical contribution to cointegration analysis or an innovative empirical application of nonlinear cointegration models to open research questions. The main findings of the four studies are summarized in their respective chapters. In the following, the studies are placed in context of related work, remaining methodological shortcomings are discussed and opportunities for future research are highlighted:

Chapter 2 focusses on the second step of the Engle-Granger procedure. This study contributes to the literature by developing a new econometric model and demonstrating its application to price transmissions in the US and German fuel markets. Whereas the majority of studies on asymmetric price transmission apply threshold cointegration model or asymmetric error correction models and test for significantly different mean-reversion in pre-specified regimes, we propose to model a quantile-dependent adjustment behaviour of the residual process. This allows us to quantify the degree of asymmetry in the adjustment to equilibrium errors without pre-specifying the number of regimes and without estimating threshold values. Instead, optimizing the loss function of a quantile regression automatically results in quantile-specific adjustment coefficients. Therefore, only the quantile has to specified to obtain, for instance, the adjustment behavior for large negative and large positive deviations from equilibrium. This approach is inspired by and uses results from the literature on quantile regression with times series.

The drawbacks of this model are directly related to the quantile regression methodology employed therein: The estimates of the adjustment coefficients cannot be used to forecast directly. The interpretation of the quantile regression estimates are given as a varying response conditional on the quantile of the dependent variable in period t. Since the state of the dependent variable is not known in period $t-1$, it is not possible to obtain point forecasts. Still, it remains an open question whether these estimates could be used for probabilistic forecasting. Another drawback related to the quantile regression technique is that estimates for different quantiles have a different degree of uncertainty assigned to them. Extreme quantiles are more difficult to estimate which results in broader confidence bands. However, this problem is also present in threshold models if a regime does not consist of a sufficiently large number of observations.

We propose a bootstrap cointegration test and the appropriate statistical tests for across-quantile comparisons and overall quantile effects. While the cointegration test is analyzed with a 'Warp-speed' simulation study employing a realistic data-generating process, the block bootstrap technique used for the follow-up analysis is based on the assumption that the equilibrium error series is stationary. Since the null hypothesis of no cointegration is rejected for all cointegration pairs, the test should remain valid.

Our empirical results suggest that asymmetries can be found in the early stages of the production chain but are not completely transferred to retail prices. This finding coincides with the theoretical expectation that the retail fuel market with numerous competitors should be more competitive than the fragmented market for fuel prestage products. Further research needs to be conducted using data from different fuel markets to test if the findings can be reproduced for smaller, less developed, fuel markets.

Chapter 3 investigates whether gold and silver are cointegrated and why previous studies on this issue produced ambiguous results. First, we discuss reasons for a long-run relationship between these precious metals which have very different industrial uses. Their substitutability on financial markets leads us to the conjecture that the relationship might be stronger in periods of financial stress and economic crisis - periods in which precious metals are particularly sought-after investments. The cointegration model should therefore account for time-varying coefficients. We address the constancy of parameters in the cointegrating regression and apply a quantile cointegration model to gold and silver prices. Under the restrictions of linear cointegration, we find sufficient evidence against the null hypothesis of cointegration, but we do not find evidence against a nonlinear long-run relationship. The quantile cointegration estimates reveal substantial asymmetry in the relationship. The response of silver (gold) prices to gold (silver) price changes is stronger if silver (gold) prices are at a high level. While previous studies could only find traces of cointegration in subsamples, which had to be pre-specified, the quantile cointegration model determines the parameter changes from the data. The periods identified by these studies match with the periods of stronger responses identified in the quantile cointegration framework.

Similarly to the study presented in Chapter 2, the quantile cointegration framework has the disadvantage that forecasting is difficult. This is particularly unfortunate considering the fact that the long-run relationship between gold and silver could be used in a trading strategy. Moreover, quantile regression produces weighted residuals which cannot be used for conventional error correction models. It is therefore not possible to analyze which price leads the quantile cointegration relationship. Although the robustness of our results across different frequencies and markets has been shown, further research could be directed at the robustness of our results to different nonlinearity concepts. For example, a Markov-switching cointegration models could be applied to gold and silver prices, to see whether the state-dependence of the long-run relationship is also found for the adjustment behaviour.

