**Justus-Liebig University Gießen**
**Department of Economics and Business Studies**

# Econometric Modelling of Energy & Financial Markets

**Doctoral Thesis**

Johannes Lips

Submission: 31st July 2019

Defended on: 6th November 2019

Supervisors

Prof. Dr. Peter Winker
Professor of Statistics and Econometrics
Justus-Liebig University Gießen

Prof. Dr. Georg Götz
Professor of Industrial Organization, Regulation and Antitrust
Justus-Liebig University Gießen

# Acknowledgements

This dissertation was written over the course of nearly five years while I was employed as research and teaching assistant at the Chair for Statistics and Econometrics (VWL VII) at the Department of Economics and Business Studies of the Justus-Liebig University Gießen.

I am thankful to my supervisor Prof. Dr. Peter Winker for giving me the opportunity to write my dissertation in the first place and providing me with the financial means necessary to live and to present my research throughout Europe. I also would like to thank him that he was always open to my research ideas and for his guidance and very constructive comments and recommendations. I am also grateful that Prof. Dr. Georg Götz was willing to serve as my second advisor and for his constructive comments.

This doctoral thesis also includes joint papers with other co-authors, namely with Christina Bannier, Corinna Ewelt-Knauer, Christoph Funk, Karol Kempa and Peter Winker. I would like to thank all of them for the very constructive and rewarding cooperation. Furthermore, I would like to thank all my colleagues at Gießen University for their support and the Frankfurt School of Finance & Management for the hospitality. I am grateful for all new friendships, critical discussions, and conversations.

Last but not least, I would like to thank my family, my parents and grandparents, who always gave me the freedom to learn, encouraged me to broaden my horizon and always believed in me, my brother Christian and his girlfriend Ellen, who helped to make the most challenging events in life bearable. Most of all, I would like to thank Renate for her support, faith and the motivation she provided me with during the whole time of writing this thesis.

# Chapter 1 Introduction

This dissertation consists of two parts with different themes, the first part is made up of three articles in the broad category of empirical energy economics. The first paper in this category analyses the impact of changes in the prices of fossil fuels on the electricity prices. The following two articles are closely related and empirically analyse the North American oil industry to see if the financial decisions and conditions of the firms and the oil industry are affected by oil price changes. The final article provides a literature review prepared for the handbook "Corruption and Fraud in Financial Markets: Malpractice, Misconduct and Manipulation." with an accompanying data analysis on the detection of financial fraud and manipulations using Benford's law.

The paper "*Do They Still Matter? – Impact of Fossil Fuels on Electricity Prices in the Light of Increased Renewable Generation*" (Chapter 2) is analysing how the German energy sector and especially its electricity market was affected by a major energy transition, the so called "Energiewende". This transition led to an increase of electricity production from renewable sources and thereby affected the whole electricity market. The aim of this analysis is to assess if there still exists a relationship between fossil fuel and electricity prices. Due to possible structural breaks in the time series, a minimum Lagrange Multiplier (LM) stationarity test is applied, which endogenously determines possible structural breaks. Subsequently, a bootstrap approach is used to estimate confidence intervals for the test statistic and the possible break dates. The stability of the cointegration vector is assessed with a test and the results indicate that the cointegration relationship is not stable over time. To incorporate these findings, the employed cointegration analysis allows structural breaks in the deterministic part of the cointegration relation. These results support the assumption that the energy transition affected the relationship between fossil fuels and electricity prices, although there still exists a relatively strong cointegration relation between fossil fuel and electricity prices in the long run. Therefore, the paper may provide lessons for countries, which are only beginning a similar transition away from fossil fuels to renewable energy sources.

The second paper "*Debt and the Oil Industry – Analysis on the Firm and Production Level*" (Chapter 3) investigates the relationship between debt and production decisions of firms active in the exploration and production of oil and gas in the US. Over the last couple of years, the development and application of innovative extraction methods led to a considerable increase in US oil production. In connection with these technological changes, another important economic development in the oil industry was largely debt-driven investment. The extensive use of debt was fostered by the macroeconomic environment in the aftermath of the financial crisis. Additionally, the rising prices in the commodities markets until mid-2014 led to higher asset valuation and a virtuous circle. This increased investment activity, especially in the US, raised the production capacity and, as a consequence, also the production of crude oil. This trend continued in spite of the oil price decline in 2014, although production reductions would be more plausible. The main research question of this paper is whether debt and leverage affect production decisions of firms. To address this question, I use a novel panel VAR approach and a dataset combining financial data on publicly listed firms and their production data on well level. This article also includes an appendix with additional analyses, which will be available online after the article is published.

The paper "*Oil Price Shocks and Cost of Debt – Evidence from Oil Firms*" (Chapter 4), co-authored with Christoph Funk and Karol Kempa, analyses the relationship between (adverse) oil price shocks and the cost of debt of US oil firms. In particular, we analyse how oil firms, which we differentiate along the oil industry's value chain, respond to oil-price shocks and how these shocks affect their borrowing decision and creditworthiness perceived by banks and capital markets. For US oil firms we collect (i) data on individual syndicated loans taken and (ii) bonds issued. We combine this data with information from these firms' corporate financial statements. Thus, we can analyse how a firm's (financial) characteristics, e.g. firm size, profitability and leverage / indebtedness, affect the credit spread of loans and bonds, i.e. the cost of debt. In addition to these firm characteristics, we consider the oil price and, in particular, oil price shocks – considering both the 2008 and 2014 oil price decline – and their effects on the firms' costs of debt. Overall, we find that the credit market tightens in the immediate aftermath of both oil price shocks, i.e. the amount of loans issued decreases, while their interest rates increases. This effect is confirmed by the firm-level analysis. Even after controlling for loan/bond and firm characteristics, oil prices, in particular oil price shocks, have an effect on a firm's cost of debt.

Finally, the handbook article "*Benford's law and its application to detecting financial fraud and manipulation.*" (Chapter 5), co-authored with Christina Bannier, Corinna Ewelt-Knauer and Peter Winker, introduces Benford's law and its history. It discusses the advantages and limitations of Benford's law with a special focus on the usage of Benford's law to detect financial fraud and manipulation. The article provides an overview of the literature on the detection of financial fraud and manipulations, where Benford's law was applied in a wide variety of cases and applications. We then empirically apply Benford's law to assess the extent of manipulation, which could be observed in the formation of the London Interbank Offered Rate (LIBOR). Interestingly, it is possible to see that suspicious manipulations were reduced by reforms introduced after a scandal surfaced during the financial crisis in 2008.

All four papers are separate works and presented as such. As the first paper is already published it is included in the layout of the journal. The second paper is accepted for publication, but not yet published, thus the accepted version is included in this thesis. The third paper currently is a working paper, not yet published, and the last article is going to be published as a Chapter in the handbook "Corruption and Fraud in Financial Markets: Malpractice, Misconduct and Manipulation." edited by Carol Alexander and Douglas Cumming.

# Chapter 2 Do They Still Matter? – Impact of Fossil Fuels on Electricity Prices in the Light of Increased Renewable Generation

**Reference for this Paper:**

Lips, J. (2017): "Do They Still Matter? – Impact of Fossil Fuels on Electricity Prices in the Light of Increased Renewable Generation". In: *Journal of Time Series Econometrics* 9 (2), pp. 1–30. DOI: 10.1515/jtse-2016-0018.

**Conferences with Review Process:**

- 16th IWH-CIREQ Macroeconometric Workshop "Challenges for Forecasting: Structural Breaks, Revisions and Measurement", Halle/Saale, December 2015 (Poster presentation).
- Annual Conference of the Verein für Socialpolitik, Augsburg, September 2016 (Presentation).

Johannes Lips*

# Do They Still Matter? – Impact of Fossil Fuels on Electricity Prices in the Light of Increased Renewable Generation

DOI 10.1515/jtse-2016-0018

**Abstract:** During the last years, the German energy sector and especially its electricity market was affected by a major energy transition, the so called „Energiewende“. This transition led to an increase of electricity production from renewable sources and thereby affected the whole electricity market. Therefore, it provides lessons for countries, which are only beginning a similar transition away from fossil fuels to renewable energy sources. The aim of this analysis is to assess if there still exists a relationship between fossil fuel and electricity prices. Due to possible structural breaks in the time series a minimum Lagrange Multiplier (LM) stationarity test is applied, which endogenously determines possible structural breaks. Subsequently a bootstrap approach is used to estimate confidence intervals (C.I.s) for the test statistic and the possible break dates. Furthermore, the stability of the cointegration vector is assessed with the test by Hansen and Johansen (1999. "Some Tests for Parameter Constancy in Cointegrated VAR-Models." *The Econometrics Journal* 2 (2):306–333.). The results indicate that the cointegration relationship is not stable over time. To incorporate these findings, the cointegration analysis is based on Johansen Mosconi, and Nielsen (2000. "Cointegration Analysis in the Presence of Structural Breaks in the Deterministic Trend." *Econometrics Journal* 3 (2):216–249), which allows structural breaks in the deterministic part of the cointegration relation. These results support the assumption that the energy transition affected the relationship between fossil fuels and electricity prices, although there still exists a relatively strong cointegration relation between fossil fuel and electricity prices in the long run.

**Keywords:** structural breaks, bootstrap, stationarity test, cointegration, energy economics, german energy transition
**JEL Classification:** C58, G14, L94, Q41

***Corresponding author: Johannes Lips,** Department of Economics, Justus-Liebig-University Giessen, Chair of Statistics and Econometrics Licher Strasse 64, 35394, Giessen, Hesse, Germany, E-mail: johannes.lips@wirtschaft.uni-giessen.de

# 1 Introduction

Over the last couple of years the German electricity market was mainly affected by the energy transition, which initially started in 2000 with the first implementation of the renewable energy act.[1] One main result of the energy transition was the increase of installed generation capacity from renewable energy sources, especially photovoltaic and wind.

As can be seen in Table 0 the share of renewable capacity nearly doubled from 26.6 % in 2007 over the course of seven years to 49.6 % in 2014. Partly, this increase can be attributed to the events triggered by the nuclear incident at the Fukushima Daiichi plant in Japan and the subsequent political reaction in Germany, which resulted in the immediate shutdown of around 60 % of the existing nuclear generation capacity in March 2011. Nevertheless, the renewable generation capacity almost tripled in absolute terms from 36 GW in 2007 to 91 GW in 2014.[2]

This increase added a lot of generation capacity to the German merit-order,[3] with no or very low marginal costs, influencing the electricity price via the merit-order effect. This merit-order effect especially affects the hours of high demand during the peak hours,[4] when also the production potential from photovoltaic is highest. (Tveten et al. 2013) Due to additional changes to the overall market design, the merit-order effect became much more important for the price determination on the spot market of the European Power Exchange (EPEX Spot). Beginning from January 2010, all of the electricity produced from renewables had to be sold over a public exchange. This led to a strong increase of traded volumes from 2009 to 2010 by 45 % and the volume kept increasing although at a much slower pace.

This increased supply of electricity with no or very low marginal costs, in connection with the price determination algorithm of the exchange, led to a marked decrease of the yearly average price for electricity during peak hours from 55 Eur/MWh in 2010 to only 41 Eur/MWh in 2014. For a comprehensive analysis on how the renewable generation capacity affects the intra-day market and the price formation on the exchange, please see Haas et al. (2013).

**1** Initial implementation of the renewable energy act (‚Erneuerbare Energien Gesetz' (EEG)) and successive amendments in 2004, 2008, 2012 and 2014.

**2** Included in this category is generation capacity from Hydro Power (2007: 5.1, 2014: 5.6), Biomass (2007: 4.7, 2017: 8.9), Wind (2007: 22.2, 2014: 38.3) and Photovoltaic (2007: 4.2, 2014: 38.2) all values in GW and taken from Burger (2016).

**3** The merit-order is the available electricity generation capacity ranked, in ascending order, according to its short-run marginal costs of production and basically represents the supply curve in the electricity market.

**4** Hours of peak demand are defined as the hours from 8:00 to 20:00.

The main consequence of the increased renewable generation capacity is less demand for electricity from fossil-fuel power stations with higher marginal costs during times of high demand. This is due to the fact that renewable energy sources, especially photovoltaic, are often able to satisfy a substantial part of demand during peak hours. Therefore, especially gas-fired power stations are crowded out of the market, because the residual demand can be satisfied with generation capacity having lower marginal costs. This might even result in a permanent shutdown of some, because the continued operation of these power plants becomes economically unviable.[5] (Haas et al. 2013, 39–41)

This becomes especially evident, when looking at the development of the full-load hours for the different energy sources over the horizon of this analysis. Full-load hours[6] are a hypothetical measurement to assess the utilization of available generation capacity. It can be interpreted as the number of hours all available generation capacity would have had to run at full utilization to generate the realized amount of electricity.

The yearly growth rate in percent of the full-load hours is displayed in Table 1.[7] It can be seen that the German moratorium on nuclear power in 2011

**Table 1:** Development of full-load hours by energy source (growth rate in %) (absolute numbers are presented in Table 7 (BDEW 2016; Burger 2016)).

| | 2007 | 2008 | 2009 | 2010 | 2011 | 2012 | 2013 | 2014 |
|---|---|---|---|---|---|---|---|---|
| Δ **Nuclear** | −18.0 | 4.1 | −9.3 | 4.1 | 31.5 | −8.2 | −2.2 | −0.4 |
| Δ **Lignite** | −0.4 | −2.6 | −3.3 | −1.0 | 16.2 | 0.3 | 1.0 | −3.5 |
| Δ **Hard Coal** | 1.0 | −14.4 | −12.7 | 4.3 | 12.2 | 5.6 | 5.9 | −8.2 |
| Δ **Gas** | 3.4 | 6.6 | −11.1 | 7.1 | −16.5 | −12.6 | −16.2 | −12.7 |
| Δ **Oil** | −8.1 | −3.3 | 7.6 | −26.8 | 13.6 | 7.1 | −5.1 | −27.2 |
| Δ **Renewables** | 12.0 | −5.3 | −14.3 | −7.9 | 0.8 | 0.8 | −2.3 | −0.8 |
| Δ Biomass | 21.2 | 4.8 | 0.0 | 2.9 | 3.2 | 13.3 | −6.5 | −0.0 |
| Δ Hydro | 6.4 | −3.6 | −10.7 | 8.6 | −21.7 | 23.0 | 4.1 | −15.2 |
| Δ Solar | −3.7 | −2.1 | −14.8 | 4.9 | 14.4 | 2.5 | 8.7 | 11.0 |
| Δ Wind | 18.2 | −5.0 | −12.4 | −7.5 | 19.6 | −3.3 | −7.7 | −1.7 |

**5** Press release by German utility e.on stating the plan to shut down two gas-fired power stations e.on Press Releases 2015, http://www.eon.com/en/media/news/press-releases/2015/3/30/no-economic-prospects-owners-of-the-irsching-4-and-5-gas-fired-power-stations-announce-their-closure.html, last accessed 31/01/1017 (30/03/2015).

**6** Defined as the total electricity produced in GWh divided by the total available generation capacity in GW.

**7** Absolute values for the full-load hours and the two variables generation capacity and actual electricity generation are presented in Tables 5, 6 and 7 in the appendix.

had big implications for the utilization of all other fossil fuel generation capacity. The rather small changes in the utilization of renewable energy sources can be attributed by the fact, that for renewables the generation capacity and the actual production grew at nearly the same pace. In case of gas-fired power plants the utilization yields a whole different picture and although the generation capacity even grew slightly the full-load hours decreased from 3,542 in 2007 to only 2,036 in 2014. This indicates that gas-fired power plants are heavily affected by the merit-order effect and hence are often crowded out of the market. Therefore, the relationship between natural gas and electricity prices is supposed to have weakened over the sample period. The aim of this paper is to empirically analyze the relationship between fossil fuels, primarily used in generation, and the wholesale price for electricity in Germany. In particular it attempts to address various questions, whereas the fundamental question is to determine if there exists any relation between the electricity prices and the fossil fuel costs, and if so, how the major transitions in the German energy sector might have affected this relationship. A detailed analysis of these issues might shed some light on how and if the wholesale market for electricity is still driven by fundamentals or if their impact became less relevant over the recent years.

There exists a wide array of empirical literature analyzing electricity prices. Whereas one strand of literature focuses on the interdependencies of the different energy commodities and the fundamental modelling of electricity prices, there exists another strand of literature, which focuses solely on modelling the electricity market. This latter research area tries to model the stochastic properties of the electricity price by incorporating, amongst other things, volatility clustering, seasonality and extreme values. Weron (2006) offers a comprehensive overview on this strand of literature. The shortcoming of these studies, however, is that they are not suited to analyze the relationship between input fuel prices and electricity prices. The first strand of literature, which focuses on the analysis of the relationship between electricity and energy commodity prices, can be differentiated along various dimensions. Most of the studies differ regarding the markets and commodities, the time horizon and empirical methodology employed. Hence, it is not possible to find a generally valid conclusion, but most of them hint at similar concluding results.

Another broad overview of the various modelling approaches in the literature, is provided by the review of electricity price forecasting in Weron (2014). This article assesses a broad variety of modelling approaches and evaluates each approach regarding its forecasting abilities. Besides classical econometric statistical approaches the authors also include agent-based computational models and computational intelligence models, using artificial intelligence and neural networks besides others, in their assessment and thereby also provide hints at future developments in this research area.

Mjelde and Bessler (2009) focus more on the short-run dynamics and include four of the major electricity generation fuel sources, namely natural gas, uranium, hard coal and crude oil. The authors use a VECM framework to assess the dynamic interactions between the prices of these commodities and U.S. electricity spot prices between 2001 and 2008. The results of their analysis show that fossil fuels are weakly exogenous in the long run and electricity together with uranium prices react to re-establish the long-run equilibrium. Mohammadi (2009), in contrast, is more interested in the long-run relationship and uses annual price time series for electricity and the fossil fuels – natural gas, hard coal and crude oil – from 1960 to 2007. It turns out that in his application of a VECM, the impact of fossil fuels in the long run is rather mute, although in the short-run electricity prices are affected by price movements in natural gas and hard coal markets.

Apart from energy markets in the U.S., several studies also analyzed liberalized markets in Europe. Fezzi and Bunn (2009) are mainly interested in the impact the European carbon trading scheme has on electricity and natural gas prices in the UK. They also use a VECM framework and conclude that, over a relatively short sample period from April 2005 to June 2006, electricity prices are driven both by carbon and natural gas prices. In contrast, Bosco et al. (2010) focuses on the question if energy markets for electricity and natural gas are integrated across nine European countries, although the markets for the Nordic countries[8] are pooled in the Nordpool market area. The results indicate that only the electricity markets of central Europe[9] are integrated, while the Spanish and the Nordpool market area seem to not share a common trend. Additionally, the authors report strong evidence of a long-run relationship between electricity and gas prices, which cannot be observed for oil prices.

Finally, Ferkingstad, Løland, and Wilhelmsen (2011) analyze the flow of dynamic price information for the Nordpool market area and Germany and also employ a VECM, which incorporates weekly prices for electricity, natural gas, hard coal and oil as endogenous variables. Their results indicate that natural gas has a stronger impact on electricity prices than hard coal and oil. An interesting result is the observation of Fell (2010), that the effect of input fuel prices varies with the demand level. In his VECM, the impact of carbon price is stronger in off-peak hours than in peak hours. Thoenes (2011) analyses the cointegration relationship between electricity, natural gas and carbon prices in

**8** Norway, Sweden, Finland and Denmark.
**9** Austria, France, Germany and the Netherlands.

Germany between 2008 and 2010 and the results indicate that electricity prices adapt to fossil fuel price changes in a long-term cointegration relationship.

The approach in this paper mostly relates to the fundamental modelling strand of literature presented above and also applies a VECM framework to analyze the question, if prices of fossil fuels still play a part in price determination of electricity markets. This analysis adds to the literature by using an econometric model, which incorporates many characteristics of electricity markets and especially takes fundamental structural changes into account.

The remainder of the paper is structured as follows. The next section describes the data used for the analysis in detail. In Section 3, a stationarity test, which allows the possibility to endogenously determine possible structural breaks, is presented and applied to the endogenous variables of the VECM framework. Then Section 4 describes this framework, which was initially developed by Johansen, Mosconi, and Nielsen (2000), in more detail. The application of a framework, which allows the possibility to allow structural breaks in the cointegration relation, is also indicated by the additionally applied test by Hansen and Johansen (1999). Afterwards, the results, obtained in the cointegration analysis, are presented and critically assessed. Finally the last section concludes the paper and presents possible routes for future research.

## 2 Data

This analysis is based on data compiled from various sources. Information regarding commodity and electricity prices are taken from Reuters Datastream. The electricity price under consideration, the Phelix Peak index, is the average day-ahead price for the delivery area of Germany and Austria during peak times and is determined on the EPEX Spot. The Phelix Peak index represents the average price for the whole period of higher load from 8:00 to 20:00 and is denominated in EUR/MWh. The primary energy sources considered in this analysis are natural gas and hard coal, because those commodities are both used as input for electricity generation and they are traded on exchanges. The natural gas price used is the European Gas Index (EGIX) for both German market areas in EUR/MWh. Hard coal for delivery in Amsterdam, Rotterdam or Antwerp (ARA) is the product primarily traded on the Intercontinental Exchange (ICE) for imports into Northwestern Europe and is therefore included in the analysis. To make the results comparable, all prices are converted into EUR/MWh (Figure 1).

The time period covered in this analysis includes all working days from September 28, 2007 to January 15, 2015, which results in a total of 1,905 observations for all variables included in this analysis, omitting observations

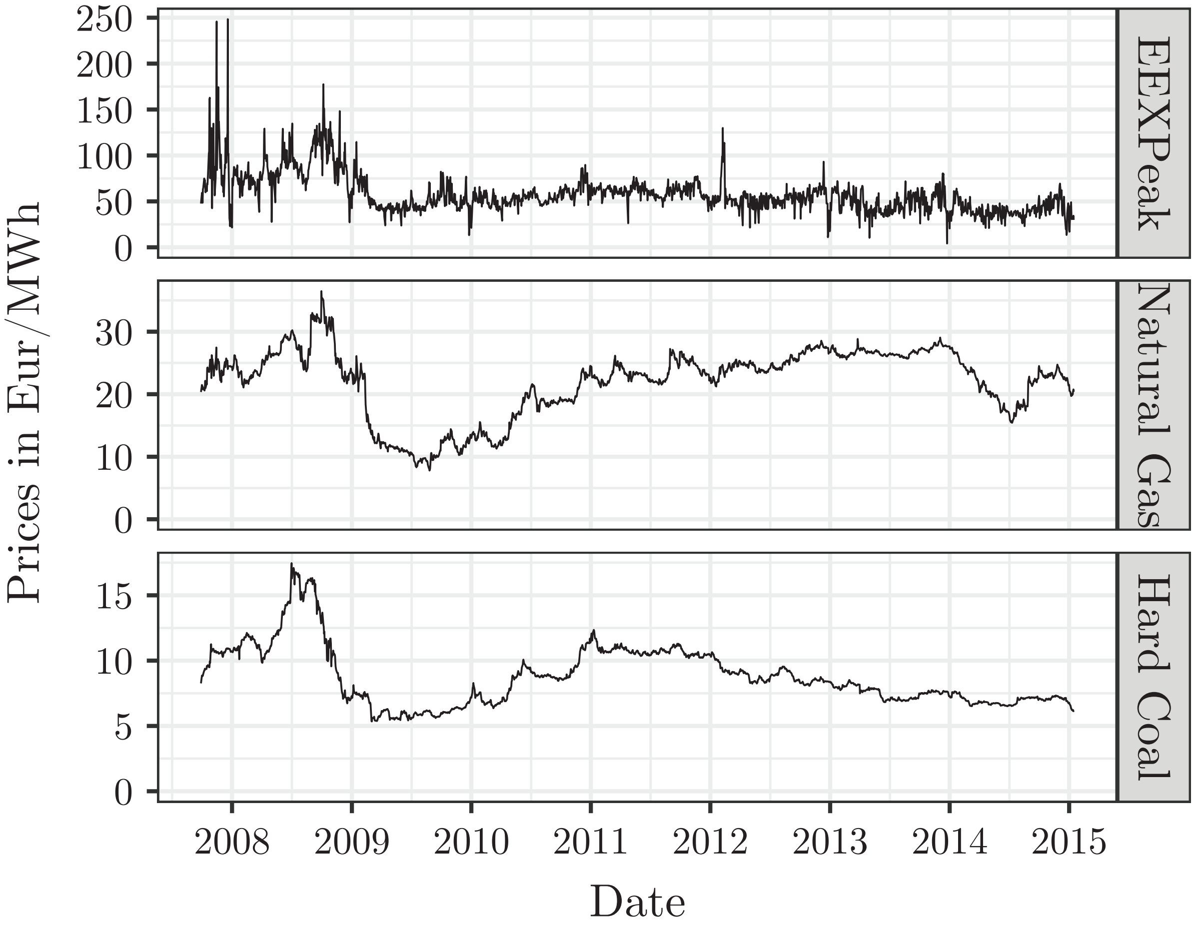



**Figure 1:** Graphical display endogenous variables covering the period from September 28, 2007 to January 15, 2015.

for weekends. Since the carbon price of the second phase of the European Emission Trading Scheme (EU-ETS) only begins to become larger than zero from mid January 2009, it is excluded. Since weather and especially temperature is one of the main exogenous factors affecting the demand for electricity and hence the price, the variables heating degree day (HDD) and cooling degree day (CDD) are included (Halvorsen 1975; Quayle and Diaz 1980). These variables are able to address the possible non-linear effect of temperature on demand by differentiating between the energy needed to heat and to cool buildings to keep the inside at a constant temperature of 18 °C throughout the year. The CDD and HDD variables are calculated, based on the average daily temperature measured across Germany by Germany's National Weather Forecast Service (Deutscher Wetterdienst (DWD) 2015). A further variable possibly affecting both electricity demand and supply and is closely related to weather and seasonality is the amount of daylight during a day. This not only affects the demand for lighting during the day, but also the potential production from photovoltaic. Therefore, the average sunshine duration across Germany, also calculated from data by the

Deutscher Wetterdienst (DWD) (2015), is incorporated into the model as an exogenous variable.

An additional indirect effect of weather, which might influence the supply of electricity, is the river temperature. Due to regulatory requirements, the power plants have to curtail their generation if the water temperature exceeds a threshold of 23 °C. Therefore, a river temperature index is calculated, based on the daily temperatures measured at 33 stations along eight major German rivers.[10] The calculation of the river temperature index is very similar to the cooling degree day (CDD)/heating degree day (HDD) variables. If the temperature for any station used is above the threshold, the absolute difference to 23 °C is calculated and weighted with the share of stations, observing temperatures above threshold on the respective day.

The seasonality of electricity prices is not only driven by weather effects, but also appears to be based on calendar effects. Therefore, dummy variables which capture the intra-week structure and all public holidays, which are observed across the whole of Germany and take place on a normal working day, are included in the analysis.

# 3 Preliminary Tests – Stationarity

## 3.1 Minimum Lagrange Multiplier (LM) Test with Structural Breaks

Since the influential paper by Perron (1989), it became clear that one has to explicitly account for possible structural breaks, when testing for stationarity or a unit root – the possibility of rejecting the unit root null hypothesis decreases when the stationary alternative is true and a structural break is not considered. In the initial implementation, Perron (1989) modified the augmented Dickey-Fuller (ADF) test and included a dummy variable to account for the known or exogenous structural break. Further extensions of this procedure allowed for an unknown breakpoint to be determined endogenously in the data. One of those procedures is the test proposed by Zivot and Andrews (1992), which chooses the breakpoint according to the minimum value of the *t*-statistic testing the null hypothesis of a unit root. Since the power of a unit root test decreases when ignoring one break, not considering a second break also results in a loss of power. Therefore, Lumsdaine and Papell (1997) extended the initial test by Zivot

**10** Included rivers are: Danube, Elbe, Ems, Main, Moselle, Neckar, Rhine and Saar.

and Andrews (1992) and allowed for the possibility of two structural breaks. One major issue in connection with these endogenous break tests is the assumption of no structural break under the unit root null hypothesis. Thus, the alternative hypothesis is that there are structural breaks in the series, which also includes the possibility of a structural break under the unit root null hypothesis. Therefore, a rejection of the null in such tests does not necessarily imply a rejection of the unit root hypothesis per se.

The minimum LM stationarity test used here was first proposed by Lee and Strazicich (2003). It has some advantages over the more commonly used tests for stationarity or a unit root. Most notably is the possibility to allow a unit root with breaks, which considerably lowers the problem of „spurious rejections". (Lee and Strazicich 2003, 1–2).

The data generating process (DGP) is based on the first-order autoregressive model described in eq. [1], where the variable $Z_t$ contains exogenous variables. The assumed DGP for the time series variable $X_t$ is modeled as a first-order autoregressive process with the error term $\varepsilon_t$ representing a white noise process.

$$y_t = \delta^T Z_t + X_t, \text{with}\, X_t = \beta X_{t-1} + \varepsilon_t, \tag{1}$$

Note that in this parametrization of the DGP, the unit root null hypothesis is represented by $\beta = 1$ in eq. [1]. Under the assumption of this hypothesis the time series $X_t$ would follow a random-walk and hence be non-stationary. Another advantage of this formulation is that structural breaks are included, both under the null and also under the alternative hypothesis, when $\beta < 1$ and hence $X_t$ is following a stationary AR(1) process. The nature of the exogenous variables included in $Z_t$ depend on both, the assumed model for the DGP and the structural break. The exogenous variables included in $Z_t$ depend on both, the assumed model for the DGP and the structural break. For the case of breaks in the intercept, the model for the DGP corresponds to model A defined in Perron (1989, 4–6) and is often referred to as the „crash" model. To appropriately incorporate such changes of the intercept into the model, $Z_t$ can be described as $Z_t = [1, t, D_{1t}, D_{2t}]^T$, where $D_{it}$ = 1 for $t \geq T_{Bj} + 1$, $\{j = 1, 2\}$, and $D_{it}$ = 0 otherwise. The date of the break is denoted by $T_{Bj}$. The second model considered in this analysis is model C from Perron (1989), which not only allows for breaks in the intercept but also in the trend of the DGP and is often referred to as the „break" model.[11] In order to account for possible changes in the trend, an additional variable $DT_{it}$ is included in $Z_t$, with $DT_{it} = t - T_{Bj}$ for $t \geq T_{Bj} + 1$, $\{j = 1, 2\}$, and

**11** The third case described by Perron (1989), Model B allows a break in trend and is called the „changing growth" model by Perron (1989, 5), but following the reasoning of Lee and Strazicich (2013, 3) it is not considered here.

$DT_{it} = 0$ otherwise. Note that the unit root null hypothesis in these models is represented by the coefficient $\beta$ in eq. [1] being equal to one. The advantage of this formulation is that structural breaks are not only included under the null, but also under the alternative hypothesis $\beta < 1$. The regression, which determines the lm stationarity test statistic can be estimated with the following equation:

$$\Delta y_t = \delta^T \Delta Z_t + \phi \tilde{S}_{t-1} + u_t, \qquad [2]$$

where $\tilde{S}_t = y_t - \tilde{\psi}_x - Z_t \tilde{\delta}$, $t = 2, \ldots, T$; $\tilde{\delta}$ are the coefficients of the regression of $\Delta y_t$ on $\Delta Z_t$ and $\tilde{\psi}_x$ is given by $y_1 - Z_1 \tilde{\delta}$, which is the restricted mle of $\psi_x (\equiv \psi + X_0)$. In the original paper by Schmidt and Phillips (1992, 259) $\psi_x$ represents the level or constant of the DGP, which in our notation is implicitly included in the vector of coefficients $\delta^T$, depending on the assumed exogenous variables in $Z_t$.

According to eq. [2], the unit root null hypothesis is expressed by $\phi = 0$ and the corresponding LM test statistics are then defined as

$$\begin{aligned} \tilde{\rho} &= T\tilde{\phi}, \\ \tilde{\tau} &= t-\text{statistic testing the null hypothesis } \phi = 0. \end{aligned} \qquad [3]$$

To account for possible autocorrelation in the residuals, augmented terms $\Delta \tilde{S}_{t-j}$, $\{j = 1, \ldots, k\}$, can be included in the test eq. [2] (Lee and Strazicich 2013, 4). In accordance with Ng and Perron (1995, 271–272), a general to specific approach is used to determine the optimal number of $k$ augmented terms. In this approach, the model initially is defined in the most general form with $k_{max}$ lags of the augmented terms. In each step of an iterative procedure, the significance of the augmented term with the highest lag-order is checked. If significant then $k = k_{max}$; otherwise the non-significant augmented term is removed and the procedure is repeated for $k_{max} - 1$ until the coefficient of the lagged augmented term becomes significant.

The location of possible break points $\lambda_j = \frac{T_{Bj}}{T}$, $\{j = 1, 2\}$ is determined by employing a grid search algorithm, minimising the unit root test $t$-statistic across all possible break locations and combinations in case of more than one break.

$$LM_\tau = \operatorname*{Inf}_{\lambda} \tilde{\tau}(\lambda) \qquad [4]$$

Due to possible endpoint problems, which are common in endogenous structural break tests, the grid search algorithm is only applied to a subsample of the total observations $\kappa$, and per default 10 % of the observations are left out at each end of the time series. An additional requirement is that the second break point can only occur at least two periods after the first break in the „crash“ model and for the „break“ model that gap needs to be at least three periods.

The critical values for the unit root null hypothesis are derived by Lee and Strazicich (2003) and depend, for the case of a break in intercept and trend, also on the location of the break $\lambda_j$. For more details on the derivation of the test procedure and its proofs, please be referred to the papers by Lee and Strazicich (2003, 2013) and Schmidt and Phillips (1992). The relevant critical values of the $LM_\tau$ test statistic for testing for a unit root hypothesis are provided in Appendix.

## 3.2 Bootstrap Procedure and Results of Minimum Lagrange Multiplier (LM) Test with Structural Breaks

In addition to the standard implementation of the minimum LM test, the bootstrap approach by Chou (2007) is employed to obtain critical values for the test statistic. Additionally, it also allows a detailed analysis of the distributional properties of the test statistic and the possible break points. The first step of the bootstrap procedure is to apply the minimum LM test on the time series, based on eq. [2], to determine the minimal test statistic and the two possible break dates. Based on these results, the test regression's coefficients are used to calculate restricted residuals, which do incorporate the possible structural breaks under the null hypothesis. These restricted residuals are then resampled and used, together with the test regression's coefficients, to construct a pseudo sample $y_t^*$. This resampling procedure is then repeated 1,000 times and the minimum lm test is applied to each of the new pseudo samples. For each run, the results are stored and it is then possible to analyze the distributional characteristics and calculate the 95 % percentile bootstrap C.I.s for the two possible break occurrences. The results of this bootstrap procedure, separately applied to each of the three time series variables, are shown in Table 2. Besides

**Table 2:** Results of bootstrap procedure of minimum LM test with the possibility of two structural breaks in trend and intercept.

| | **Test stat.**[a] | $T_{B1}$[b] | $T_{B2}$[b] |
|---|---|---|---|
| 7 EEXPeak | −11.28 | 19/12/2008 | 08/10/2010 |
| | [−23.70] | [11/07/2008, 07/02/2012] | [24/11/2008, 26/02/2014] |
| 7 Natural Gas | −3.49 | 13/02/2009 | 07/12/2010 |
| | [−20.06] | [04/08/2008, 16/03/2010] | [10/03/2010, 16/11/2012] |
| 7 Hard Coal | −3.61 | 30/10/2008 | 07/12/2010 |
| | [−20.20] | [14/07/2008, 04/04/2011] | [14/04/2010, 02/10/2013] |

99 % bootstrapped one-sided lower confidence limits in squared brackets
95 % bootstrapped two-sided C.I.s in squared brackets.

the point estimate for the two break dates $T_{B1}$ and $T_{B2}$ two-sided C.I.s are provided. Especially the C.I.s indicate that the dates for the structural breaks don't seem to be statistically significantly different, since for both breaks the C.I. s overlap quite substantially. Therefore, it is possible to assume that the date for each of the three time series the structural breaks happen on the same dates. Based on these findings the break dates are set rather ad-hoc and the first break is set to have occurred on the break found for the electricity price on December 19, 2008, whereas the second break is assumed to have happened on December 7, 2010, since this break is found independently in both time series of the natural gas and hard coal prices. These results, shown in Table 2, indicate, that the break dates don't seem to be statistically significantly different and therefore it is possible to assume that the two structural breaks do occur on the same dates for all time series. In appendix B histograms with the relative frequency of breaks are provided, which also indicate that it's not possible to assume different break dates for all three time series. Therefore, it is assumed that the two structural breaks occur for all time series on December 19, 2008 and December 7, 2010 respectively. This decision is rather ad-hoc and the first break is set to occur on the found break for the electricity price. The second break is assumed to happen on the December 7, 2010, since this break is found independently in both time series of the natural gas and hard coal prices.

In case of the test statistic, one-sided 99 % C.I.s are calculated and shown in Table 2. Based on these results it is not possible to reject the unit root hypothesis for all three time series, because the lower confidence limits are not exceeded by any of the test statistic (Figure 2).