Chapter 4 mainly contributes to the field of theoretical time series econometrics. We develop a new cointegration test with SETAR adjustment allowing for the presence of a structural break in the cointegrating vector. This test is residual-based and extends the Engle-Granger framework at both steps. Modelling structural breaks of the cointegrating regression accounts for the possibility of multiple equilibria and modelling the adjustment behavior with a SETAR model accounts for nonlinear responses to equilibrium errors caused, for example, by transaction costs or collusive agreements. We derive the asymptotic distribution of the test statistics and analyze the finite sample properties of the test. This test could be a useful tool for researchers working on asymmetric price transmission models in an unstable environment as demonstrated by the empirical example using US fuel market data. Since the procedure to determine the timing of the structural break is based on a statistic of the auxiliary regression, the test is easily modified to account for multiple structural breaks and alternative structural break models. However, the computational costs of higher dimensional grid searches might restrict these considerations.

Further research needs to be conducted on cointegration models with MTAR adjustment under the presence of structural breaks. This would extend the second test given in Enders and Siklos (2001). Although the model specification is similar, asymptotic theory in this case is potentially very different. Moreover, the restriction of a fixed threshold value needs to be relaxed to allow for empirical applications where the threshold value is unknown. Again, this leads to some theoretical difficulties which result from the test construction. Chan (1993) proved the consistency of a threshold estimate that is based on minimizing the sum of squared errors over a set of data-dependent threshold values. However, the structural breaks in our procedure are found as a supremum of Wald-type statistics or, conceptionally, as a model specifications that speaks against the null hypothesis of no cointegration. Since the configuration of the structural break dummy alters the residual series, these two approaches cannot be merged straightforwardly.

Chapter 5 presents a Markov-switching approach to answer the question of whether the global crude oil market is unified or regionalized. We analyze the long-run relationship between the different regional crude oil benchmark prices and find that prices are reacting to each other. The degree of market integration, however, seems to be state-dependent. Two specifications of a MSVECM reveal that prices react more strongly to each other in periods of economic crises. Furthermore, the price leadership changes from one regime to the next. The most commonly used benchmark prices, WTI and Brent, are not always price setters. Instead, Dubai emerges as the only exogenous price in all regimes.

Although a Markov-switching approach presents a convenient way to model state-dependencies, the immediate drawback is the large number of parameters that have to be estimated. We show the robustness of our results with respect to the number of regimes and their allocation, using a two-regime and a three-regime model. Further model specifications with more regimes are not considered because the degrees of freedom would be reduced below a reasonable threshold in these cases. Also, the dynamic specification is altered and the results do not seem to be sensitive in this regard. Again, a lower number of lags is chosen to prevent a reduction in the degrees of freedom. We match the allocation of the 'crisis regime' with spikes in the volatility index (VXO) and a newly developed indicator for macroeconomic uncertainty, to underline the interpretation of a stronger market integration in periods of economic uncertainty. Further research might be directed to enable inferential analysis of this finding.

Overall, it should be emphasized that the purpose of the nonlinear extensions in this thesis is not to improve the statistical fit of a given cointegration model to the empirical data. Instead, the nonlinear cointegration models are motivated by economic theory. Long-run effects of asymmetric price transmissions, for example, cannot be investigated if the cointegration model does not allow for nonlinear adjustment behaviour. Furthermore, the restrictive nature of conventional cointegration models is not suitable for the analysis of long-run relationships which do not stay constant over the observational period. Chapter 2 deals with gradually changing speed of adjustment in the context of quantile regressions, while Chapter 4 improves existing threshold cointegration models under the presence of structural breaks which might be caused by policy changes, technological changes or events such as economic crises. Chapter 3 presents an empirical study in which a weaker concept than Engle-Granger cointegration is used to model the nonlinear long-run relationship between gold and silver which is motivated by a time-varying demand for precious metals as investments. Finally, Chapter 5 proposes the use of a Markov-switching VECM to model the long-run relationships on a volatile crude oil market. Thus, nonlinear cointegrations models are used to reveal information about the long-run relationship beyond the average behavior and aim to capture the true dynamics of these relationships more closely than their linear benchmark models.

Bibliography

ADELMAN, M. A. (1984). International oil agreements. *The Energy Journal*, **5** (3), 1–9.

AGYEI-AMPOMAH, S., GOUNOPOULOS, D. and MAZOUZ, K. (2014). Does gold offer a better protection against losses in sovereign debt bonds than other metals? *Journal of Banking and Finance*, **40**, 507–521.

AL-GUDHEA, S., KENC, T. and DIBOOGLU, S. (2007). Do retail gasoline prices rise more readily than they fall?: A threshold cointegration approach. *Journal of Economics and Business*, **59** (6), 560–574.

ALEXANDER, C. (1999). Optimal Hedging using cointegration. *Philosophical Transactions: Mathematical, Physical and Engineering Sciences*, **357** (1758), 2039–2058.