# 4 Cointegration Analysis

## 4.1 Methodology

In this part of the analysis, a possible cointegration relationship between the variables is investigated and especially the possible structural breaks indicated by the minimum LM stationarity test are explicitly considered in more detail. Due to the fundamental changes the market for electricity in Germany underwent during the sample period, the possible cointegration relationship might have changed as well. Therefore, before conducting the cointegration analysis in detail, the test by Hansen and Johansen (1999) is employed to analyze, in a VECM framework, if the assumed cointegration relationship is stable over time. This LM type of test makes it possible to identify structural breaks in a multivariate

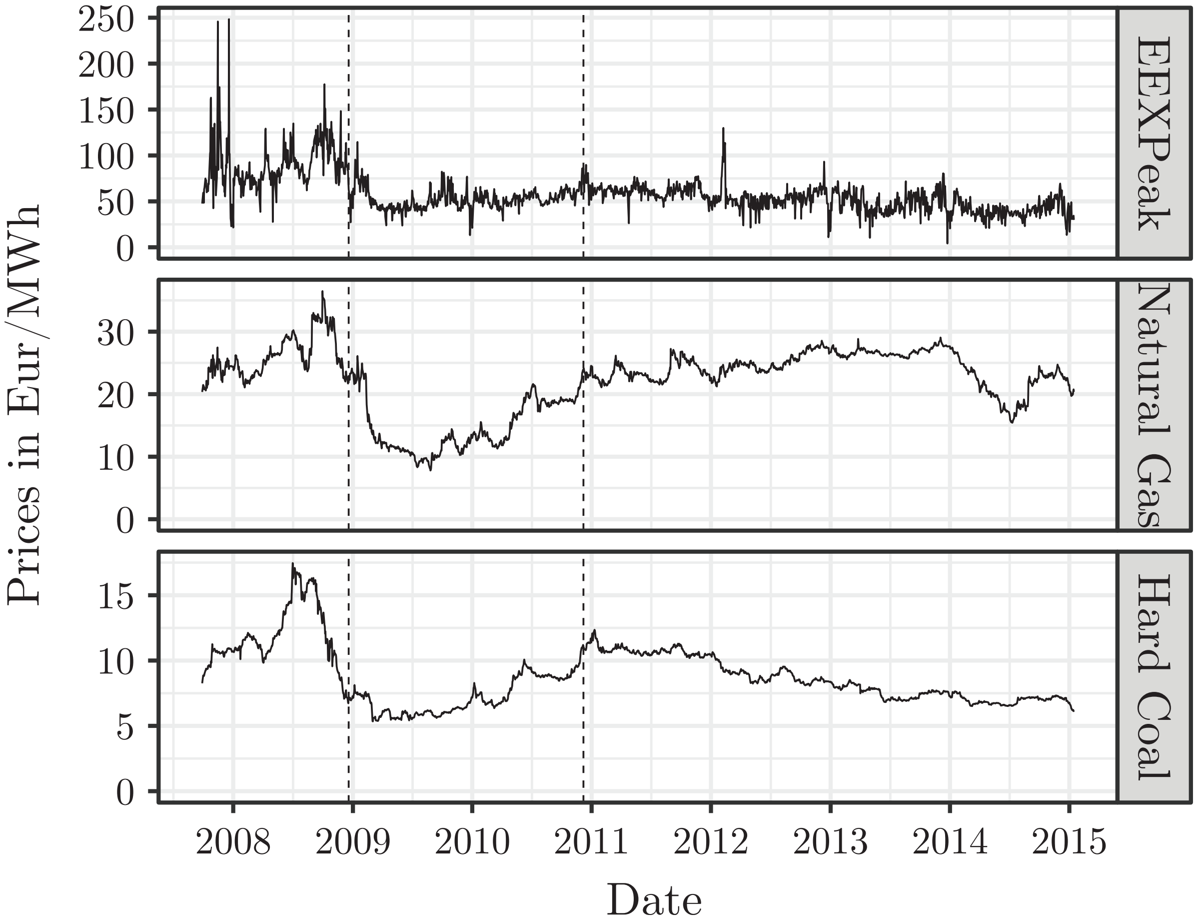



**Figure 2:** Graphical display endogenous variables with the break dates identified by the minimum LM test. (Dashed vertical lines indicate the break dates on December 19, 2008 and December 7, 2010).

framework, which were already indicated by the univariate minimum LM stationarity test. The basic idea of this test is the recursive estimation of a basic VECM, which assesses the constancy of the long-run parameter $\beta$, given that the short-run dynamics are held constant over time. (Hansen and Johansen 1999)

It is important to note that with this test it is only possible to reject the null hypothesis of a stable cointegration parameter, because it does not formulate a specific alternative hypothesis (Hansen and Johansen 1999, 307). Figure 3 shows the recursively estimated test statistic for the cointegration vector $\beta$. Additionally, the vertical lines depict the dates of structural breaks indicated by the minimum LM stationarity test. It is striking that those relatively closely match the period of high values for the test statistic between the end of 2008 and 2010. Since the maximum value of the test statistic 4.797 for the cointegration vector $\beta$, is far greater than the 5 % critical value of 2.44, the null hypothesis of a constant $\beta$ can be safely rejected. Based on the results of the minimum LM stationarity and the stability test of Hansen and Johansen (1999), the detailed analysis of the possible cointegration relationship is conducted using the method initially developed by Johansen, Mosconi, and

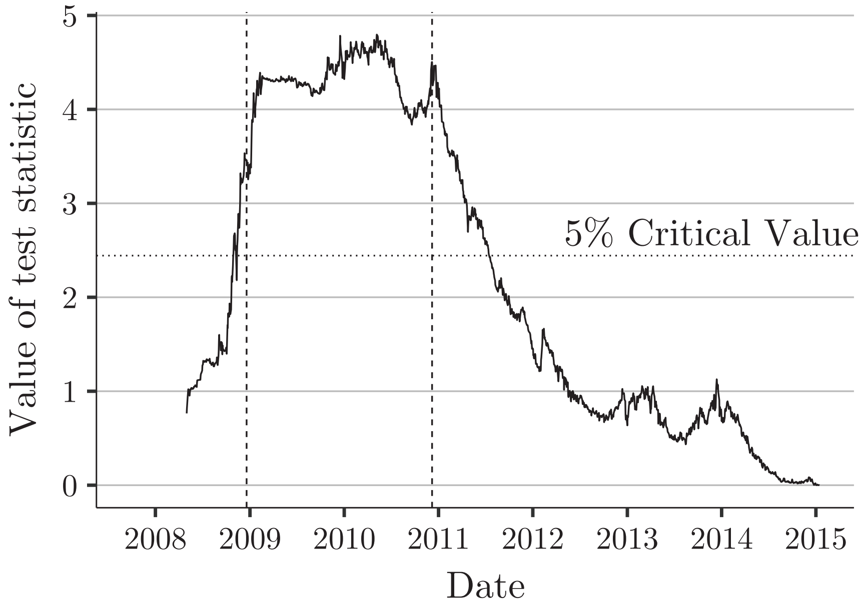



**Figure 3:** Recursively estimated test statistic for $\beta$ constancy (Hansen and Johansen 1999). (dotted horizontal line indicates 5 % critical value at 2.44; dashed vertical lines indicate the break dates on December 19, 2008 and December 7, 2010; first 150 observations not used in the recursive calculation).

Nielsen (2000). Furthermore, it is also used to estimate the whole VECM to determine the nature of the cointegration vector. This method is a generalization of their maximum likelihood cointegration test developed earlier in Johansen (1988, 1991) and allows to consider structural breaks at known points in time. In the following part, the main building blocks of the model are introduced briefly. In order to incorporate consider the structural breaks when testing for the cointegration rank, it is necessary to define $q-1$ intervention and indicator dummies, which indicate each structural break between each subsample $q$. Following the results of the minimum lm stationarity and the stability test of Hansen and Johansen (1999), indicating the presence of at least two structural breaks, which consequently leads to three distinct subsamples, setting $q=3$ in this analysis. The definition of intervention and indicator dummies follows the notation used by Joyeux (2007) and are defined as follows: The intervention dummies are defined as follows:

$$D_{j,t} = \begin{cases} 1 & for \;\; T_{B,j-1} \leq t \leq T_{B,j}, \\ 0 & \text{otherwise}, \end{cases} \qquad \text{for } j=2, \ldots, q,$$

and

$$D_{j,t-k} = \begin{cases} 1 & for \;\; T_{B,j-1} + k + 1 \leq t \leq T_{B,j} + k, \\ 0 & \text{otherwise}, \end{cases} \qquad \text{for } j=2, \ldots, q.$$

The indicator dummies need to be defined according to the following statement:

$$I_{j,t} = \begin{cases} 1 & \mathit{for}\ t = T_{B,j-1} + 1, \\ 0 & \text{otherwise}, \end{cases} \qquad \text{for } j = 2, \ldots, q.$$

The model by Johansen Mosconi, and Nielsen (2000, 218–219) allows to distinguish between three different cointegration hypotheses, whereas in this case only the most general is considered. In this most general form, originally denoted as $H_l(r)$,[12] all time series follow a trending pattern, but it not only allows breaks in the trend of each individual time series, but also in the cointegrating relationship. This implies, that although the cointegration vector $\beta$ is constant over the whole sample period, structural breaks in the trend of the cointegrating relationship are explicitly modeled by including the intervention and indication dummies.

If the following vectors are defined: $D_t = (1, \ldots, D_{q,t})^T, \mu = (\mu_1, \ldots, \mu_q)$, $\gamma = (\gamma_1^T, \ldots, \gamma_q^T)^T$ it is possible to express the model for all $q$ subsamples in a condensed form similar to eq. [5]. The lagged intervention dummy $D_{t-k}$ multiplied with a time trend $t$ is part of the cointegration relationship and has the coefficient $\gamma$.

$$\Delta Y_t = \alpha \begin{pmatrix} \beta \\ \gamma \end{pmatrix}^T \begin{pmatrix} Y_{t-1} \\ tD_{t-k} \end{pmatrix} + \mu D_{t-k} + \sum_{i=1}^{k-1} \Gamma_i \Delta Y_{t-i} + \sum_{i=0}^{k-1} \sum_{j=2}^{q} \kappa_{j,i} I_{j,t-i} + \delta X_t + \varepsilon_t \qquad [5]$$

Due to the generalization, new asymptotic critical values are needed, since the asymptotic distribution of the test statistic now also depends on the locations of the structural breaks in the sample[13] and the difference between the number of time series $p$ and the cointegration rank $r$. To calculate the new critical values and the respective p-values, the procedures implemented by Giles and Godwin (2012) are used.

## 4.2 Empirical Results

### 4.2.1 Empirical Results for the Whole Sample with Structural Breaks

Besides the endogenous price series of electricity, natural gas and hard coal, all the additional variables discussed in Section 2 are included in the model as

---

**12** $r$ denotes the cointegration rank.

**13** Breakpoints are denoted as $\lambda_j = \frac{T_{B,j}}{T}$, where $T$ is total number of observations and $T_{B,j}$ is the last observation of subsample $j$, with $j = 1, 2, \ldots, q$.

exogenous variables to account for possible effects of weather and seasonality. The VECM is implemented as presented in the previous section, with no constant or trend in the cointegration relation and the lag order for the endogenous variables is set to five, according to the Hannan-Quinn information criteria. This additionally allows to properly model the weekly structure of the data, but at the same time preserves the model's parsimony.

In a first step, the cointegration rank of the system is determined based on the trace test statistic. The results of this test can be found in Table 3, together with the calculated asymptotic critical values. Based on these results, the null hypothesis of no cointegration rank can be rejected and it is therefore safe to assume that at least one cointegration vector exists. Starting from these results, the VECM is estimated with the restriction of only one cointegration rank. In Table 4 the cointegration vector $\beta$, which reports the long-run relation between the variables, and the $\alpha$ vector, which indicates if and how the variables react to deviations from the long-run relationship, are presented.

**Table 3:** Trace test statistic to determine the cointegration rank. Critical values are derived according to Giles and Godwin (2012).

| Rank | Trace test statistic | 10 % | 5 % | 1 % |
|---|---|---|---|---|
| $r \leq 2$ | 2.88 | 21.03 | 23.60 | 28.94 |
| $r \leq 1$ | 23.58 | 42.20 | 45.54 | 52.27 |
| $r = 0$ | 236.20*** | 67.02 | 71.08 | 79.11 |

**Table 4:** Cointegration relationship for a VECM with a cointegrating rank $r = 1$, including the loading parameters in the $\hat{\alpha}$-vector and the coefficients in the stacked vector of $\hat{\beta}$ and $\hat{\gamma}$, which incorporates the coefficients of the endogenous and the intervention variables for the two structural breaks.

| | $\hat{\alpha}$-vector | | $\hat{\beta}$- and $\hat{\gamma}$-vector | |
|---|---|---|---|---|
| | Parameter | t-stat | Parameter | t-stat |
| EEXPeak | −0.3239*** | −14.9779 | 1.0000 | — |
| Natural Gas | −0.0002 | −0.18 | −1.7372*** | −7.22 |
| Hard Coal | 0.0000 | 0.09 | −3.3350*** | −6.29 |
| $tD_{1,t-5}$ | – | – | 0.0235** | 2.524 |
| $tD_{2,t-5}$ | – | – | 0.0037 | 0.9749 |

The estimated cointegration vector, $(\hat{\beta}, \hat{\gamma})^T$, in Table 4 shows that both price time series, natural gas and hard coal, are part of a long-run relationship and are important drivers of the electricity price during times of high demand. The coefficients of the two fossil fuels have the theoretically expected negative sign, which implies that an increase in one of these input factors leads to an increase in electricity prices. Given that the energy efficiency of power plants is only around 33 % for hard coal and 41 % for natural gas, meaning that only this share of the energy input is transformed into electricity.[14] It is interesting to see that apparently, electricity prices in the long run react over proportionally to price changes of the fossil fuels. Furthermore, when looking at the coefficients of the $tD_{j,t-5}$ intervention dummies, which take the structural breaks into account, it can be seen that only the dummy covering the second subsample from December 2008 to December 2010 is significant.

In order to analyze if the natural gas and hard coal prices are weakly exogenous for the electricity price, a Likelihood Ratio (LR) test, based on Johansen (1991), is applied on the $\alpha$ vector, which models the speed of adjustment to the long-run equilibrium. It is not possible to reject the simultaneous linear restriction that both coefficients are actually zero. Therefore, it is safe to assume, that both fossil fuels are weakly exogenous in the short run.

To assess the long-run relationship in more detail, a similar LR test is also applied to the cointegration vector $\beta$. For this purpose, a linear restriction is imposed, which restricts the coefficients of natural gas and hard coal prices to zero. As expected from the values in Table 4 it is possible to strongly reject the imposed restrictions and it can be assumed that in the long-run electricity prices are influenced by changes in natural gas and hard coal prices.

The Figures 4, 5 and 6 show the the impulse response functions (IRFs) of the electricity price to a shock in all three endogenous time series. (IRFs) for the impact on natural gas and hard coal are shown in appendix (C) for completeness. It can be seen that shocks to the electricity price are corrected within a relatively short time period. While a positive change of natural gas prices increases the electricity price in the short-run, a price increase in the hard coal market does only affect the electricity price with a delay of a couple of days and the effect is not statistically significant. These differences in the reaction of hard coal and natural gas fired power plants could possibly be attributed to different technological frictions between coal-burning and natural gas fired power plants. These technical frictions mainly consist of higher maintenance costs for switching fuels, varying load or other changes to the electricity production. These costs

**14** Calculations of energy efficiencies for various energy sources is based on average operating heat rates as published by the U.S. Energy Information Administration (EIA) (2016).

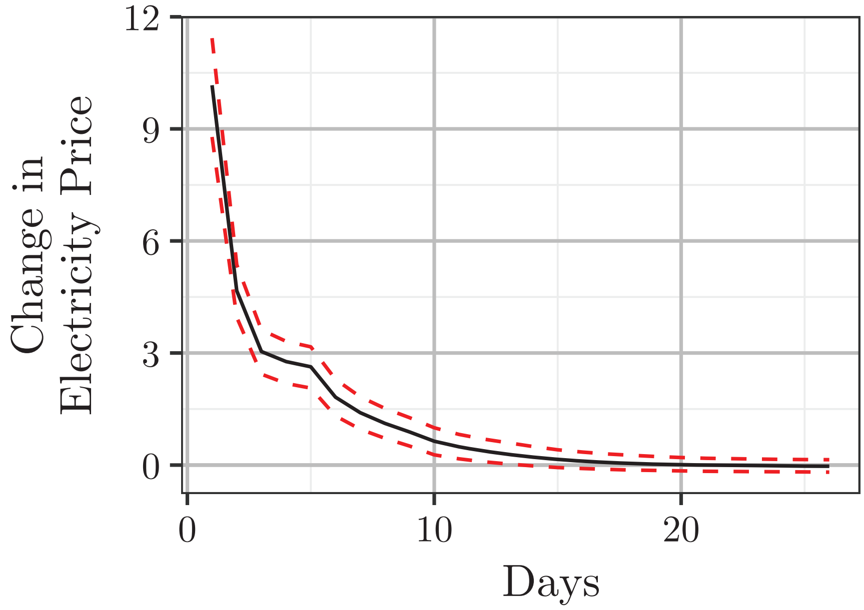



**Figure 4:** Response of EEXPeak electricity price to a shock in the EEXPeak electricity price. (dashed red lines indicate the 95 % bootstrapped C.I.s).

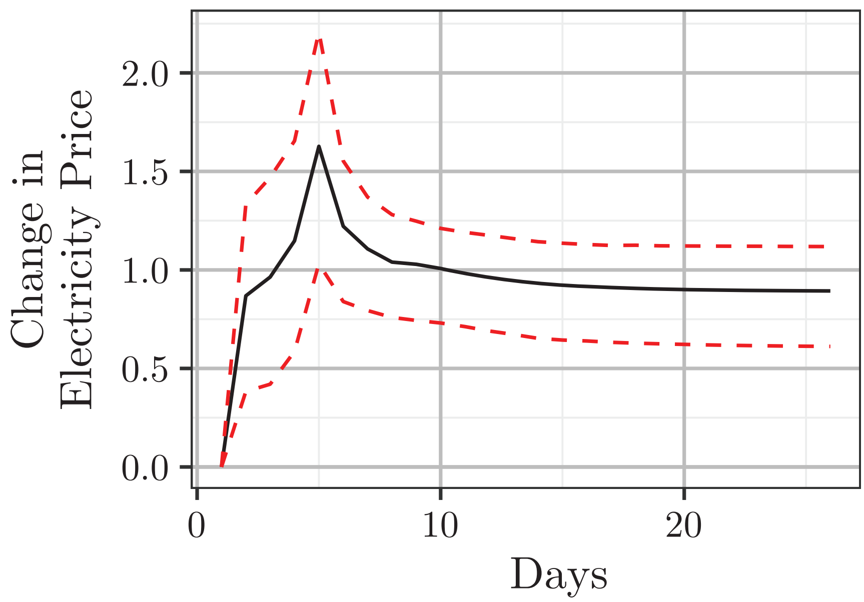



**Figure 5:** Response of EEXPeak electricity price to a shock in the natural gas price. (dashed red lines indicate the 95 % bootstrapped C.I.s).

are considerably higher for coal-fired power plants than for power plants using natural gas. (Matisoff, Noonan, and Cui 2014, 3)

The Breusch-Godfrey test, amongst others, indicates the presence of serial correlation in the residuals, hence the results need to be interpreted with caution. Additionally, a test for possible ARCH effects in the residuals also allows to reject the null hypothesis of no ARCH effects. Although, according to the results of Silvapulle and Podivinsky (2000) possible ARCH or GARCH effects

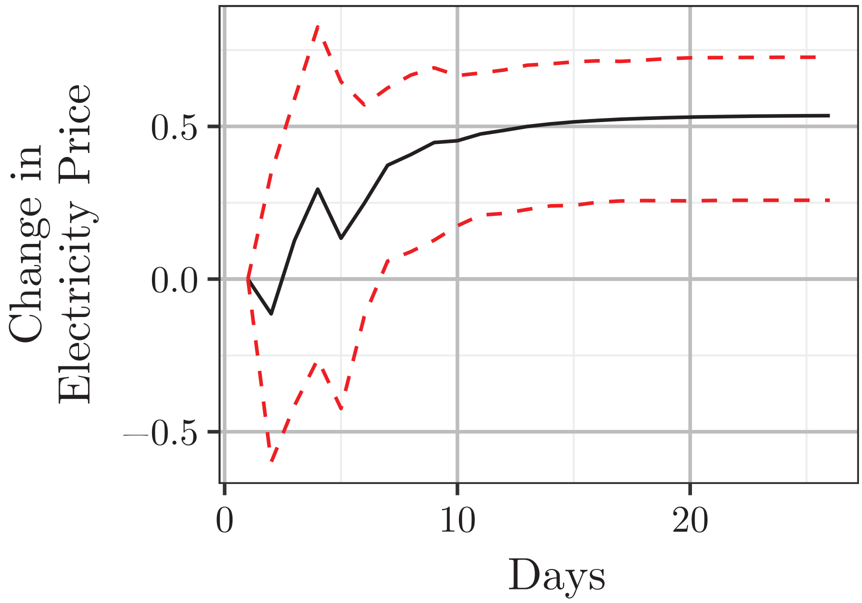



**Figure 6:** Response of EEXPeak electricity price to a shock in the hard coal price. (dashed red lines indicate the 95 % bootstrapped C.I.s).

might not affect the results of the cointegration analysis too much. Nevertheless, to increase the efficiency of the obtained results it might be useful to extend the VECM with a GARCH error structure, which incorporates the structure of the residuals explicitly.

### 4.2.2 Empirical Results for Each Subsample

An implementation of the VECM framework for each of the three subsamples supports the previous results that the cointegration relationship underwent considerable changes over the course of the sample period.

The long-run coefficient of natural gas decreased from −6.662 in the first period to −2.223 in the second and is a mere −0.866 in the last period. This is a strong indication that the relationship between the electricity and the natural gas price became less important over the horizon of this analysis, although it's not a definitive proof of this hypothesis.

Another interesting result, when looking at the development of the cointegration vector $\beta$, is that over the sample horizon the impact of hard coal prices also changed. It changed, however, differently than that of natural gas prices. It is not significant in the first subsample and turns positive in the second period, which is not in line with the assumed theoretical relationship, because it means that an increase in hard coal prices leads to decreasing electricity prices. For the last period, the coefficient becomes negative and highly significant. This indicates that the impact of hard coal prices, in contrast to the natural gas price, increased over the sample horizon. One possible explanation for these two opposing developments could be the merit-order effect. The increasing generation capacity of renewable energy sources shifts the merit-order to the right. This then leads to the possibility to satisfy the electricity demand using generation capacity with lower marginal costs, namely the substitution of natural gas fired power plants by hard coal fired ones.

# 5 Conclusion and Further Research

This analysis examined the long-run relations and the short-run dynamics between major fossil fuels used for electricity generation and electricity prices, during a time of fundamental changes to the German electricity market. The econometric model incorporated these fundamental changes into the analysis and allowed to show that there still exists a strong cointegration relation between the prices of fossil fuels and electricity, even when taking structural breaks in the cointegration relation into account. There is strong evidence for a significant long-run impact of natural gas and hard coal prices on the price for electricity during times of high load. This, however, has strong policy implications for the aim reducing the reliance on fossil fuels, especially on hard coal or lignite, and thereby curbing carbon emissions. Since these policies might lead to increasing costs for hard coal generation capacity and hence also to higher electricity prices. The short-run dynamics are characterized by a significant and instant impact of shocks to natural gas prices on electricity prices. Whereas, an increase in hard coal prices only has a significant impact after seven days. These differences in the short-run dynamics can most probably be attributed to different characteristics of the markets, since the trading and transportation properties of coal are less flexible than the entry-exit regime of the German natural gas market.

The results of this analysis are potentially useful for other countries, which are at a different stage of replacing fossil fuels with renewable energy sources to satisfy electricity demand and reduce carbon emissions. Although most electricity markets and energy sectors differ between countries or regions, so the results are only valid for the German electricity market and can't be easily transferred to other countries.

In preparation for the cointegration analysis the non-stationarity hypothesis was examined using the test proposed by Lee and Strazicich (2003, 2013), which allows two structural breaks in the time series. In connection with the subsequently employed bootstrap approach of this test critical values for the test statistic and C.I.s for the break dates were calculated. Based on these results it was not possible to reject the unit-root null hypothesis and two structural breaks in intercept and trend were identified. Further indication for structural breaks, not only in the univariate time series, but also in the multivariate VECM framework were given by the applied stability test of Hansen and Johansen (1999).

Since the aim of this analysis mainly was to assess if fossil fuels still influence electricity prices, a couple of additional influencing factors were not included. Moreover, in some cases it was a problem of data availability, which prevented the inclusion of prices for carbon emission certificates and the cross-border flows of electricity. Furthermore, it would be interesting to assess the influence of lignite, since it constitutes a sizable amount of generation capacity and also is used to generate around 30 % of all electricity Germany (BDEW 2016). However, since there is no liquid market for lignite and most of it is directly burned close to the mining site, no market price for lignite exists. A probably more important effect might be the actual production from renewable sources. If sufficient data on actual production from renewables would be available, it could directly be incorporated in the econometric model. In this case it would be possible to get a better understanding of suspected non-linearities in the relationship, depending on the amount of electricity generated from renewable energy sources.

Econometric methods modelling these non-linearities, include for example the threshold cointegration methods by Balke and Fomby (1997). If these non-linearities are themselves functions of exogenous variables, as it might be the case here, the application of an open-loop threshold autoregressive system (TARSO) model might be beneficial (Tong 1990).

Another area for future research could be to explicitly model the time-varying nature of the cointegration relationship. This would be possible by using the time-varying VECM framework, by Bierens and Martins (2010), which is an extension of the methods proposed by Johansen (1988, 1991, 1995) and allows to analyze the development of the cointegration relationship over time. Another promising approach in this research area might be using a Bayesian framework, for example Koop, Leon-Gonzalez, and Strachan (2011) also offer the possibility to explicitly allow the cointegration space to evolve over time

# Appendix

## A. Development of Electricity Production and Installed Capacity by Energy Source

**Table 5:** Electricity production by energy source in GWh (BDEW 2016).

| | 2007 | 2008 | 2009 | 2010 | 2011 | 2012 | 2013 | 2014 |
|---|---|---|---|---|---|---|---|---|
| **Nuclear** | 133 229 | 140 710 | 127 690 | 132 971 | 102 241 | 94 180 | 92 127 | 91 800 |
| **Lignite** | 142 328 | 138 090 | 133 653 | 134 169 | 137 888 | 148 147 | 149 163 | 144 328 |
| **Hard Coal** | 130 799 | 114 423 | 98 773 | 107 357 | 103 177 | 106 755 | 116 755 | 108 670 |
| **Gas** | 75 447 | 86 244 | 78 236 | 86 560 | 83 505 | 74 000 | 65 265 | 58 911 |
| **Oil** | 9011 | 8722 | 9058 | 7860 | 6364 | 6785 | 6446 | 5031 |
| **Renewables** | 81 779 | 86 433 | 88 303 | 97 789 | 116 650 | 135 997 | 144 124 | 153 684 |
| Hydro | 20 751 | 20 098 | 18 697 | 20 650 | 17 304 | 21 697 | 22 654 | 19 322 |
| Biomass | 18 359 | 21 463 | 24 492 | 27 734 | 31 011 | 37 402 | 38 907 | 41 121 |
| Wind | 39 594 | 40 452 | 38 531 | 37 677 | 48 736 | 50 518 | 51 553 | 57 185 |
| Solar | 3075 | 4420 | 6583 | 11 728 | 19 599 | 26 380 | 31 010 | 36 056 |

**Table 6:** Electricity generation capacity by energy source in GW (Burger 2016).

| | 2007 | 2008 | 2009 | 2010 | 2011 | 2012 | 2013 | 2014 |
|---|---|---|---|---|---|---|---|---|
| **Nuclear** | 21.3 | 21.6 | 21.5 | 21.5 | 12.1 | 12.1 | 12.1 | 12.1 |
| **Lignite** | 22.5 | 22.4 | 22.4 | 22.7 | 19.9 | 21.3 | 21.2 | 21.3 |
| **Hard Coal** | 29.3 | 29.6 | 29 | 30.2 | 25.7 | 25.1 | 25.9 | 26.2 |
| **Gas** | 21.3 | 22.8 | 23.1 | 23.8 | 27.1 | 27.2 | 28.2 | 28.9 |
| **Oil** | 5.4 | 5.4 | 5.2 | 5.9 | 4.2 | 4.1 | 4.1 | 4.2 |
| **Renewables** | 36.2 | 40.4 | 47.6 | 57.0 | 67.5 | 78.1 | 84.7 | 91.0 |
| Hydro | 5.1 | 5.2 | 5.3 | 5.4 | 5.6 | 5.6 | 5.6 | 5.6 |
| Biomass | 4.7 | 5.3 | 6 | 6.6 | 7.2 | 7.6 | 8.4 | 8.9 |
| Wind | 22.2 | 23.8 | 25.7 | 27.1 | 28.8 | 30.8 | 34.0 | 38.3 |
| Solar | 4.2 | 6.1 | 10.6 | 17.9 | 26.0 | 34.1 | 36.7 | 38.2 |

**Table 7:** Full-load hours by energy source (own calculations based on BDEW (2016) and Burger (2016)).

| | 2007 | 2008 | 2009 | 2010 | 2011 | 2012 | 2013 | 2014 |
|---|---|---|---|---|---|---|---|---|
| **Nuclear** | 6255 | 6514 | 5939 | 6185 | 8471 | 7803 | 7633 | 7606 |
| **Lignite** | 6326 | 6165 | 5967 | 5911 | 6946 | 6965 | 7033 | 6792 |
| **Hard Coal** | 4464 | 3866 | 3406 | 3555 | 4016 | 4246 | 4503 | 4149 |
| **Gas** | 3542 | 3783 | 3387 | 3637 | 3084 | 2720 | 2313 | 2036 |

(*continued*)

**Table 7:** (*continued*)

| | 2007 | 2008 | 2009 | 2010 | 2011 | 2012 | 2013 | 2014 |
|---|---|---|---|---|---|---|---|---|
| **Oil** | 1669 | 1615 | 1742 | 1332 | 1526 | 1639 | 1557 | 1187 |
| **Renewables** | 2258 | 2142 | 1856 | 1715 | 1728 | 1742 | 1702 | 1689 |
| Hydro | 4037 | 3895 | 3501 | 3817 | 3074 | 3868 | 4031 | 3463 |
| Biomass | 3890 | 4080 | 4082 | 4202 | 4337 | 4954 | 4643 | 4641 |
| Wind | 1785 | 1698 | 1501 | 1392 | 1694 | 1639 | 1518 | 1492 |
| Solar | 737 | 722 | 623 | 654 | 755 | 774 | 845 | 943 |

# B. Histograms of Bootstrap Results Stationarity Test

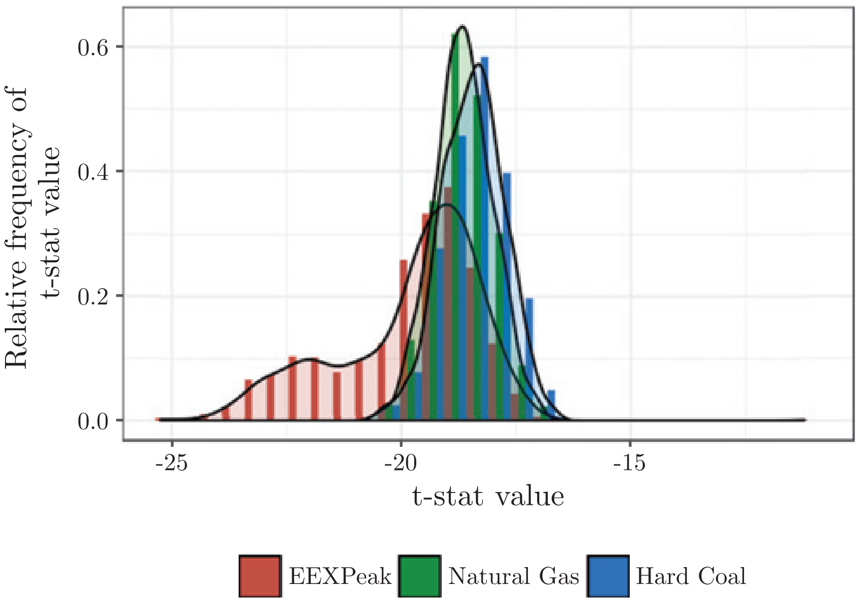



**Figure 7:** Histogram for the value of the t-statistic.

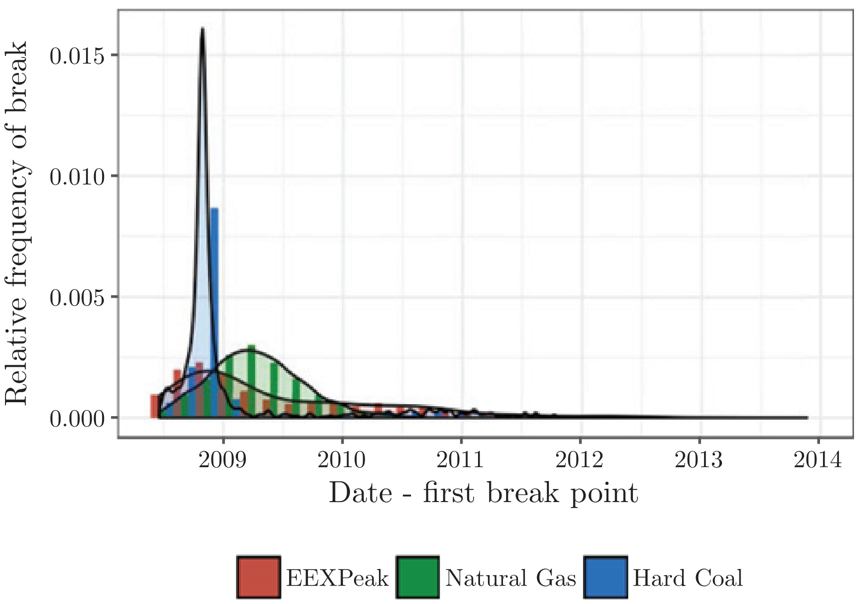



**Figure 8:** Histogram for the location of the first break.

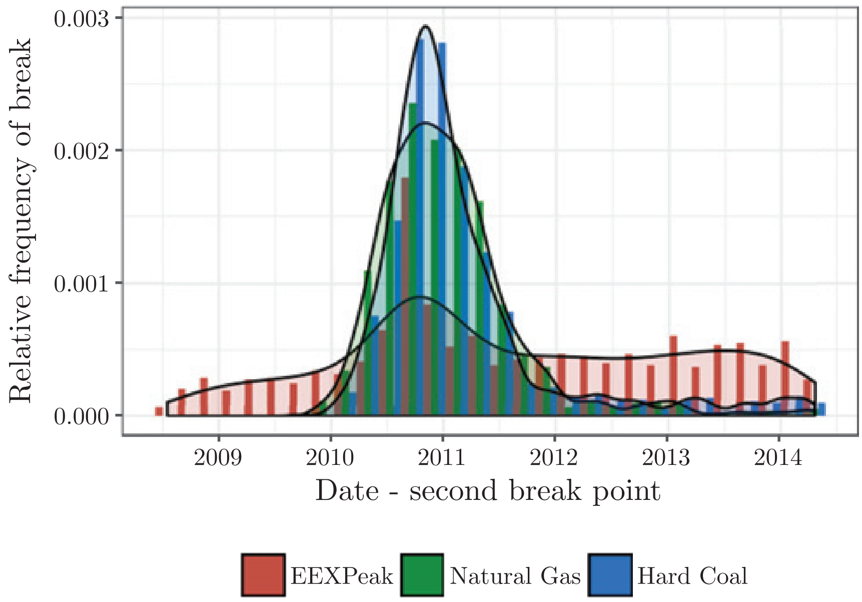



**Figure 9:** Histogram for the location of the second break.

## C. Impulse Response Functions for Natural Gas and Hard Coal

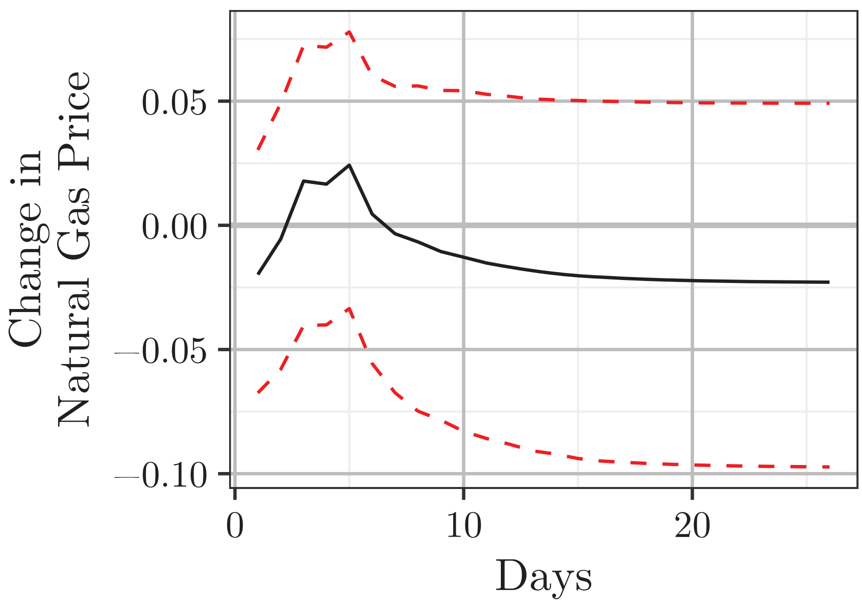



**Figure 10:** Response of the natural gas price to a shock in the EEXPeak electricity price (dashed red lines indicate the 95 % bootstrapped C.I.s).

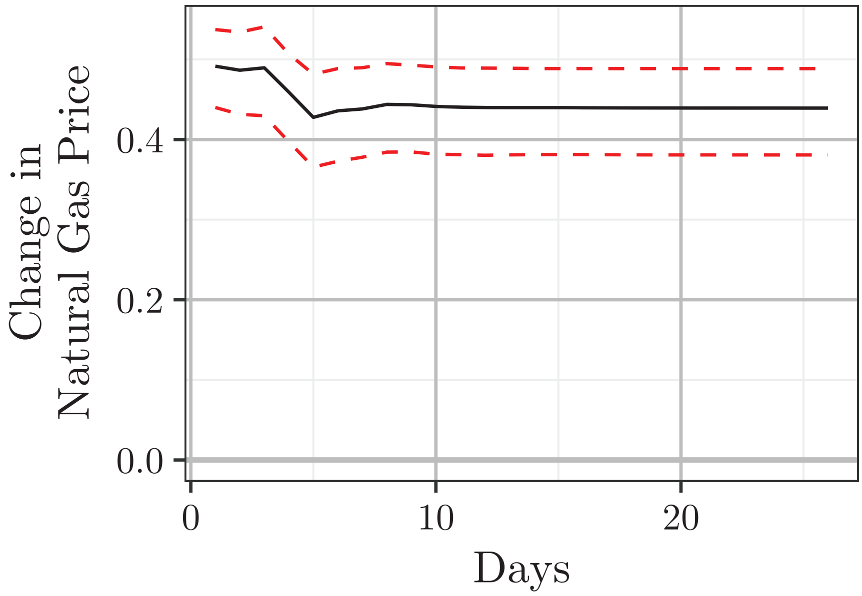



**Figure 11:** Response of the natural gas price to a shock in the natural gas price (dashed red lines indicate the 95 % bootstrapped C.I.s).

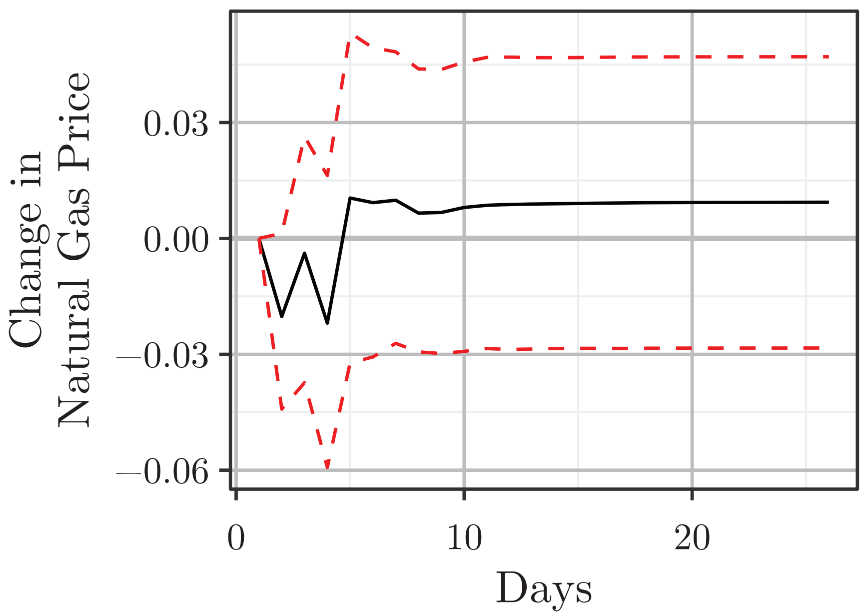



**Figure 12:** Response of the natural gas price to a shock in the hard coal price (dashed red lines indicate the 95 % bootstrapped C.I.s).