AMERICAN PETROLEUM INSTITUTE (2017). State Gasoline Tax Reports. `http://www.api.org/oil-and-natural-gas/consumer-information/motor-fuel-taxes`, [Online; accessed 10-February-2017].

ANDREWS, D. W. K. (1984). Non-Strong Mixing Autoregressive Processes. *Journal of Applied Probability*, **21** (4), 930–934.

— (1991). Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation. *Econometrica*, **59** (3), 817–858.

ARAI, Y. and KUROZUMI, E. (2007). Testing for the Null Hypothesis of Cointegration with a Structural Break. *Econometric Reviews*, **26** (6), 705–739.

AWIROTHANANON, T. and CHEUNG, W.-K. A. (2009). On Joint Determination of the Number of States and the Number of Variables in Markov-Switching Models: A Monte Carlo Study. *Communications in Statistics - Simulation and Computation*, **38** (8), 1757–1788.

BACHMEIER, L. J. and GRIFFIN, J. M. (2003). New Evidence on Asymmetric Gasoline Price Responses. *The Review of Economics and Statistics*, **85** (3), 772–776.

BACON, R. W. (1991). Rockets and feathers: The asymmetric speed of adjustment of UK retail gasoline prices to cost changes. *Energy Economics*, **13** (3), 211–218.

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

BANERJEE, A., DOLADO, J. J., HENDRY, D. F. and SMITH, G. W. (1986). Exploring Equilibrium Relationships in Econometrics through Static Models: Some Monte Carlo Evidence. *Oxford Bulletin of Economics and Statistics*, **48** (3), 253–277.

BAUR, D. G. and LUCEY, B. M. (2010). Is Gold a Hedge or a Safe Haven? An Analysis of Stocks, Bonds and Gold. *Financial Review*, **45** (2), 217–229.

— and MCDERMOTT, T. K. (2010). Is gold a safe haven? International evidence. *Journal of Banking & Finance*, **34** (8), 1886–1898.

— and TRAN, D. T. (2014). The long-run relationship of gold and silver and the influence of bubbles and financial crises. *Empirical Economics*, **47** (4), 1525–1541.

BENTZEN, J. (2007). Does OPEC influence crude oil prices? Testing for co-movements and causality between regional crude oil prices. *Applied Economics*, **39** (11), 1375–1385.

BERA, A. K., GALVAO, A. F. and WANG, L. (2014). On testing the equality of mean and quantile effects. *Journal of Econometric Methods*, **3** (1), 47–62.

BILLINGSLEY, P. (1999). *Convergence of Probability Measures*. New York: Wiley, 2nd edn.

BORENSTEIN, S., CAMERON, A. C. and GILBERT, R. (1997). Do Gasoline Prices respond asymmetrically to Crude Oil Price Changes. *The Quarterly Journal of Economics*, **112** (1), 305–339.

BUSETTI, F. and HARVEY, A. (2001). Testing for the presence of a random walk in series with structural breaks. *Journal of Time Series Analysis*, **22** (2), 127–150.

BUYUKSAHIN, B. and HARRIS, J. (2011). Do speculators drive crude oil futures prices? *The Energy Journal*, **32** (2), 167–202.

CANDELON, B., JOËTS, M. and TOKPAVI, S. (2013). Testing for Granger causality in distribution tails: An application to oil markets integration. *Economic Modelling*, **31**, 276–285.

CANER, M. and HANSEN, B. E. (2001). Threshold Autoregression with a Unit Root. *Econometrica*, **69** (6), 1555–1596.

CARRION-I SILVESTRE, J. L. and SANSO, A. (2006). Testing the Null of Cointegration with Structural Breaks. *Oxford Bulletin of Economics and Statistics*, **68** (5), 623–646.

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

CHANG, Y., PARK, J. Y. and SONG, K. (2006). Bootstrapping cointegrating regressions. *Journal of Econometrics*, **133** (2), 703–739.

CHARLES, A., DARNÉ, O. and KIM, J. H. (2015). Will precious metals shine? A market efficiency perspective. *International Review of Financial Analysis*, **41**, 284–291.

CHERNOZHUKOV, V., FERNANDEZ-VAL, I. and GALICHON, A. (2010). Quantile and probability curves without crossing. *Econometrica*, **78** (3), 1093–1125.

CHOI, I. and KUROZUMI, E. (2012). Model selection criteria for the leads-and-lags cointegrating regression. *Journal of Econometrics*, **169** (2), 224–238.

CINER, C. (2001). On the long run relationship between gold and silver prices: A note. *Global Finance Journal*, **12** (2), 299–303.