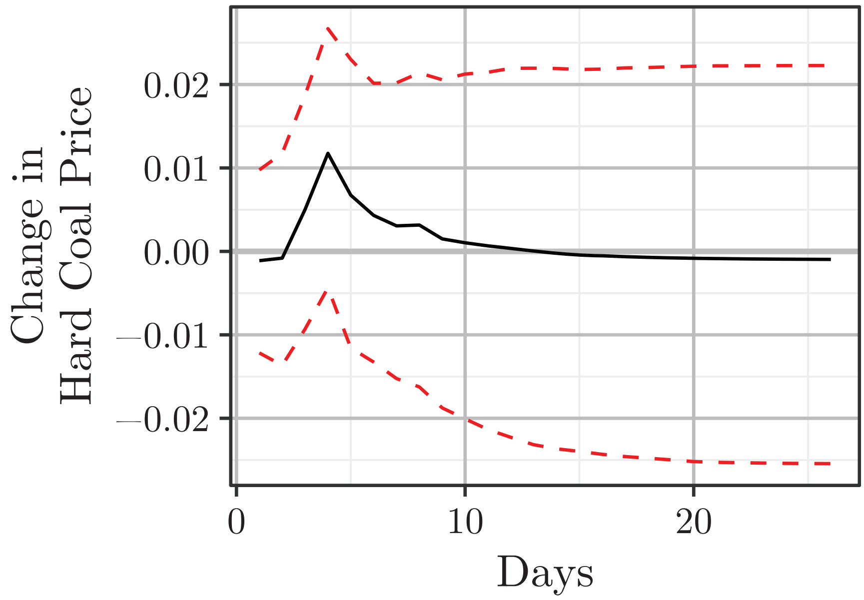



**Figure 13:** Response of the hard coal price to a shock in the EEXPeak electricity price (dashed red lines indicate the 95 % bootstrapped C.I.s).

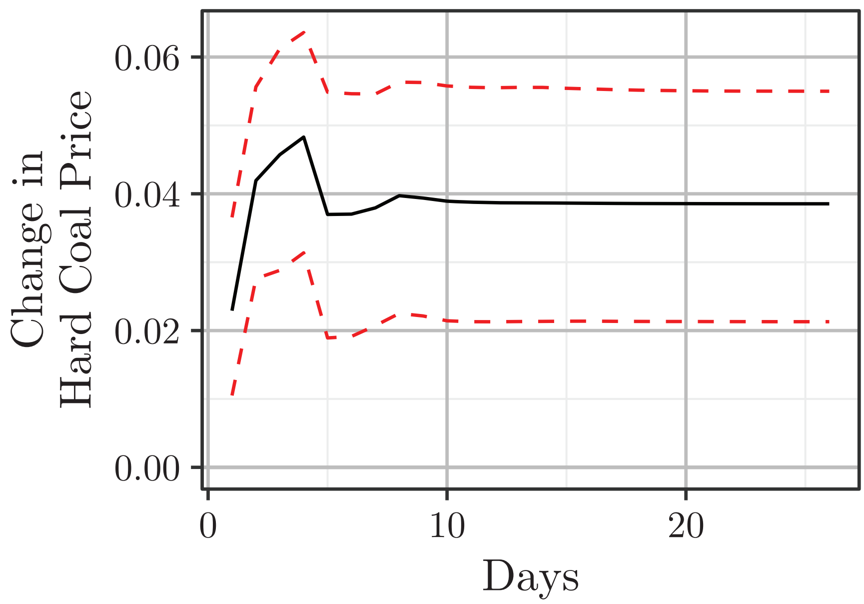



**Figure 14:** Response of the hard coal price to a shock in the natural gas price (dashed red lines indicate the 95 % bootstrapped C.I.s).

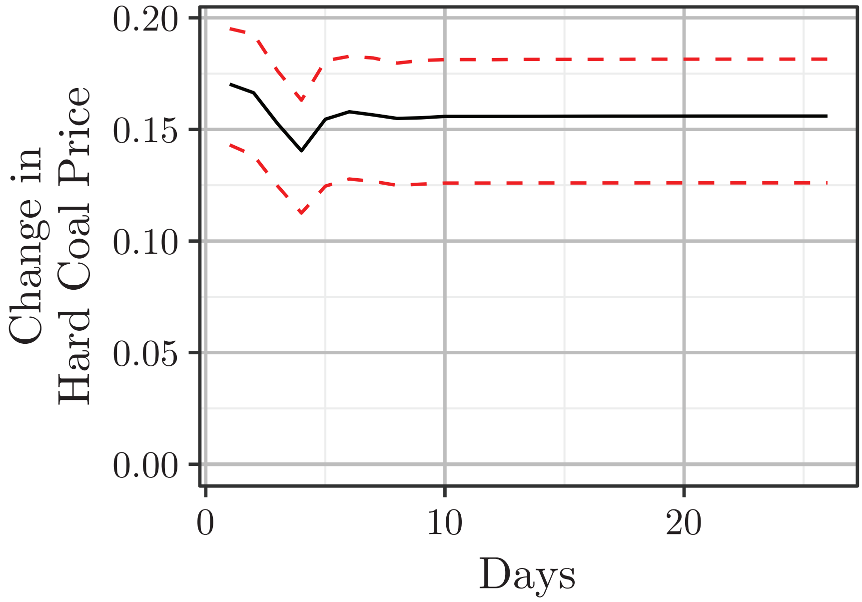



**Figure 15:** Response of the hard coal price to a shock in the hard coal price (dashed red lines indicate the 95 % bootstrapped C.I.s).

# D. Results of Cointegration Analysis for Different Subsamples

## Initial Period

**Table 8:** Trace test statistic to determine the cointegration rank for the first period from 28/09/2007 to 19/12/2008.

| Rank | Trace test stat | 10 % | 5 % | 1 % |
|---|---|---|---|---|
| r ≤ 2 | 3.81 | 6.5 | 8.18 | 11.65 |
| r ≤ 1 | 25.52 | 15.66 | 17.95 | 23.52 |
| r = 0 | 88.77*** | 28.71 | 31.52 | 37.22 |

**Table 9:** Cointegration parameters for the first period from 28/09/2007 to 19/12/2008.

| | $\alpha$-vector | | $\beta$-vector | |
|---|---|---|---|---|
| | Parameter | t-stat | Parameter | t-stat |
| EEXPeak | −0.4769*** | −7.88 | 1.0000 | — |
| Natural gas | 0.0046* | 1.70 | −6.6616*** | −8.64 |
| Hard coal | 0.0025** | 2.20 | −0.2421 | −0.18 |

## Second Period

**Table 10:** Trace test statistic to determine the cointegration rank for the second period from 22/12/2008 to 07/12/2010.

| Rank | Trace test stat | 10 % | 5 % | 1 % |
|---|---|---|---|---|
| r ≤ 2 | 0.4 | 6.5 | 8.18 | 11.65 |
| r ≤ 1 | 16.92 | 15.66 | 17.95 | 23.52 |
| r = 0 | 80.59*** | 28.71 | 31.52 | 37.22 |

**Table 11:** Cointegration parameters for the second period from 22/12/2008 to 07/12/2010.

| | $\alpha$-vector | | $\beta$-vector | |
|---|---|---|---|---|
| | Parameter | t-stat | Parameter | t-stat |
| EEXPeak | −0.2926*** | −7.43 | 1.0000 | — |
| Natural gas | 0.0062** | 2.03 | −2.2230*** | −6.02 |
| Hard coal | −0.0008 | −0.90 | 2.7142** | 2.42 |

## Third Period

**Table 12:** Trace test statistic to determine the cointegration rank for the third period from 08/12/2010 to 15/01/2015.

| Rank | Trace test stat | 10 % | 5 % | 1 % |
|---|---|---|---|---|
| r ≤ 2 | 0.48 | 6.5 | 8.18 | 11.65 |
| r ≤ 1 | 3.66 | 15.66 | 17.95 | 23.52 |
| r = 0 | 152.81*** | 28.71 | 31.52 | 37.22 |

**Table 13:** Cointegration parameters for the third period from 08/12/2010 to 15/01/2015.

| | $\alpha$-vector | | $\beta$-vector | |
|---|---|---|---|---|
| | Parameter | t-stat | Parameter | t-stat |
| EEXPeak | −0.3568*** | −17.96 | 1.0000 | — |
| Natural gas | 0.0013 | 1.37 | −0.8660** | −2.05 |
| Hard coal | 0.0009*** | 3.54 | −4.4924*** | −6.12 |

# References


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

BDEW (2016): "Electricity Generation in Germany by Energy Sources 2000 to 2015." https://www.bdew.de/internet.nsf/id/DE_Energiedaten{#}3-5-netto-stromerzeugung-in-deutschland-nach-erzeugern-de.

Bierens, H. J., and L. F. Martins. 2010. "Time-Varying Cointegration." *Econometric Theory* 26 (05):1453–1490.

Bosco, B., L. Parisio, M. Pelagatti, and F. Baldi. 2010. "Long-Run Relations in European Electricity Prices." *Journal of Applied Econometrics* 25 (5):805–832.

Burger, B. (2016): "Energy-Charts." www.energy-charts.de. Accessed: 2016-06-23.

Chou, W. L. 2007. "Performance of LM-Type Unit Root Tests with Trend Break: A Bootstrap Approach." *Economics Letters* 94 (1):76–82.

Cottrell, A., and R. Lucchetti. 2015. *Gretl User's Guide*.

Deutscher Wetterdienst (DWD). 2015. *Daily Climate Data for 73 Stations*. German Meteorological Service: Offenbach am Main, Germany.

Energy Information Administration (EIA) (2016): *Average Operating Heat Rate for Selected Energy Sources*. http://www.eia.gov/electricity/annual/html/epa_08_01.html. Accessed: 2016-02-23.

Fell, H. 2010. "EU-ETS and Nordic Electricity: A CVAR Analysis." *The Energy Journal* 31 (2):1–26.

Ferkingstad, E., A. Løland, and M. Wilhelmsen. 2011. "Causal Modeling and Inference for Electricity Markets." *Energy Economics* 33 (3):404–412.

Fezzi, C., and D. Bunn. 2009. "Structural Interactions of European Carbon Trading and Energy Prices." *The Journal of Energy Markets* 2 (4):53 69.

Giles, D. E., and R. T. Godwin. 2012. "Testing for Multivariate Cointegration in the Presence of Structural Breaks: P-Values and Critical Values." *Applied Economics Letters* 19 (16):1561–1565.

Haas, R., G. Lettner, H. Auer, and N. Duic. 2013. "The Looming Revolution: How Photovoltaics Will Change Electricity Markets in Europe Fundamentally." *Energy* 57:38–43.

Halvorsen, R. 1975. "Residential Demand for Electric Energy." *The Review of Economics and Statistics* 57 (1):12–18.

Hansen, H., and S. Johansen. 1999. "Some Tests for Parameter Constancy in Cointegrated VAR-Models." *The Econometrics Journal* 2 (2):306–333.

Johansen, S. 1988. "Statistical Analysis of Cointegration Vectors." *Journal of Economic Dynamics and Control* 12 (2–3):231–254.

Johansen, S. 1991. "Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models." *Econometrica* 59.6:1551–1580.

Johansen, S. 1995. *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models*. Advanced texts in econometrics. Oxford and New York: Oxford University Press.

Johansen, S., R. Mosconi, and B. Nielsen. 2000. "Cointegration Analysis in the Presence of Structural Breaks in the Deterministic Trend." *Econometrics Journal* 3 (2):216–249.

Joyeux, R. 2007. "How to Deal with Structural Breaks in Practical Cointegration Analysis." In *Cointegration for the Applied Economist*, 2nd ed., Edited by B. B. Rao, 256. Basingstoke: Palgrave Macmillan.

Koop, G., R. Leon-Gonzalez, and R. W. Strachan. 2011. "Bayesian Inference in a Time Varying Cointegration Model." *Journal of Econometrics* 165 (2):210–220.

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

Lee, J., and M. C. Strazicich. 2013. "Minimum LM Unit Root Test with One Structural Break." *Economics Bulletin* 33 (4):2483–2492.

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

Matisoff, D. C., D. S. Noonan, and J. Cui. 2014. "Electric Utilities, Fuel Use, and Responsiveness to Fuel Prices." *Energy Economics* 46:445–452.

Mjelde, J.W., and D. A. Bessler. 2009. "Market Integration among Electricity Markets and Their Major Fuel Source Markets." *Energy Economics* 31 (3):482–491.

Mohammadi, H. 2009. "Electricity Prices and Fuel Costs: Long-Run Relations and Short-Run Dynamics." *Energy Economics* 31 (3):503–509.

Ng, S., and P. Perron. 1995. "Unit Root Tests in ARMA Models with Data-Dependent Methods for the Selection of the Truncation Lag." *Journal of the American Statistical Association* 90 (429):268–281.

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

Quayle, R. G., and H. F. Diaz. 1980. "Heating Degree Day Data Applied to Residential Heating Energy Consumption." *Journal of Applied Meteorology* 19 (3):241–246.

R Core Team. 2016. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.

Schmidt, P., and P. C. B. Phillips. 1992. "LM Tests for a Unit Root in the Presence of Deterministic Trends." *Oxford Bulletin of Economics and Statistics* 54 (3):257–287.

Silvapulle, P. S., and J. M. Podivinsky. 2000. "The Effect of Non-Normal Disturbances and Conditional Heteroskedasticity on Multiple Cointegration Tests." *Journal of Statistical Computation and Simulation* 65 (1-4):173–189. http://dx.doi.org/10.1080/00949650008811997.

Thoenes, S. (2011): *Understanding the Determinants of Electricity Prices and the Impact of the German Nuclear Moratorium in 2011*. EWI Working Papers 2011–6. Energiewirtschaftliches Institut, Universität Köln.

Tong, H. 1990. *Non-Linear Time Series: A Dynamical System Approach*. Oxford statistical science series, Oxford: Clarendon.

Tveten, Å. G., T. F. Bolkesjø, T. Martinsen, and H. Hvarnes. 2013. "Solar Feed-In Tariffs and the Merit Order Effect: A Study of the German Electricity Market." *Energy Policy* 61:761–770.

Weron, R.. 2006. *Modeling and Forecasting Electricity Loads and Prices: A Statistical Approach (The Wiley Finance Series)*, Vol. 403. Chichester & Hoboken, NJ: John Wiley & Sons.

Weron, R.. 2014. "Electricity Price Forecasting: A Review of the State-Of-The-Art with A Look into the Future." *International Journal of Forecasting* 30 (4):1030–1081.

Zivot, E., and D. W. K. Andrews. 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.

# Chapter 3 Debt and the Oil Industry – Analysis on the Firm and Production Level

**Reference for this Paper:**

Forthcoming as: Lips, J. (2019): "Debt and the oil industry: analysis on the firm and production level". In: *Journal of Energy Markets* 12 (4), pp. 1–29. DOI: 10.21314/JEM.2019.189.

**Conferences with Review Process:**

- 15th IAEE European Conference, Vienna, September 2017. (Presentation)
- 10th CFE-CMStatistics 2017, London, December 2017. (Presentation)
- Annual Conference of the Verein für Socialpolitk, Freiburg, September 2018. (Presentation)

# Debt and the Oil Industry – Analysis on the Firm and Production Level *

Johannes Lips†

This version: September 2019

This paper analyzes the relationship between debt and the production decision of companies active in the exploration and production of oil and gas in the United States (US). Over the last couple of years, the development and application of innovative extraction methods led to a considerable increase in US oil production. In connection with these technological changes, another important economic development in the oil industry was largely debt-driven investment in the oil sector. The extensive use of debt was fostered by the macroeconomic environment in the aftermath of the financial crisis. Additionally, the rising prices in the commodities markets until mid-2014 led to higher asset valuation and a virtuous circle. This increase in investment activity, especially in the US, raised the production capacity and as a consequence also the production of crude oil. This trend continued in spite of the oil price decline in 2014, although production reductions would have been more plausible.
The main research question of this paper is whether debt and leverage affect production decisions of companies and addresses this question using a novel panel vector autoregressive (VAR) approach and a data set combining financial data on publicly listed firms and their production data on well level.

**Keywords:** Corporate Finance, Oil Industry, Debt, Leverage, Panel VAR, Dynamic Panel Data

**JEL classification:** C33, C58, G01, G30, Q40

*I am grateful to Karol Kempa, Christoph Funk, Daniel Grabowski, Peter Winker and participants of the 15th IAEE European Conference, Vienna 2017, the 11th CFE, London (2017) and the Annual Conference of the Verein fuer Socialpolitik, Freiburg (2018) for helpful comments and support. I am especially thankful to DrillingInfo for providing me with free access to their valuable database (http://info.drillinginfo.com/).
†Department of Economics, Justus-Liebig University of Giessen, Chair of Statistics and Econometrics (VWL VII), Licher Str. 64, 35394 Giessen, Germany. E-mail: Johannes.Lips@wirtschaft.uni-giessen.de

# 1 Introduction

In 2014, the price of crude oil markedly declined following a period of relative stability, during which it stayed at around $100. The price recovered relatively quickly from the subsequent decline following the financial crisis in 2008. In comparison with previous episodes of oil price declines, it appears to be more difficult to identify a single underlying cause explaining the persistently low prices of crude oil. It rather seems to be a result of the interplay between multiple factors, both on the demand and supply side of the global market for crude oil. It appears that market participants underestimated the expected crude oil production and at the same time overestimated the demand for oil, which was mainly subdued by weaker than expected global growth. On the demand-side of the market, the major determinant for decreasing oil prices was an unexpectedly sharp deterioration in global economic activity (Baumeister and Kilian 2016). An additional effect on the demand side identified by Baffes et al. (2015) is the relatively strong appreciation of the US dollar, which makes dollar denominated crude oil imports more expensive in local currencies and thus could lead to a lower demand. However, this hypothesis is contested and the estimated impact of this effect varies between studies, e.g., Baumeister and Kilian (2016) are skeptical of any explanation based on exchange-rate movements.

In the global context of oil-producing countries, the most important decision affecting the supply of oil was an announcement by the Organization of the Petroleum Exporting Countries (OPEC) to not curtail their production in November 2014, which might have resulted in a loss of market share. Additionally, the easing of geopolitical tensions resulted in higher than expected production in the Middle East. The impact of sanctions and counter-sanctions following the conflict between Russia and Ukraine on European oil and natural gas markets was also weaker than expected (Baffes et al. 2015, 13). Another development on the supply-side was the emergence of the US shale industry, which repeatedly surprised markets by exceeding the estimates for the crude oil production and thus also put downward pressure on crude oil prices. However, the supply from these unconventional sources might be more price elastic, since they are less capital-intensive and their life-cycle is much shorter, compared to conventional oil projects (Baffes et al. 2015, 13). These characteristics and the observation of a sharp reduction in active oil rigs already led some to the conclusion that the shale oil producer in the US might have replaced Saudi-Arabia as the swing producer for the world crude oil market.[1] Although, as noted by Cakir Melek (2015), a reduction in rig count does not necessarily translate into a corresponding decline in oil production, since efficiency gains in the processes can offset these contrarily moving developments.