DAVIES, R. B. (1987). Hypothesis testing when a nuisance parameter is present only under the alternative. *Biometrika*, **74** (1), 33–43.

DOUGLAS, C. C. (2010). Do gasoline prices exhibit asymmetry? Not usually! *Energy Economics*, **32** (4), 918–925.

— and HERRERA, A. M. (2010). Why are gasoline prices sticky? A test of alternative models of price adjustment. *Journal of Applied Econometrics*, **25**, 903–928.

EKNER, L. E. and NEJSTGAARD, E. (2013). Parameter identification in the logistic STAR model. *Discussion Paper No. 13-07, Department of Economics of the University of Copenhagen*, pp. 1–22.

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.

— and SIKLOS, P. L. (2001). Cointegration and Threshold Adjustment. *Journal of Business & Economic Statistics*, **19** (2), 166–176.

ENGLE, R. F. and GRANGER, C. W. J. (1987). Co-Integration and Error Correction: Representation, Estimation and Testing. *Econometrica*, **55** (2), 251–276.

— and YOO, B. S. (1987). Forecasting and Testing in Co-integrated Systems. *Journal of Econometrics*, **35** (1), 143–159.

ESCRIBANO, A. and GRANGER, C. W. J. (1998). Investigating the relationship between gold and silver prices. *Journal of Forecasting*, **17** (2), 81–107.

EUROPEAN COMMISSION (2016). Renewable Energy Directive. `https://ec.europa.eu/energy/en/topics/renewable-energy/renewable-energy-directive`, [Online; accessed 05-August-2016].

EUROSTAT (2017). Passenger cars in the EU. `http://ec.europa.eu/eurostat/statistics-explained/index.php/Passenger_cars_in_the_EU`, [Online; accessed 10-February-2017].

FATTOUH, B. (2006). The origins and evolution of the current international oil pricing system: A critical assessment. In R. Mabro (ed.), *Oil in the 21st Century - Issues, Challenges and Opportunities*, *3*, Oxford: Oxford University Press, pp. 41–101.

— (2007). WTI Benchmark Temporarily Breaks Down: Is It Really a Big Deal? *Oxford Energy Comment*, pp. 1–7.

— (2010). The dynamics of crude oil price differentials. *Energy Economics*, **32** (2), 334–342.

FOSTEN, J. (2012). Rising household diesel consumption in the united states: A cause for concern? Evidence on asymmetric pricing. *Energy Economics*, **34** (5), 1514–1522.

FREY, G. and MANERA, M. (2007). Econometric Models of Asymmetric Price Transmission. *Journal of Economic Surveys*, **21** (2), 349–415.

FUELSEUROPE (2014). Statistical Report 2014. `https://www.fuelseurope.eu/uploads/Modules/Resources/statistical_report_fuels_europe-_v25_web.pdf`, [Online; accessed 25-September-2015].

GFMS (2016a). *World Gold Survey 2016*. London: Thomson Reuters.

— (2016b). *World Silver Survey 2016*. London: Thomson Reuters.

GIACOMINI, R., POLITIS, D. N. and WHITE, H. (2013). A Warp-Speed Method for Conducting Monte Carlo Experiments Involving Bootstrap Estimators. *Econometric Theory*, **29** (03), 567–589.

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

— and — (2006b). Threshold effects in multivariate error correction models. In H. Hassani, T. C. Mills and K. Patterson (eds.), *Handbook of Econometrics Volume 1: Econometric Theory*, *15*, Palgrave Macmillan UK, pp. 578–609.

GRANGER, C. W. J. (1981). Some Properties of Time Series Data and their use in Econometric Model Specification. *Journal of Econometrics*, **16** (1), 121–130.

— (1986). Developments in the Study of Cointegrated Economic Variables. *Oxford Bulletin of Economics and Statistics*, **48** (3), 213–228.

— and TERASVIRTA, T. (1993). *Modelling nonlinear economic relationships*. Oxford University Press.

GRASSO, M. and MANERA, M. (2007). Asymmetric error correction models for the oil–gasoline price relationship. *Energy Policy*, **35** (1), 156–177.

GREENWOOD-NIMMO, M. and SHIN, Y. (2013). Taxation and the asymmetric adjustment of selected retail energy prices in the UK. *Economics Letters*, **121** (3), 411–416.

GREGORY, A. W. and HANSEN, B. E. (1992). Residual-based tests for cointegration in models with regime shifts. *Queen's Economics Department Working Paper*, (862), 1–32.