Baumeister and Kilian (2016) emphasize the importance of unexpected movements in oil supply. Especially, if a curtailment of the oil production is widely expected, then a positive oil supply shock leads to additional price fluctuations in the crude oil market. Accordingly, Baffes et al. (2015, 20) identify the main driver of the recent oil price drop on the supply-side of the market. The demand side-related factors that decreased the oil price had their biggest impact at the end of 2014 and thus cannot explain the prolonged period of low crude oil prices from 2015 to 2017.

It is critical to disentangle the different effects on the demand and supply side of the crude oil market in order to react accordingly. This is particularly important for central banks looking to anticipate movements in the price level and ensure financial stability. Following the great recession, quantitative easing in connection with low interest rates led to an increase of corporate loans via the risk-taking channel of monetary policy. This in turn also has implications for financial stability, since a crash in the corporate bond market in the oil producing sector could

[1] The Economist - The Economist (2015)

have severe implications for the whole financial sector.

To identify underlying mechanisms and reactions of the companies to exogenous price shocks it thus is important to analyze the relationship between oil production and debt as proposed by Domanski et al. (2015). They analyze how the buildup of debt in the oil industry following the great recession[2] and the decline in oil prices might affect the production decisions in the oil industry. This price decline mainly has two effects, it leads to lower valuation of oil companies' assets and of course reduces the cashflow of companies substantially, especially if they have not sold their production via futures contracts. In connection with the much higher debt levels in the industry this led to increased leverage and financial pressure. The oil companies can respond in two ways. They can either scale down on debt-financed investment or sell assets, which subsequently would lead to lower production in the future. Nevertheless, in order to generate enough cashflow to service their debt, oil companies could attempt to keep up the production levels or even increase them. This increases downward pressure on oil prices. It is thus particularly important to further analyze the companies' resilience and the main factors preventing the occurrence of contagious illiquidity episodes, which could jeopardize the soundness of the whole sector.[3]

Using quarterly data for over 300 companies from 2000 to 2016, this paper empirically analyzes the relationship between the financial situation of oil and gas exploration and production (E&P) companies and their production of hydrocarbons. To the best of my knowledge this is the first attempt modeling the relationship between the financial situation and the production of oil companies using detailed data on the well level. This makes it possible to disentangle the different financial conditions affecting the production decision. As the data covers both the oil price decline in 2008 and the last one in late 2014, it is possible to compare the firms' behavior in the aftermath of both events. It thus expands previous research, e.g. Lehn and Zhu (2016), by (i) using a more detailed data set and (ii) applying a different, more suitable empirical methodology, namely panel Vector Autoregressive (VAR) model.

The analysis in this paper focuses for the most part on companies active in the E&P of oil and gas. Since most companies have both oil and gas operations, it is not possible to focus solely on oil companies. Therefore, if not stated otherwise, oil industry refers to companies active in both, the E&P of oil and natural gas, hence there is no distinction made between the two different hydrocarbons. In addition, the term oil well refers to all wells for the production of oil or natural gas, no matter for which of the two they were initially drilled.

The following Section 2 reviews the literature on the impact of supply and demand shocks on the oil industry and discusses the theoretical and empirical corporate finance literature on the relationship between companies' capital structure and their performance or production decisions. In Section 3 the data set and the empirical methodology are introduced. Based on this, Section 4 presents descriptive and exploratory results as well as the results from the panel VAR approach. Section 5 concludes.

[2]International Energy Agency (IEA) (2014) provides a summary of the recent trends in energy investments.

[3]Domanski et al. (2015) focus not only on oil companies in the US, but also analyze the reactions of oil exporting countries.

# 2 Theoretical Considerations & Related Literature

## 2.1 Economics of Oil and Gas Production

In order to empirically address the hypotheses raised in the article by Domanski et al. (2015), it is necessary to first give an overview over the specific characteristics of the oil and gas E&P industry. Therefore, the following part focuses on the limitations by geological and technological boundaries and their economic implications and how this changed following the increased usage of hydraulic fracturing, commonly referred to as ‚fracking‘ and horizontal or directional drilling. These two technologies were already known in the industry for quite some time, early hydraulic fracturing for example was developed during the 1940s, although not widely used (Fitzgerald 2013).[4]

It was the discovery of unconventional reservoirs and the technological improvements to the directional drilling and fracking processes that increased the production and led to the ‚shale gas boom‘. This, of course, was also driven by the economics of relatively high natural gas prices during the early 2000s and the declining productivity of conventional US gas production, which provided an additional stimulus for the application of the novel combination of directional drilling and fracking (Rogers 2011).

These changes to the industry also have implications for the investment decisions being faced by companies, as they have increased the responsiveness of the oil supply by reducing the time lag between investment decisions being made and production. In addition, lower investment costs and a shorter life of a shale oil well reduce the problem of sunk costs and thus make it easier to lower production in response to price signals (Dale 2016). Nevertheless, the costs of the drilling and fracturing process increased during the first decade of the 2000s, since the use of more sophisticated drilling technologies makes it necessary to use more expensive rig equipment. This effect is reinforced by the fact that the well servicing industry is very concentrated and only few companies control a major share of the market. Additionally, the hydraulic stimulation of the reservoir prior to the first production adds to the drilling costs. (Fitzgerald 2013)

Gilje et al. (2017) address this implication empirically and they find that even during periods of severe contango, companies do not immediately adjust their production. Even though it would be better to curtail present production to sell it for a higher price in the future. This might be driven by sunk costs of unconventional oil wells and, in particular conventional wells, which have a longer life-cycle.

Due to these technological boundaries in the reactions of production and the irreversibility of investment decisions, the oil industry is a prime subject for empirically studying real options theory. This theory was developed to explain companies' investment decisions, when sunk costs are involved. Using drilling activities of companies, Kellogg (2014) is able to show that changes to the price volatility do impact the drilling activity and the magnitude is consistent with the optimal response postulated by the theoretical model. However, the period studied only covers the years from 1993 to 2003 and thus structural changes in the last years are not taken into account.

In an earlier paper, Hurn and Wright (1994) also apply this theory on investment decisions on North Sea oil operations and, contrary to Kellogg (2014) they conclude that, in contrast to the oil price and the level of reserves, the volatility of oil prices does not affect the time to exploitation. In related studies Dunne and Mu (2010) and Moel and Tufano (2002) empirically

[4] For a more detailed explanation on the technological details and developments, please see Fitzgerald (2013) and the references therein.

test the real options theory on mines and investment decisions of oil refineries.

A possible explanation for the non-responsiveness of oil production to changes in the oil price is offered by Anderson et al. (2018). The non-responsiveness is based on the empirical observation that over the period from 1990 to 2007 the oil production from existing oil wells in Texas was inelastic to either changes in the spot or expected future prices. The authors discover that indeed the drilling activity of companies, in contrast to production, is highly correlated with oil prices. Therefore, the authors use Hotelling's (1931) model of exhaustible resource extraction and reformulate it as a drilling problem, since the companies can decide when to drill, but cannot influence the reservoir pressure and thus production. Although, after 2007 the production from unconventional sources increased considerably and this probably made the supply more elastic to changes in prices. In connection with the Hotelling principle, Thompson (2001) analyzes the impact of backwardation in non-renewable resource markets and shows that oil companies face two decisions. First, they need to decide on the investment in the production capacity and subsequently need to determine the level of production.

Gilje et al. (2017) also address the hypotheses by Domanski et al. (2015) and empirically analyze the relationship between companies' drilling decisions and their leverage. Using detailed project-level data, they are able to show that highly leveraged firms tend to move forward with project completion, even though it would have been more profitable to protract the completion during contango periods. One explanation for this can be found in the decision of equity holders to sacrifice long-term returns in order to enhance collateral in the short term, because this behavior is more pronounced just before debt renegotiations.

Moreover, Lehn and Zhu (2016) show that the price decline affects oil companies differently, according to their leverage. Their results indicate that highly leveraged companies reduce their investments and at the same time increase the production from existing investments. The focus of their paper is only on the period from 2011 to 2015 and thus only includes the most recent decline in crude oil prices. The present paper is closely related to the two studies last mentioned and thus builds on their research, extending their analyses and methodologies.

## 2.2 Relationship between Financial Situation and Production Decisions

Moving away from the literature on the decision making process and the distinctive characteristics of companies' investments in the E&P sector, it is important to provide an overview on the determinants of the structure of the liability side of the balance sheet of companies and how the debt level and investments affect the production decision.

Frank and Goyal (2008) give a comprehensive overview on different theories on the determinants of debt financing, which can be subsumed under the two umbrella terms *trade-off* and *pecking order* theory. The trade-off theory assumes that a companies' decision maker needs to balance the trade-off between the tax benefits of debt and the dead-weight costs of bankruptcy to reach an optimal level of leverage. This balancing leads to a target leverage ratio and deviations from this target are gradually eliminated over time. [5]

The pecking order theory hypothesizes that firms prefer internal over external finance and if external finance is used, then it prefers debt to equity. Frank and Goyal (2008) provide a summary on the motivation of this theory based on the adverse selection and the agency theory behind it.

These same authors empirically examine different factors, which affect capital structure decisions of companies. Besides company-specific factors, they also identify industry-specific

[5]For a detailed discussion on the differences of static and dynamic trade-off theory and the empirical research, please see Frank and Goyal (2008).

ones, which are relevant for an empirical study (Frank and Goyal 2009). These factors and their effect on leverage are median industry leverage (+), market-to-book assets ratio (-), tangibility (+), profits (-), log assets (+), and expected inflation (+). Kayhan and Titman (2007) find empirical evidence for the trade-off theory and that additional variables might affect the determination of the leverage ratio.

The decision on how much to produce is of course not only influenced by the capital structure of the company, but it is even more closely related to a companies' investment decisions, especially past ones. Therefore, it is important to identify factors influencing the level of investment. One contested variable is the level of cashflow: there are a series of papers from two groups of authors arguing over the importance and implications of the relationship between cashflow levels and investment (Fazzari, Hubbard, and Petersen 2000; Fazzari, Hubbard, Petersen, et al. 1988; Kaplan and Zingales 1997, 2000).

Another strand of literature studies the relationship of market structure, capital structure and the output decision of a company. These studies show that the structure of the product market and the capital structure of a company influence its output decision. In this literature an important factor is the limited liability effect of debt, which basically creates an incentive for the equity holder to only use debt financing for investments (Brander and Lewis 1986; Phillips 1995). Fosu (2013) also focuses on the relationship between leverage and the degree of competition within an industry and shows that leverage increases with higher competition.

On an aggregate level there is another important factor which increased the debt-level in the energy sector, namely the quantitative easing of the Federal Reserve Bank in the US. This risk-taking channel of the monetary policy in connection with the relatively high oil prices contributed to increased capital flows into the energy sector and the corporate bond market, please see Borio and Zhu (2012), Delis et al. (2017), and Dell'Ariccia et al. (2017)

# 3 Empirical Analysis Framework

## 3.1 Combining the Data Set

An analysis of the relationship between the financial conditions of companies and their production decision requires not only financial data, but also detailed data on their production. This made it inevitable to compile the data set from two distinct data sources, since all available data sets were not sufficient for an in depth analysis of this topic.

The quarterly financial data is taken from the CapitalIQ database and covers all companies headquartered in the US or Canada falling under the Standard Industrial Classification (SIC) code 1311, which includes companies primarily engaged in the exploration of oil and gas field properties. The selection of this quite narrow definition is done to solely focus on the relationship between the financial situation and the production decision.

Since detailed production data is not provided in the CapitalIQ database, the data on oil production is taken from an industry-specific database provided by Enverus for the period from 2000 to 2016[6]. This database has the advantage that it includes not only the base data of the oil well, but also detailed production data for oil, natural gas and water. The base data of an oil well consists of information on the location, like basin, reservoir, formation and field, and political subdivisions like state and county. Additionally, in most cases it also includes the drilling technology, which allows to differentiate between directionally, horizontally, and

[6] DrillingInfo is a private company based in Austin, Texas providing detailed oil industry data. Please see http://www.enverus.com for more information.

vertically drilled wells. This allows to analyze the impact of new technologies and their firm specific effects.

The combination of the two data sets is achieved by using a hybrid matching approach, initially using R (R Core Team 2018) in connection with the *stringdist* package developed by Loo (2014) to automatically generate matches based on the similarity of companies' names. In the next step, each match is manually checked using additional base data on the companies. In all cases, where a match could not be completely verified by a manual check, the data was discarded and not included in the final data set. This procedure resulted in an unbalanced quarterly data set covering the period from Q1 2000 to Q2 2016 and consisting of 339 different companies. From the initially 153 companies in Q1 2000, 53 are present throughout the whole sample period, while 186 companies enter into the sample after the start of the sample period. Together with the 170 companies dropping out of the sample, this results on average in around 145 companies per quarter. Even though there is quite some fluctuation in the data set, the average duration of a company in the sample is marginally above 27 quarters or nearly seven years. In Figure 1 the reasons for a company dropping out of the sample are shown over time. It can be seen that Acquisitions & Mergers with a total 102 companies are by far the main reason for a company to drop out of the sample. Over the horizon of the analysis only five companies filed for bankruptcy and only two companies were liquidated.[7] Interestingly, these events occur shortly after the collapse of the crude oil prices in 2008 and 2014. Apparently, these numbers understate the overall numbers of bankruptcies and liquidations in the E&P industry following the oil price decline, since the „Oil Patch Bankruptcy Monitor“ by Haynes and Boone, LLP (2017) already lists 44 bankruptcy filings for 2015 and 70 for the whole of 2016.

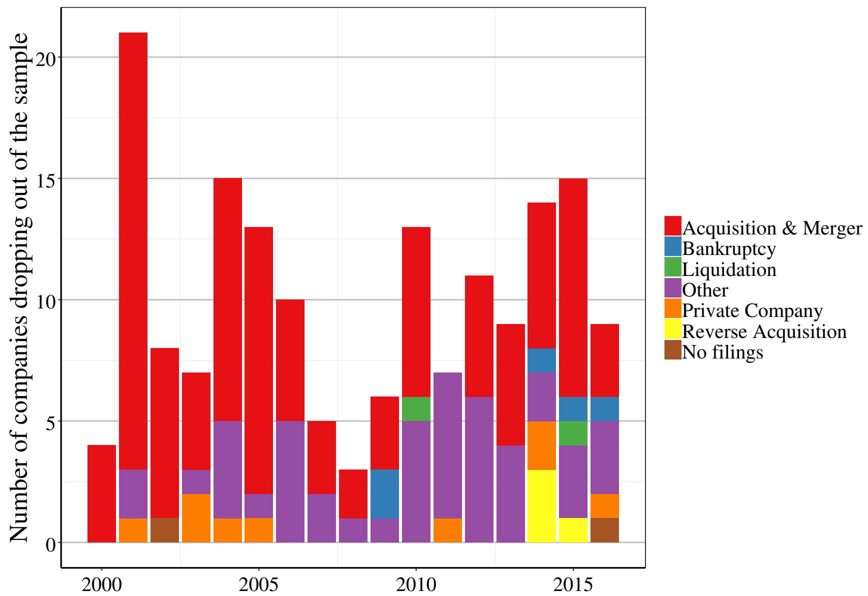



Figure 1: Number of companies dropping out of the sample per year and the respective reason.

In order to analyze companies' behavior, price time series for crude oil and natural gas are included in the empirical analysis. In case of crude oil, the spot price of West Texas Intermediate (WTI) measured at Cushing, Oklahoma in $ per Barrel (bbl) is used. This is the benchmark for crude oil in the continental US. In case of natural gas, this role is fulfilled by the Henry Hub

[7] For the rest of the companies the reason for dropping out of the sample are given by: Other (46), Going private (9), Reverse takeover (4) or no more fundamental filings (2).

distribution point in Erath, Louisiana, which is reported in $ per million British thermal units (mmBtus).

To assess the extent of contango or backwardation in both markets, New York Mercantile Exchange (NYMEX) futures prices for delivery in the four consecutive months following the trade date are included. All price time series are obtained from the U.S. Energy Information Administration (EIA).

## 3.2 Empirical Methodology

The empirical analysis of the relationship between debt and the production of fossil fuels faces several challenges, of which endogeneity, inherent in most corporate finance data sets, is the most important one. Roberts and Whited (2013) provide a comprehensive overview on the causes of endogeneity and how these can be overcome. Two common problems that arise from corporate finance data can be summarized as measurement errors and endogeneity and simultaneity (Roberts and Whited 2013). The latter two can be addressed using a panel VAR model. This methodology additionally offers the possibility to explicitly account for the persistence observable in corporate financial data. Since panel VAR models incorporate lagged endogenous variables, the estimated coefficients suffer from Nickell bias and thus, it is necessary to use generalized method of moments (GMM) techniques to estimate these models (Nickell 1981). Holtz-Eakin et al. (1988) were the first to apply the well-established VAR techniques to panel data. Especially the improvements to GMM estimation of single equation dynamic panel data models by Arellano and Bond (1991), Arellano and Bover (1995), and Blundell and Bond (1998) also influenced the panel VAR estimation techniques. Binder et al. (2005) is one of the few theoretical papers solely concerned with the estimation of panel data VAR models. All these improvements to the panel VAR methodology led to an increase in empirical applications of which Love and Zicchino (2006) and later Abrigo and Love (2016) are probably most important, since they also provided code to apply their methodology.

Following the discussion of different challenges the empirical modelling of this research is facing, the panel VAR methods by Sigmund and Ferstl (2019) are employed, since they offer the possibility to apply the latest estimation techniques and additional tools to visualize the relationship between the variables. To remove the unobserved individual effects in a dynamic panel data model there exist two different approaches, taking first differences or the calculation of the forward orthogonal deviations. Arellano and Bover (1995) show that the GMM estimator is not affected by the transformation chosen to remove the individual effects. Although, these results only hold if the transformation matrix is upper triangular and all the available instruments are used. In empirical applications, these conditions are rarely met, since including too many instruments deteriorates the finite sample behavior and thus might bias the GMM estimator. Therefore, the choice of the transformation is vital and Hayakawa (2009) uses a simulation study to compare the performance of the two transformations. This is also supported by the result of the Monte Carlo simulation by Phillips (2019). Since both of these results indicate that forward orthogonal deviation performs better in those cases most similar to the data set at hand, the forward orthogonal deviations is used in this study to remove the unobserved individual effects.

To assess the performance of various estimation techniques developed to counteract the biases introduced in dynamic panel data, Flannery and Hankins (2013) create simulated corporate finance data. They are trying to include all data related issues, normally observed in such data, like missing, correlated or endogenous independent variables. Based on these results they can show that the best estimation technique strongly depends on the issues present in the data, although it seems that the estimation technique developed by Blundell and Bond (2000) appears to be best in most cases. However, one has to keep in mind that the application of GMM estimation techniques might lead to the issue of too many instruments (Roodman 2009).

The estimated panel VAR is specified, following the notation of Sigmund and Ferstl (2019), as:

$$\mathbf{y}_{i,t} = (\mathbf{I}_m - \sum_{l=1}^{p} \mathbf{A}_l)\mu_i + \sum_{l=1}^{p} \mathbf{A}_l \mathbf{y}_{i,t-l} + \mathbf{B}\mathbf{x}_{i,t} + \mathbf{C}\mathbf{s}_{i,t} + \varepsilon_{i,t} \tag{1}$$

The vector of endogenous variables $\mathbf{y}_{i,t}$ includes the logarithmized variables of total assets as a measure of company size, the leverage ratio and the quarterly oil production, $i$ and $t$ indicate the company and time respectively. Since oil price in this specification is assumed to be exogenous the natural logarithm of the last available quarterly WTI oil price is included in the vector of exogenous variables $\mathbf{x}_{i,t}$. The vector $\mathbf{s}_{i,t}$ would cover strictly exogenous variables, however these are not present in this application.