— and — (1996a). Residual-based tests for cointegration in models with regime shifts. *Journal of Econometrics*, **70** (1), 99–126.

— and — (1996b). Tests for Cointegration in Models with Regime and Trend Shifts. *Oxford Bulletin of Economics and Statistics*, **58** (3), 555–560.

GUELEN, G. (1999). Regionalization in the World Crude Oil Market: Further Evidence. *The Energy Journal*, **20** (1), 125–139.

HALL, P. and HEYDE, C. C. (1980). *Martingale Limit Theory and Its Application*. Academic Press.

HALL, S., PSARADAKIS, Z. and SOLA, M. (1997). Switching error-correction models of house prices in the United Kingdom. *Economic Modelling*, **14** (4), 517–527.

HAMILTON, J. D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica*, **57** (2), 357–384.

— (1994). *Time Series Analysis*. Princeton University Press.

HAMMOUDEH, S. M., EWING, B. T. and THOMPSON, M. A. (2008). Threshold cointegration analysis of crude oil benchmarks. *The Energy Journal*, **29** (4), 79–95.

HANSEN, B. E. (1992). Tests for Parameter Instability in Regressions with I(1) Processes. *Journal of Business & Economic Statistics*, **10** (3), 321–335.

— and SEO, B. (2002). Testing for two-regime threshold cointegration in vector error-correction models. *Journal of Econometrics*, **110** (2), 293–318.

HARRIS, D., LEYBOURNE, S. J. and TAYLOR, A. M. R. (2016). Tests of the co-integration rank in VAR models in the presence of a possible break in trend at an unknown point. *Journal of Econometrics*, **192** (2), 451–467.

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

HONARVAR, A. (2010). Modeling of asymmetry between gasoline and crude oil prices: A monte carlo comparison. *Computational Economics*, **36** (3), 237–262.

HÖÖK, M., HIRSCH, R. and ALEKLETT, K. (2009). Giant oil field decline rates and their influence on world oil production. *Energy Policy*, **37** (6), 2262–2272.

INOUE, A. (1999). Tests of cointegrating rank with a trend-break. *Journal of Econometrics*, **90** (2), 215–237.

INTERCONTINENTAL EXCHANGE (2016). ICE Crude Oil. `https://www.theice.com/publicdocs/ICE_Crude_Oil.pdf`, [Online; accessed 05-August-2016].

JI, Q. and FAN, Y. (2015). Dynamic integration of world oil prices: A reinvestigation of globalisation vs. regionalisation. *Applied Energy*, **155**, 171–180.

JOHANSEN, S. (1988). Statistical Analysis of Cointegration Vectors. *Journal of Economic Dynamics and Control*, **12**, 231–254.

— (1991). Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models. *Econometrica*, **59** (6), 1551–1580.

—, MOSCONI, R. and NIELSEN, B. (2000). Cointegration analysis in the presence of structural breaks in the deterministic trend. *Econometrics Journal*, **3** (2), 216–249.

JURADO, K., LUDVIGSON, S. C. and NG, S. (2015). Measuring uncertainty. *American Economic Review*, **105** (3), 1177–1216.

KAPETANIOS, G., SHIN, Y. and SNELL, A. (2003). Testing for Cointegration in Nonlinear STAR Error Correction Models. *Queen's Economics Department Working Paper*, (497), 1–24.

—, — and — (2006). Testing for Cointegration in Nonlinear Smooth Transition Error Correction Models. *Econometric Theory*, **22**, 279–303.

KAUFMANN, R. K. and LASKOWSKI, C. (2005). Causes for an asymmetric relation between the price of crude oil and refined petroleum products. *Energy Policy*, **33** (12), 1587–1596.

KILIÇ, R. (2011). Testing for co-integration and nonlinear adjustment in a smooth transition error correction model. *Journal of Time Series Analysis*, **32** (6), 647–660.

KOENKER, R. and BASSETT, G. (1978). Regression quantiles. *Econometrica*, **46** (1), 33–50.

— and XIAO, Z. (2004). Unit root quantile autoregression inference. *Journal of the American Statistical Association*, **99** (467), 775–787.

— and — (2006). Quantile autoregression. *Journal of the American Statistical Association*, **101** (475), 980–990.

KOENKER, R. W. and D'OREY, V. (1987). Algorithm as 229: Computing regression quantiles. *Applied Statistics*, **36** (3), 383–393.

KRISHNAKUMAR, J. and NETO, D. (2015). Testing for the cointegration rank in threshold cointegrated systems with multiple cointegrating relationships. *Statistical Methodology*, **26**, 84–102.

KRISTENSEN, D. and RAHBEK, A. (2010). Likelihood-based inference for cointegration with nonlinear error-correction. *Journal of Econometrics*, **158** (1), 78–94.

— and — (2013). Testing and Inference in Nonlinear Cointegrating Vector Error Correction Models. *Econometric Theory*, **29** (6), 1238–1288.

KROLZIG, H. M. (1997). *Markov-Switching Vector Autoregressions: Modelling, Statistical Inference, and Application to Business Cycle Analysis*. Springer.

KUCK, K. and SCHWEIKERT, K. (2017). A Markov regime-switching model of crude oil market integration. *Journal of Commodity Markets*, **6**, 16–31.

KURTZ, T. G. and PROTTER, P. (1991). Weak limit theorem for stochastic integrals and stochastic differential equations. *The Annals of Probability*, **19** (3), 1035–1070.

LEE, J. and STRAZICICH, M. C. (2003). Minimum Lagrange Multiplier Unit Root Test with Two Structural Breaks. *The Review of Economics and Statistics*, **85** (4), 1082–1089.

LEHKONEN, H. (2015). Stock Market Integration and the Global Financial Crisis. *Review of Finance*, **19** (5), 2039–2094.

LI, H. and MADDALA, G. S. (1997). Bootstrapping cointegrating regressions. *Journal of Econometrics*, **80** (2), 297–318.

LIU, L., CHEN, C.-C. and WAN, J. (2013). Is world oil market "one great pool"?: An example from China's and international oil markets. *Economic Modelling*, **35**, 364–373.

LU, F.-B., HONG, Y.-M., WANG, S.-Y., LAI, K.-K. and LIU, J. (2014). Time-varying Granger causality tests for applications in global crude oil markets. *Energy Economics*, **42**, 289–298.

LUCEY, B. M. and TULLY, E. (2006). The evolving relationship between gold and silver 1978-2002: evidence from a dynamic cointegration analysis: a note. *Applied Financial Economics Letters*, **2** (1), 47–53.

LUMSDAINE, R. L. and PAPELL, D. H. (1997). Multiple Trend Breaks and the Unit-Root Hypothesis. *The Review of Economics and Statistics*, **79** (2), 212–218.

LÜTKEPOHL, H., SAIKKONEN, P. and TRENKLER, C. (2003). Comparison of tests for the cointegrating rank of a VAR process with a structural shift. *Journal of Econometrics*, **113** (2), 201–229.

—, — and TRENKLER, C. (2004). Testing for the Cointegration Rank of a VAR Process with Level Shift at Unknown Time. *Econometrica*, **72** (2), 647–662.

MACKINNON, J. G. (2010). Critical values for cointegration tests. *Queen's Economics Department Working Paper*, (1227), 1–17.

MAKI, D. (2013). Detecting cointegration relationships under nonlinear models: Monte Carlo analysis and some applications. *Empirical Economics*, **45** (1), 605–625.

— and KITASAKA, S.-I. (2015). Residual-based tests for cointegration with three-regime TAR adjustment. *Empirical Economics*, **48** (3), 1013–1054.

MANN, J. and SEPHTON, P. (2016). Global relationships across crude oil benchmarks. *Journal of Commodity Markets*, **2** (1), 1–5.

MANNING, D. N. (1991). Petrol prices, oil price rises and oil price falls: Some evidence for the UK since 1972. *Applied Economics*, **23** (9), 1535–1541.

MEYER, J. and CRAMON-TAUBADEL, S. (2004). Asymmetric Price Transmission: A Survey. *Journal of Agricultural Economics*, **55** (3), 581–611.

MEYLER, A. (2009). The pass through of oil prices into euro area consumer liquid fuel prices in an environment of high and volatile oil prices. *Energy Economics*, **31** (6), 867–881.

MOKKADEM, A. (1988). Mixing properties of ARMA processes. *Stochastic Processes and their Applications*, **29** (2), 309–315.

NBER (2017). Us business cycle expansions and contractions (table). *International Energy Statistics*. `http://www.nber.org/cycles.html`, [Online; accessed 08-February-2017].

NICHOLLS, D. F. and QUINN, B. G. (1982). *Random Coefficient Autoregressive Models: An Introduction*. Springer-Verlag.

NTIM, C. G., ENGLISH, J., NWACHUKWU, J. and WANG, Y. (2015). On the efficiency of the global gold markets. *International Review of Financial Analysis*, **41**, 218–236.

OSTERWALD-LENUM, M. (1992). A Note with Quantiles of the Asymptotic Distribution of the Maximum Likelihood Cointegration Rank Test Statistics. *Oxford Bulletin of Economics and Statistics*, **54** (3), 461–472.