The estimation includes four lags of the endogenous variables in order to incorporate the seasonality and autocorrelation of the quarterly observations.

To analyze the relationship between different variables in more details it might be interesting to test for granger causality between different variables, using the approach described by Dumitrescu and Hurlin (2012). Another option would be to use a difference-in-difference approach like Gilje et al. (2017), with two different treatments. The first treatment is high and low leverage and the second treatment is the occurrence of contango or backwardation. The implementation of this methodology would of course allow the comparison of the results and determine if there are differences between the decision of drilling new oil wells and the level of production.

It is important to complement the empirical analysis with some robustness checks to make sure that the results are not statistical artifacts. Especially, since Frank and Goyal (2008) highlight the problems associated with using book leverage and its implications for econometric modelling. Additionally, in early empirical work Titman and Wessels (1988) found evidence that leverage varies with the companies' size.

# 4 Empirical Analysis Results

## 4.1 Exploratory Data Analysis

This section summarizes the data set and highlights various aspects, which are already offering interesting insights and are important for the subsequent empirical analysis as well. In order to examine the validity of the constructed data set, the aggregate crude oil production of the individual companies in the data set is compared to official data on the total crude oil production in the US.

Figure 2 depicts the development of US crude oil production. It shows that the observable increase in total crude oil production, starting in 2008, is mainly driven by the increased production from unconventional sources. In order to provide further evidence for the validity of the company level data set, Figure 3 is based on the aggregated production data and shows the total volume of crude oil differentiated across the different drilling technologies used in the production. However, the aggregate volume in the sample comprises between 20% and 38% of the total production in the US[8], the overall development of the oil production, especially the

[8] The share ranges from 22% in Q2 2002 to 38% in Q1 2015, although for most quarters after 2008 the share is

increase after 2008, is well represented in the company level data.

In addition, when looking at the different technologies and the development of their production volume over time, it is apparent that the production from horizontally and directionally drilled oil wells can be used as a proxy for production from unconventional sources. In particular, the rise in oil production at the company level can be attributed to increasing crude oil production from horizontally drilled oil wells. The strong visual conformity between the two oil production time series from unconventional sources is also underpinned by a really strong correlation of 0.9913. Additionally, when comparing the oil production of the sample with the total US oil production in Figure 4, it is apparent that the oil production in the sample comprises a considerable share of the total US oil production, the relative share over the whole horizon varies between 22% and 37%.

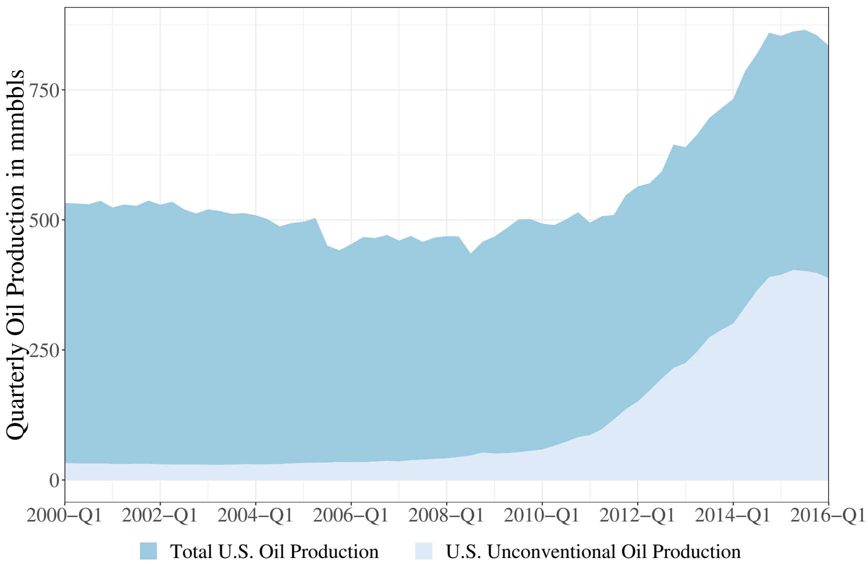



Figure 2: Development of Conventional and Unconventional US Oil Production.
Source: Crude oil production (EIA 2017a) and tight oil production estimates (EIA 2017c))

The development of well productivity is depicted in Figure 5. Starting in 2009, the productivity of unconventional wells starts to increase, while the productivity from conventional wells over the same time period is decreasing. This also is in line with Roll and Dahl (2017), who show that the main driver of productivity growth in the oil sector were unconventional sources and the technologies used to develop them.

The development of WTI crude oil and natural gas prices for the US is displayed in Figure 6. The main difference between the two price time series is that, unlike the price for crude oil, the price for natural gas does not quickly recover following the price decline in 2008. The differences in trajectory of the price time series is also expressed by the diverging development of the contango following 2008's peak in high prices. The Henry Hub natural gas spot price is in contango until 2013. So during these periods, the futures prices were higher than the spot prices, which provides an incentive to curtail production to exploit resources at a later point in time. This incentive was much greater in the case of natural gas, since the periods of contango were much longer and the price did not recover as much as in the case of crude oil.

The observable periods of contango and backwardation are similar to those studied by Gilje

above 30%.

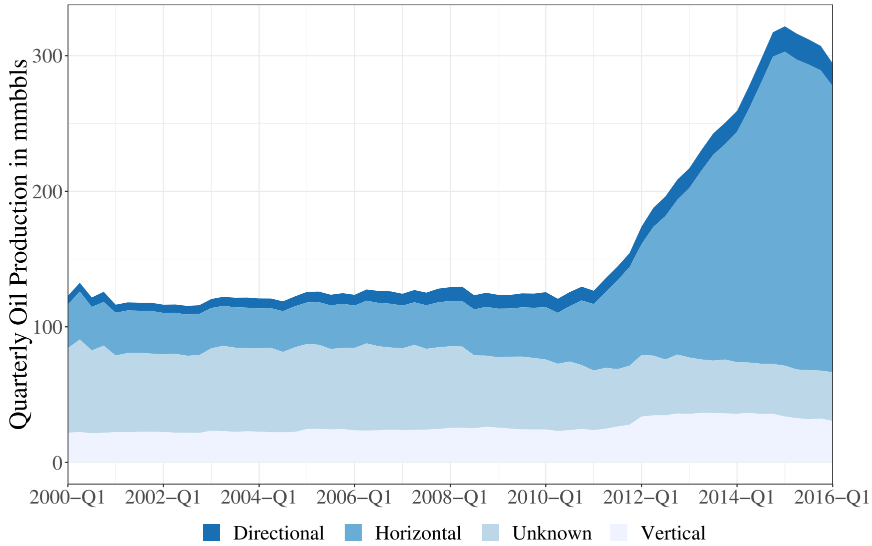



Figure 3: Development of aggregated oil production for different drilling technologies in the sample used in this analysis.
Source: Own calculations based on data provided by DrillingInfo

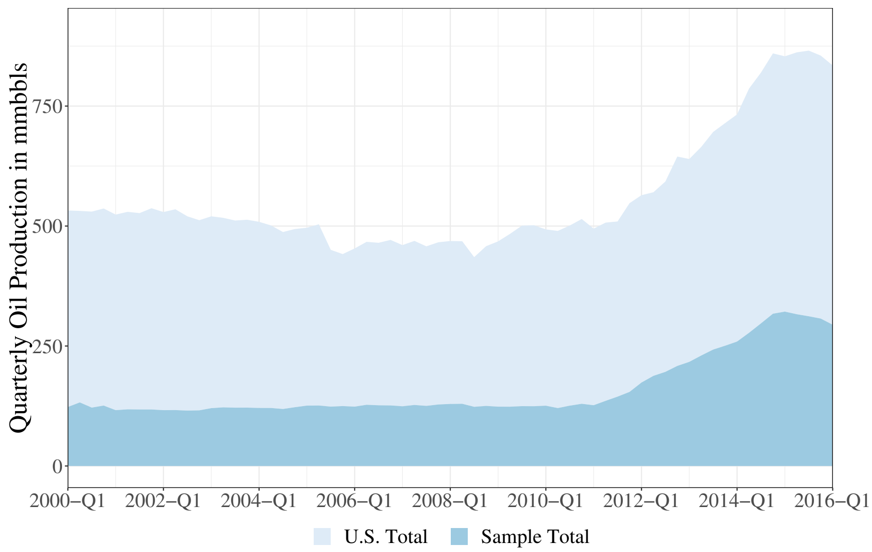



Figure 4: Total US oil production and oil production in sample
Source: Own calculations based on data provided by DrillingInfo and EIA.

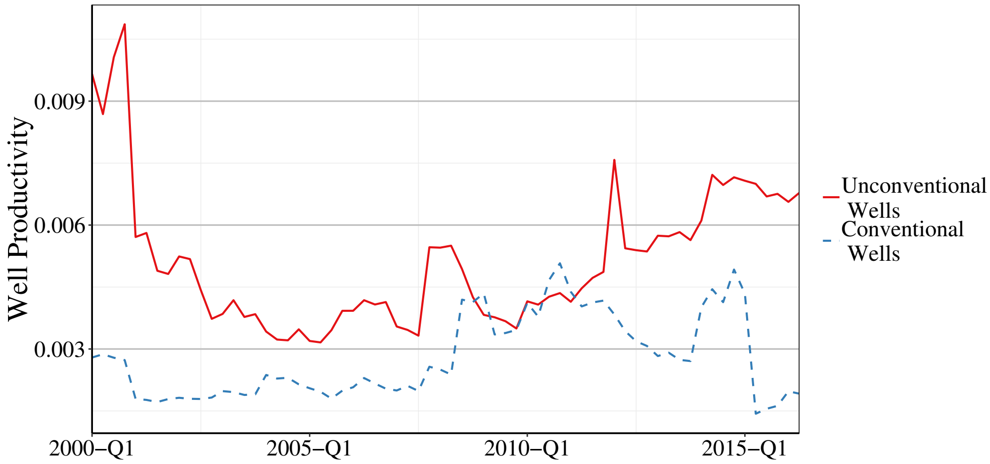



Figure 5: Development of oil well productivity differentiated by conventional and unconventional drilling technology.
Source: Own calculations based on data provided by DrillingInfo

et al. (2017), although the actual numbers and the extent of contango differ because of different time horizons of the future contracts used in the calculation.

The diverging trajectories of the two fossil fuel prices are especially interesting, since they allow us to distinguish between firms' reactions to these two different price changes. Especially, it is interesting to analyze the reaction of companies to the protracted period of lower prices in the natural gas market starting in 2008. This episode likely provides insights into the response of the companies to the period of lower crude oil prices following the decline in the second half of 2014.

In this analysis, leverage is based on the book value and defined as the sum of the total long term debt and debt in current liabilities divided by the total value of assets, so basically it is the debt-to-asset ratio of a company. In Figure 7, the development of average leverage across all companies in the sample is depicted. Beginning in 2000, the leverage decreases until reaching the lowest point in the third quarter of 2005. After a peak of nearly 0.3 during the great recession, it again falls until in 2011, when it starts to increase again and in 2016 reaches the level of 0.35, previously only seen at the start of the 2000s. The development of leverage in this sector also reflects the impact of the risk-taking channel, since the increase in leverage is mostly due to increasing levels of debt and not solely caused by deteriorating asset valuations over this horizon.

To analyze the impact leverage might have on production and the adoption of new technologies there is further empirical evidence provided in Section App.A of the online appendix. It is observable that the level of leverage does not negatively affect the adoption of new technologies.

Using a different measure for the indebtedness of companies, namely the ratio of debt to earnings before interest, taxes, depreciation, and amortization (EBITDA) the severeness of the price declines in 2008 and 2014 and their impact on companies is evident. The development of the ratio is depicted in Figure 8. Nevertheless, the companies are able to stabilize their income and return to a positive EBITDA relatively quick after the decline in 2008. The visual inspection of Figure 8 indicates that the level of the ratio is higher after the recovery following the great recession. This is also confirmed by a comparison of the median values of the ratio. In the period after the great recession from Q4 2008 to Q3 2014 the median value of the ratio

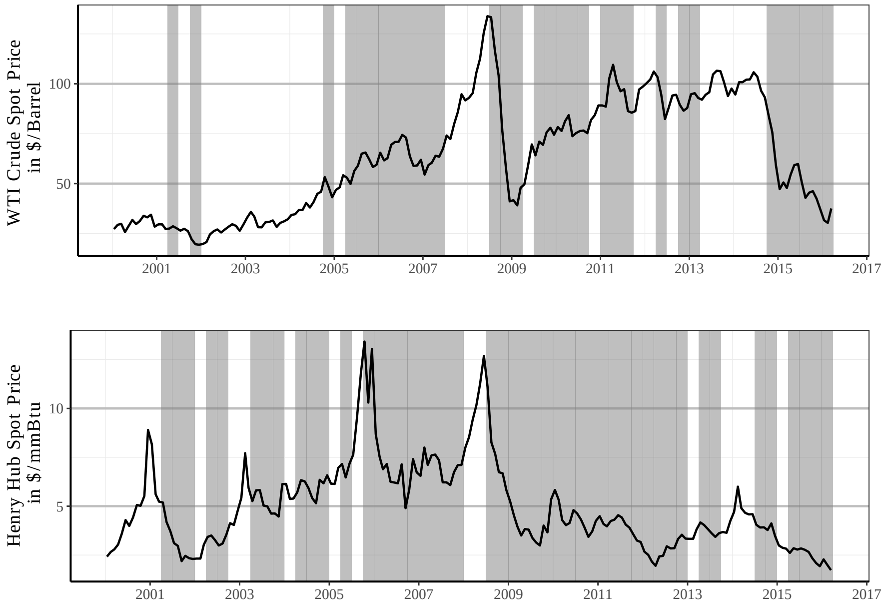



Figure 6: Development of WTI crude oil and Henry Hub natural gas spot prices. Shaded areas indicate quarters during which the futures prices were higher than the spot price.
Data source: WTI price time series (EIA 2017d) and Henry Hub Natural Gas price time series (EIA 2017b)

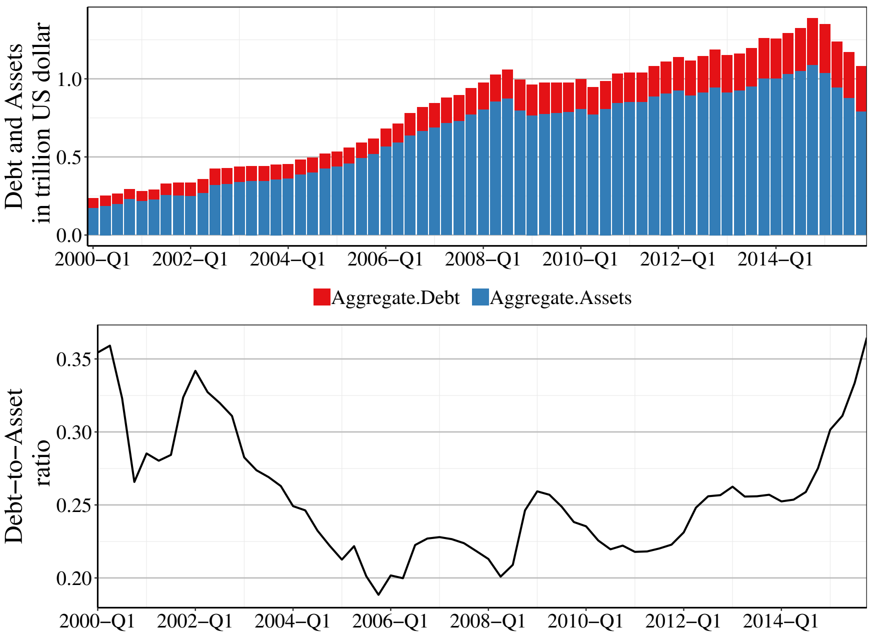



Figure 7: Development of the aggregate debt and asset level and the associated average debt-to-asset ratio across all companies.
Source: Own calculations based on data provided by Compustat

is 6.21 whereas before it is only 4.52.

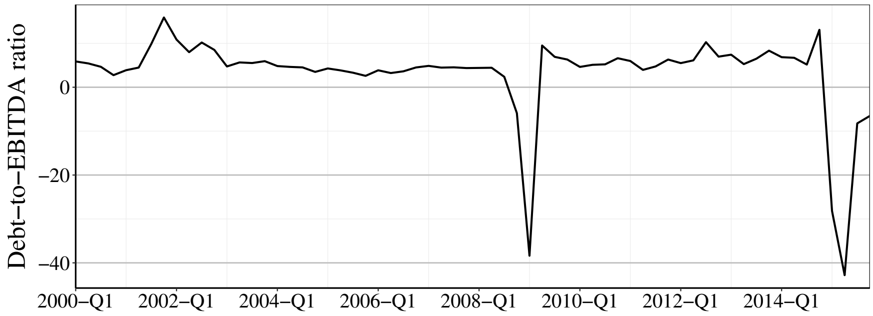



Figure 8: Development of the average debt to EBITDA ratio across all companies.
Source: Own calculations based on data provided by Compustat

| | Mean | Median | Std. Dev. | MAD | Min | Max |
|---|---|---|---|---|---|---|
| Quarterly Oil Production | 1.22 | 0.06 | 3.44 | 0.08 | 0 | 32 |
| Quarterly Gas Production | 15 074.74 | 926.46 | 40 594.10 | 1373.57 | 0 | 554 792 |
| Leverage | 0.36 | 0.27 | 0.97 | 0.19 | 0 | 40 |
| Debt-to-EBITDA ratio | 5.00 | 4.11 | 191.78 | 6.09 | −9703.227 | 4514 |
| Assets | 4893.41 | 644.35 | 13 971.61 | 927.54 | 0.007 | 190 155 |
| Debt | 1256.40 | 206.62 | 3074.79 | 305.84 | 0 | 35 707 |
| WTI Spot Price | 64.91 | 65.94 | 28.86 | 42.71 | 19.960 | 140 |
| HHUB Spot Price | 4.94 | 4.29 | 2.35 | 1.93 | 1.730 | 13 |

Table 1: Descriptive statistics for main variables used in the analysis. Oil Production is measured in million Barrels (mmBbls) per quarter, gas production in mmBtus and the financial data is reported in million US dollar.

It is important to note that small companies in this sample actually are quite large, since a lot of small companies are not publicly listed (Bond et al. 2004, 24). This can be seen in Table 1, since the mean value of a companies' assets is nearly five billion US dollars and a median value indicating that 50% of the companies have more than 644 million US dollars in assets. This highlights a possible selection bias and creates additional problems in connection with the survivorship bias, because only surviving companies are present over the whole sample period. However, it is possible to address this question in more detail and determine the factors which influence the probability of a company dropping out of the sample.

## 4.2 Results of the Panel VAR

Table 2 shows the estimated coefficients for each of the three endogenous variables. The results are based on a total of 8373 company quarter observations, which are made up of 330 different companies; this means that, on average, there are 25.37 quarters per company in the sample. The results show that there are only minor interdependencies and the variables are mainly affected by lagged variables of their own. The only statistically significant effect of the leverage ratio on the oil production can be observed with a lag of four quarters. It implies that a higher leverage ratio lowers the oil production in subsequent quarters.