PARK, J. Y. and PHILLIPS, P. C. B. (2001). Nonlinear regression with integrated time series. *Econometrica*, **69** (1), 117–161.

PERDIGUERO-GARCÍA, J. (2013). Symmetric or asymmetric oil prices? A meta-analysis approach. *Energy Policy*, **57**, 389–397.

PERRON, P. (1989). The Great Crash, the Oil Price Shock, and the Unit Root Hypothesis. *Econometrica*, **57** (6), 1361–1401.

PETRUCCELLI, J. D. and WOOLFORD, S. W. (1984). A Threshold AR(1) Model. *Journal of Applied Probability*, **21** (2), 270–286.

PHILLIPS, P. C. B. (1986). Understanding spurious regressions in econometrics. *Journal of Econometrics*, **33**, 311–340.

— (1987a). Asymptotic Expansions in Nonstationary Vector Autoregressions. *Econometric Theory*, **3** (1), 45–68.

— (1987b). Time Series Regression with a Unit Root. *Econometrica*, **55** (2), 277–301.

— and DURLAUF, S. N. (1986). Multiple Time Series Regression with Integrated Processes. *Review of Economic Studies*, **53** (4), 473–495.

— and OULIARIS, S. (1990). Asymptotic Properties of Residual Based Tests for Cointegration. *Econometrica*, **58** (1), 165–193.

PIERDZIOCH, C., RISSE, M. and ROHLOFF, S. (2014). On the efficiency of the gold market: Results of a real-time forecasting approach. *International Review of Financial Analysis*, **32**, 95–108.

PIPPENGER, M. K. and GOERING, G. E. (2000). Additional Results on the Power of Unit Root and Cointegration Tests under Threshold Processes. *Applied Economics Letters*, **7** (10), 641–644.

POLITIS, D. N. and ROMANO, J. P. (1994). The stationary bootstrap. *Journal of the American Statistical Association*, **89** (428), 1303–1313.

PORTNOY, S. and KOENKER, R. (1997). The gaussian hare and the laplacian tortoise: computability of squared-error versus absolute-error estimators. *Statistical Science*, **12** (4), 279–300.

PSARADAKIS, Z., SOLA, M. and SPAGNOLO, F. (2004). On Markov error-correction models, with an application to stock prices and dividends. *Journal of Applied Econometrics*, **19** (1), 69–88.

— and SPAGNOLO, N. (2006). Joint Determination of the State Dimension and Autoregressive Order for Models with Markov Regime Switching. *Journal of Time Series Analysis*, **27** (5), 753–766.

QUINTOS, C. E. (1997). Stability tests in error correction models. *Journal of Econometrics*, **82** (2), 289–315.

— and PHILLIPS, P. C. B. (1993). Parameter Constancy in Cointegrating Regressions. *Empirical Economics*, **18**, 675–706.

REBOREDO, J. C. (2011). How do crude oil prices co-move? A copula approach. *Energy Economics*, **33** (5), 948–955.

RODRIGUEZ, A. E. and WILLIAMS, M. D. (1993). Is the World Oil Market ' One Great Pool '? A test. *Energy Studies Review*, **5** (2), 121–130.

SAID, S. E. and DICKEY, D. A. (1984). Testing for unit roots in autoregressive-moving average models of unknown order. *Biometrika*, **71** (3), 599–607.

SAIKKONEN, P. (1991). Asymptotically efficient estimation of cointegration regressions. *Econometric Theory*, **7** (1), 1–21.

— (2005). Stability results for nonlinear error correction models. *Journal of Econometrics*, **127** (1), 69–81.

— (2008). Stability of Regime Switching Error Correction Models under Linear Cointegration. *Econometric Theory*, **24** (1), 294–318.

— and CHOI, I. (2004). Cointegrating Smooth Transition Regressions. *Econometric Theory*, **20** (2), 301–340.

— and LÜTKEPOHL, H. (2000). Testing for the Cointegrating Rank of a VAR Process with Structural Shifts. *Journal of Business & Economic Statistics*, **18** (4), 451–464.

SEO, B. (1998). Tests for Structural Change in Cointegrated Systems. *Econometric Theory*, **14** (2), 222–259.

SEO, M. H. (2006). Bootstrap testing for the null of no cointegration in a threshold vector error correction model. *Journal of Econometrics*, **134** (1), 129–150.

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

— (2011). Estimation of Nonlinear Error Correction Models. *Econometric Theory*, **27** (02), 201–234.

SHIN, Y. (1994). A residual-based test of the null of cointegration against the alternative of no cointegration. *Econometric Theory*, **10** (1), 91–115.

SMITH, G. (2002). Tests of the random walk hypothesis for London gold prices. *Applied Economics Letters*, **9** (10), 671–674.

STATISTA (2010). Kraftstoffverbrauch in Deutschland bis 2025. `https://de.statista.com/statistik/daten/studie/198562/umfrage/verbrauch-von-otto-und-dieselkraftstoffen-in-deutschland/`, [Online; accessed 25-September-2015].

STOCK, J. H. (1987). Asymptotic Properties of Least Squares Estimators of Cointegrating Vectors. *Econometrica*, **55** (5), 1035–1056.

TANG, K. and XIONG, W. (2012). Index Investment and the Financialization of Commodities. *Financial Analysts Journal*, **68** (6), 54–74.

TERASVIRTA, T. (1994). Specification, estimation, and evaluation of smooth transition autoregressive models. *Journal of the American Statistical Association*, **89** (425), 208–218.

TONG, H. (1983). *Threshold models in non-linear time series analysis*. Springer.

— (1990). *Non-linear time series: A dynamical system approach*. Oxford University Press.

TRENKLER, C., SAIKKONEN, P. and LÜTKEPOHL, H. (2007). Testing for the cointegrating rank of a VAR process with level shift and trend break. *Journal of Time Series Analysis*, **29** (2), 331–358.

TSAY, R. S. (1998). Testing and Modeling Multivariate Threshold Models. *Journal of the American Statistical Association*, **93** (443), 1188–1202.

US ENERGY INFORMATION AGENCY (2013). Refinery Capacity Report. `http://www.eia.gov/petroleum/refinerycapacity/archive/2013/refcap2013.cfm`, [Online; accessed 25-September-2015].

US ENERGY INFORMATION AGENCY (2015). Finished Petroleum Products. `http://www.eia.gov/dnav/pet/pet_cons_psup_dc_nus_mbbl_a.htm`, [Online; accessed 25-September-2015].

US ENERGY INFORMATION AGENCY (2016). Spot Prices. `http://www.eia.gov/dnav/pet/pet_pri_spt_s1_d.htm`, [Online; accessed 05-August-2016].

US ENERGY INFORMATION AGENCY (2017). Total petroleum and other liquids production (table). *International Energy Statistics*. `http://www.eia.gov/beta/international/data/browser/`, [Online; accessed 08-February-2017].

VAN DIJK, D., TERASVIRTA, T. and FRANSES, P. H. (2002). Smooth transition autoregressive models - A survey of recent developments. *Econometric Reviews*, **21** (1), 1–47.

WAGNER, M. (2010). Cointegration analysis with state space models. *Advances in Statistical Analysis*, **94** (3), 273–305.

WEINER, R. J. (1991). Is the World Oil Market "One Great Pool"? *The Energy Journal*, **12** (3), 95–108.

— (1993). The World Oil Market is not "One Great Pool": A Reply to Rodriguez and Williams. *Energy Studies Review*, **5** (3), 225–230.

WESTERLUND, J. and EDGERTON, D. L. (2007). New Improved Tests for Cointegration with Structural Breaks. *Journal of Time Series Analysis*, **28** (2), 188–224.

WILMOT, N. A. (2013). Cointegration in the Oil Market among Regional Blends. *International Journal of Energy Economics and Policy*, **3** (4), 424–433.

WLAZLOWSKI, S., HAGSTRÖMER, B. and GIULIETTI, M. (2011). Causality in crude oil prices. *Applied Economics*, **43** (24), 3337–3347.

XIAO, Z. (2009). Quantile cointegrating regression. *Journal of Econometrics*, **150** (2), 248–260.

— and PHILLIPS, P. C. B. (2002). A cusum test for cointegration using regression residuals. *Journal of Econometrics*, **108**, 43–61.

ZHANG, Y.-J. and ZHANG, L. (2015). Interpreting the crude oil price movements: Evidence from the Markov regime switching model. *Applied Energy*, **143**, 96–109.

ZHOU, S. and KUTAN, A. M. (2011). Is the evidence for PPP reliable? A sustainability examination of the stationarity of real exchange rates. *Journal of Banking and Finance*, **35** (9), 2479–2490.

ZIVOT, E. and ANDREWS, D. W. K. (1992). Further Evidence on the Great Crash, the Oil-Price Shock, and the Unit-Root Hypothesis. *Journal of Business & Economic Statistics*, **10** (3), 251–270.