The exogenously modeled oil price has the theoretically expected impact on the assets and the oil production, namely that both assets and oil production increase in response to an increasing oil price. The price elasticity of the oil production has a value of 0.1236.

| | Dependent Variables | | |
|---|---|---|---|
| | Leverage | log.Assets.Total | log.DI.Oil.Prod.Total.Sum |
| $\text{Leverage}_{t-1}$ | 0.6510*** | 0.0044 | −0.0409** |
| | (0.1895) | (0.0075) | (0.0208) |
| $\text{log.Assets.Total}_{t-1}$ | −0.1108*** | 1.0231*** | −0.0023 |
| | (0.0308) | (0.0413) | (0.0660) |
| $\text{log.DI.Oil.Prod.Total.Sum}_{t-1}$ | 0.0064* | −0.0014 | 0.5677*** |
| | (0.0038) | (0.0034) | (0.0842) |
| $\text{Leverage}_{t-2}$ | 0.6125 | −0.0217* | 0.0309 |
| | (0.7097) | (0.0115) | (0.0224) |
| $\text{log.Assets.Total}_{t-2}$ | 0.0851 | 0.1590*** | −0.0487 |
| | (0.2862) | (0.0375) | (0.1763) |
| $\text{log.DI.Oil.Prod.Total.Sum}_{t-2}$ | 0.0028 | −0.0044 | 0.0549 |
| | (0.0067) | (0.0058) | (0.0411) |
| $\text{Leverage}_{t-3}$ | −0.3166 | −0.0031 | 0.0404 |
| | (0.4607) | (0.0115) | (0.0312) |
| $\text{log.Assets.Total}_{t-3}$ | 0.3068 | −0.2863*** | −0.0198 |
| | (0.4685) | (0.1068) | (0.2162) |
| $\text{log.DI.Oil.Prod.Total.Sum}_{t-3}$ | 0.0058 | 0.0031 | 0.1935 |
| | (0.0085) | (0.0087) | (0.1198) |
| $\text{Leverage}_{t-4}$ | 0.1956*** | −0.0192 | −0.0700* |
| | (0.0745) | (0.0218) | (0.0364) |
| $\text{log.Assets.Total}_{t-4}$ | −0.2563 | 0.0556 | 0.0681 |
| | (0.2133) | (0.0698) | (0.1475) |
| $\text{log.DI.Oil.Prod.Total.Sum}_{t-4}$ | −0.0084 | 0.0008 | 0.0855 |
| | (0.0091) | (0.0085) | (0.1174) |
| log.Last.Quarterly.WTI.Spot.Price | −0.0383 | 0.0630*** | 0.1236*** |
| | (0.0253) | (0.0104) | (0.0423) |
| const | −0.0309 | 0.0884* | −0.7077** |
| | (0.0792) | (0.0471) | (0.2824) |
| Observations | 8373 | | |
| Number of Groups | 330 | | |
| Avg. Obs. Group | 25.37 | | |
| Min. Obs. Group | 1 | | |
| Max. Obs. Group | 66 | | |

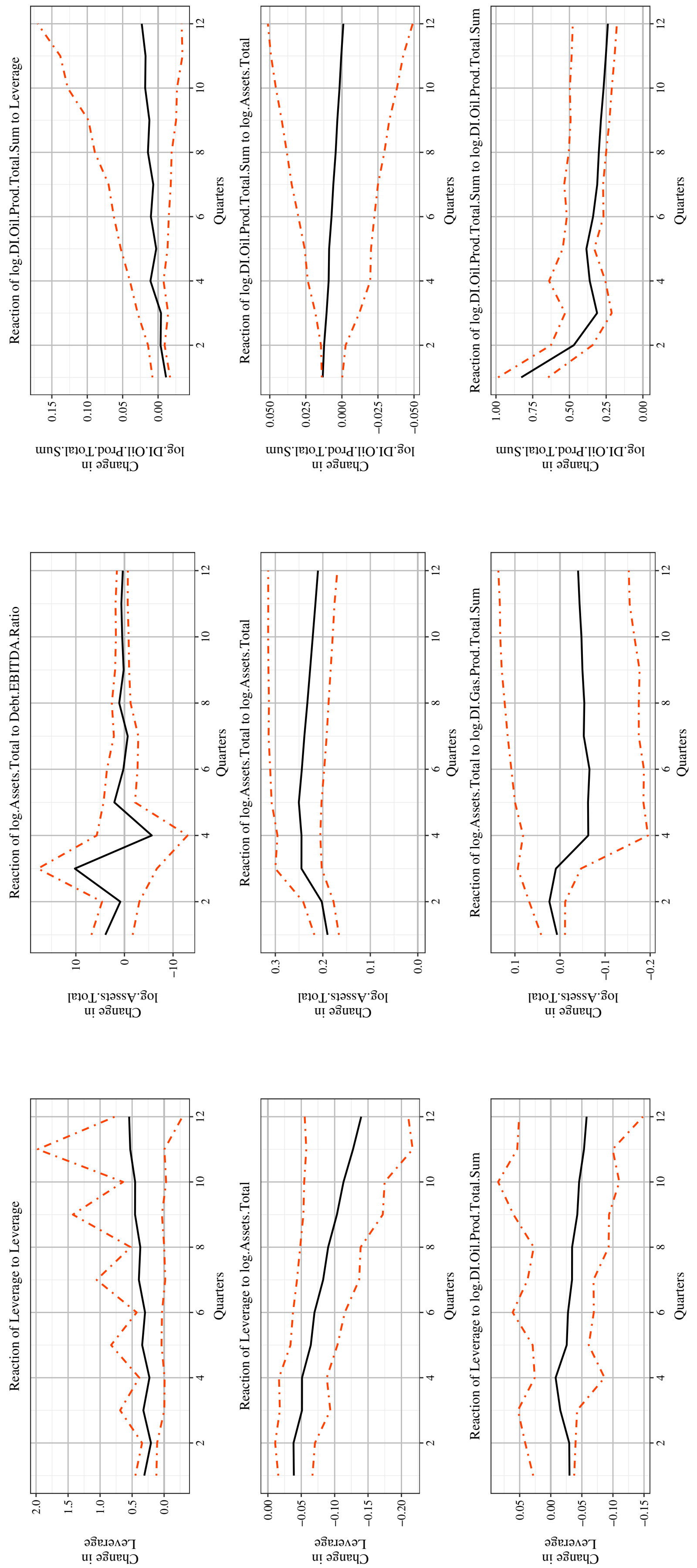
*Note:* *p<0.1; **p<0.05; ***p<0.01
Corrected standard errors are reported in parentheses.
Variable transformation: Forward Orthogonal Deviation

Table 2: Results of the panel VAR approach for the oil production.

In the context of VAR models, the preferred way of analyzing the relationship and interdependences between variables is the calculation of impulse response function (IRF). Since it is difficult to come up with theoretical assumptions on the contemporaneous effects, instead of orthogonalized IRF, the generalized IRF introduced by Pesaran and Shin (1998) are calculated for 12 quarters and are depicted in Figures 9a, 9b and 9c. Each of the three figures depicts the reaction of one dependent variable to shocks in one of the three endogenous variables. The bootstrapped 95% confidence intervals are calculated using the procedure by Kapetanios (2008). In Figure 9a it can be observed that in reaction to a shock the leverage ratio does not return to its equilibrium value. This is in contrast to the IRF of the debt to EBITDA ratio, which is shown in Figure App.10a it can be seen that in reaction to a shock of itself the debt to EBITDA ratio returns to its equilibrium value very quickly. This is also in line with the observations of the average debt to EBITDA and the leverage ratio in Figures 8 and 7, where it is obvious that over the horizon of this analysis the debt to EBITDA ratio mostly remained fairly constant and only deviated strongly during the extreme price declines in 2008 and 2014, whereas the average leverage ratio is less constant over time.

In the case of a shock to either assets or the oil production the leverage ratio, after a short period of adjustment, is moving to a lower level. This is also in line with the theoretical consideration that an increase in assets and oil production should, in the medium term, increase assets and thus also lower it relative to the debt of company. Although only the reaction to a shock in assets is statistically different from zero. The reaction of assets to shocks in the other endogenous variables are shown in Figure 9b. Although, assets fluctuate quite strongly in response to a shock of the leverage ratio the effect dies down rather quick. A positive shock of assets leads to a persistent increase over the course of the 12 quarters analyzed. Interestingly a shock increasing the oil production leads to an immediate increase in assets, although the change is not persistent and not significantly different from zero. In case of the oil production, depicted in Figure 9c, it can be observed that both a shock to the leverage ratio and to the level of assets does not really have any impact on the oil production. A positive shock to oil production is only reversed slowly, although a reversion to the equilibrium appears to happen.

Further, the same analysis was conducted using an alternative measure for indebtedness, namely the debt to EBITDA ratio, and the results are reported in section App.E of the online appendix. To shed further light on the potential determinants of production decisions, the sample is divided into subsamples and it is analyzed how determinants vary across subsamples. Specifically, the analysis focuses on companies' variations in (1) leverage, (2) share of unconventional production. The result for this analysis are provided in Section App.G of the online appendix.




(a) Impact on leverage

(b) Impact on total assets

(c) Impact on oil production.

Figure 9: Generalized IRF for the whole sample. Dashed red lines indicate the bootstrapped 95% confidence intervals based on 500 iterations.
Source: Own calculations

# 5 Concluding Remarks and Outlook

This paper analyzes the relationship between the leverage of companies and their production decision using a novel data set. In the first part of the paper the theoretical background on the economics of the crude oil and natural gas production and the possible connection to the financial situation of a company is provided. Additionally, the possible problems a researcher faces, when empirically analyzing financial data of companies over time are discussed. The novel data set is then described and the relatively new empirical methodology of a panel VAR is introduced. The exploratory data analysis, besides other interesting insights is able to show that the data set on the company level is capable of describing a sizable part of the domestic crude oil production in the US. Using the panel VAR approach to analyze the data set it is possible to disentangle the relationship between the endogenous variables and the impact the oil price has on the production decisions of oil producing companies.

To exploit the information in the data set still further, the sample is then divided into subsamples. In a first step the companies are divided into companies with a low and high level of leverage, to see if the interdependencies of the variable changes. To analyze the impact of unconventional production technologies, like directional drilling or hydraulic fracturing, and if this might have changed the economics of oil and gas exploration, the companies are, in a second step, also differentiated according to the share of unconventional oil on their total oil production. It can be shown that especially for companies with a high leverage and a high share of production from unconventional sources the price of oil has a much bigger impact, since in these two subsamples the price elasticity is much higher, than for companies with low leverage and a smaller share of production from unconventional sources. These results lend further support to the hypothesis that especially the shale oil producing companies might be able to provide flexible oil production capacity. The detailed results for this can be found in the online appendix.

Additionally, it might be interesting to check if a differentiation according to company age instead of company size might yield interesting results, as discussed in Fort et al. (2013). Furthermore, it would be probably worthwhile to extent the horizon of the analysis into more recent quarters, since the resilience probably weakens the longer the prices stay at lower levels. In an extension of this research it might be beneficial to check if during periods of divergence between the price for WTI and Brent oil the reaction of oil producers might have changed. Especially, since for oil producers in the US the WTI price might have become endogenous, during the periods when transporting capacity was insufficient. Kilian (2016) provides additional information on why the price for WTI became decoupled from Brent and why this was mainly due to domestic developments in the US.

The results of the panel VAR approach look promising although they cannot really lend their support to the hypothesis of Domanski et al. (2015) that high levels of debt or leverage might be responsible for the observable resilience of the oil producer in the US. Nevertheless, it has to be noted that the data set created for this analysis might not be perfect and thus the conclusions based on this data need to be taken with a grain of salt. Additionally, the currently employed empirical methodology lacks a rigorous testing for the stationarity properties of the data series and the properties of the estimation residuals. These shortcomings need to be addressed in future research. It would also be interesting to see if the analysis could replicate results from the corporate finance literature in order to increase the validity of the results obtained herein.

# References


Abrigo, M. R. and I. Love (2016): "Estimation of panel vector autoregression in Stata." In: *The Stata Journal* 16 (3), pp. 778–804.

Anderson, S. T., R. Kellogg, and S. W. Salant (2018): "Hotelling under Pressure." In: *Journal of Political Economy* 126 (3), pp. 984–1026. DOI: 10.1086/697203.

Arellano, M. and S. Bond (1991): "Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations." In: *The Review of Economic Studies* 58 (2), pp. 277–297.

Arellano, M. and O. Bover (1995): "Another look at the instrumental variable estimation of error-components models." In: *Journal of Econometrics* 68 (1), pp. 29–51. DOI: 10.1016/0304-4076(94)01642-D.

Baffes, J., M. A. Kose, F. Ohnsorge, and M. Stocker (2015): "The Great Plunge in Oil Prices: Causes, Consequences, and Policy Responses." In: *SSRN Electronic Journal*. DOI: 10.2139/ssrn.2624398.

Baumeister, C. and L. Kilian (2016): "Understanding the Decline in the Price of Oil since June 2014." In: *Journal of the Association of Environmental and Resource Economists* 3 (1), pp. 131–158. DOI: 10.1086/684160.

Binder, M., C. Hsiao, and M. H. Pesaran (2005): "Estimation and inference in short panel vector autoregressions with unit roots and cointegration." In: *Econometric Theory* 21 (4), pp. 795–837. DOI: 10.1017/s0266466605050413.

Blundell, R. and S. Bond (1998): "Initial Conditions and Moment Restrictions in Dynamic Panel Data Models." In: *Journal of Econometrics* 87 (1), pp. 115–143. DOI: 10.1016/S0304-4076(98)00009-8.

– (2000): "GMM Estimation with persistent panel data: an application to production functions." In: *Econometric Reviews* 19 (3), pp. 321–340. DOI: 10.1080/07474930008800475.

Bond, S., A. Klemm, R. Newton-Smith, M. Syed, and G. W. Vlieghe (2004): *The Roles of Expected Profitability, Tobin's Q and Cash Flow in Econometric Models of Company Investment*. SSRN Scholarly Paper ID 641241. Rochester, NY: Social Science Research Network. DOI: 10.2139/ssrn.641241.

Borio, C. and H. Zhu (2012): "Capital regulation, risk-taking and monetary policy: A missing link in the transmission mechanism?" In: *Journal of Financial Stability* 8 (4), pp. 236–251. DOI: 10.1016/j.jfs.2011.12.003.

Brander, J. A. and T. R. Lewis (1986): "Oligopoly and Financial Structure: The Limited Liability Effect." In: *American Economic Review* 76 (5), pp. 956–970.

Cakir Melek, N. (2015): "What could lower prices mean for US oil production?" In: *Economic Review - Federal Reserve Bank of Kansas City*, pp. 51–69.

Dale, S. (2016): "New Economics of Oil." In: *ONE J: Oil and Gas, Natural Resources, and Energy Journal* 1, pp. 365–379.

Delis, M. D., I. Hasan, and N. Mylonidis (2017): "The Risk-Taking Channel of Monetary Policy in the U.S.: Evidence from Corporate Loan Data." In: *Journal of Money, Credit and Banking* 49 (1), pp. 187–213. DOI: 10.1111/jmcb.12372.

Dell'Ariccia, G., L. Laeven, and G. A. Suarez (2017): "Bank Leverage and Monetary Policy's Risk-Taking Channel: Evidence from the United States." In: *The Journal of Finance* 72 (2), pp. 613–654. DOI: 10.1111/jofi.12467.

Domanski, D., J. Kearns, M. J. Lombardi, and H. S. Shin (2015): "Oil and Debt." In: *BIS Quarterly Review*.

Dumitrescu, E.-I. and C. Hurlin (2012): “Testing for Granger non-causality in heterogeneous panels.” In: *Economic Modelling* 29 (4), pp. 1450–1460. DOI: 10.1016/j.econmod.2012.02.014.

Dunne, T. and X. Mu (2010): “Investment Spikes and Uncertainty in the Petroleum Refining Industry.” In: *The Journal of Industrial Economics* 58 (1), pp. 190–213. DOI: 10.1111/j.1467-6451.2010.00407.x.

EIA (2017a): *Crude Oil Production*. URL: https://www.eia.gov/dnav/pet/pet_crd_crpdn_adc_mbbl_m.htm.

– (2017b): *Henry Hub Natural Gas Spot and Futures Price*. URL: https://www.eia.gov/dnav/ng/ng_pri_fut_s1_d.htm.

– (2017c): *Tight oil production estimates by play*. URL: https://www.eia.gov/energyexplained/oil-and-petroleum-products/data/US-tight-oil-production.xlsx.

– (2017d): *WTI Spot and Futures Prices*. URL: https://www.eia.gov/petroleum/data.php#prices.

Fazzari, S. M., R. G. Hubbard, and B. C. Petersen (2000): “Investment-Cash Flow Sensitivities are Useful: A Comment on Kaplan and Zingales.” In: *The Quarterly Journal of Economics* 115 (2), pp. 695–705. DOI: 10.1162/003355300554773.

Fazzari, S. M., R. G. Hubbard, B. C. Petersen, A. S. Blinder, and J. M. Poterba (1988): “Financing Constraints and Corporate Investment.” In: *Brookings Papers on Economic Activity* 1988 (1), pp. 141–206.

Fitzgerald, T. (2013): “Frackonomics: Some Economics of Hydraulic Fracturing.” In: *Case Western Reserve Law Review* 63, pp. 1337–1362.

Flannery, M. J. and K. W. Hankins (2013): “Estimating dynamic panel models in corporate finance.” In: *Journal of Corporate Finance* 19, pp. 1–19. DOI: 10.1016/j.jcorpfin.2012.09.004.

Fort, T. C., J. Haltiwanger, R. S. Jarmin, and J. Miranda (2013): “How Firms Respond to Business Cycles: The Role of Firm Age and Firm Size.” In: *IMF Economic Review* 61 (3), pp. 520–559. DOI: 10.1057/imfer.2013.15.

Fosu, S. (2013): “Capital structure, product market competition and firm performance: Evidence from South Africa.” In: *The Quarterly Review of Economics and Finance* 53 (2), pp. 140–151. DOI: 10.1016/j.qref.2013.02.004.

Frank, M. Z. and V. K. Goyal (2008): “Chapter 12 - Trade-Off and Pecking Order Theories of Debt.” In: *Handbook of Empirical Corporate Finance*. Ed. by B. E. Eckbo. Handbooks in Finance. San Diego: Elsevier, pp. 135–202. DOI: 10.1016/B978-0-444-53265-7.50004-4.

– (2009): “Capital Structure Decisions: Which Factors Are Reliably Important?” In: *Financial Management* 38 (1), pp. 1–37. DOI: 10.1111/j.1755-053X.2009.01026.x.

Gilje, E., E. Loutskina, and D. P. Murphy (2017): *Drilling and Debt*. SSRN Scholarly Paper ID 2939603. Rochester, NY: Social Science Research Network. DOI: 10.2139/ssrn.2939603.

Hayakawa, K. (2009): “First Difference or Forward Orthogonal Deviation- Which Transformation Should be Used in Dynamic Panel Data Models?: A Simulation Study.” In: *Economics Bulletin* 29 (3), pp. 2008–2017.

Haynes and Boone, LLP (2017): *Haynes and Boone - Oil Patch Bankruptcy Monitor*. http://www.haynesboone.com/~/media/files/energy_bankruptcy_reports/2017/2017_oil_patch_monitor_20170731.ashx.

Holtz-Eakin, D., W. Newey, and H. S. Rosen (1988): “Estimating Vector Autoregressions with Panel Data.” In: *Econometrica* 56 (6), pp. 1371–1395.

Hurn, A. S. and R. E. Wright (1994): "Geology or Economics? Testing Models of Irreversible Investment Using North Sea Oil Data." In: *The Economic Journal* 104 (423), pp. 363–371.

International Energy Agency (IEA) (2014): *World Energy Investment Outlook: Special Report*. Tech. rep. https://www.iea.org/publications/freepublications/publication/WEIO2014.pdf. International Energy Agency.

Kapetanios, G. (2008): "A bootstrap procedure for panel data sets with many cross-sectional units." In: *Econometrics Journal* 11 (2), pp. 377–395. DOI: 10.1111/j.1368-423x.2008.00243.x.

Kaplan, S. N. and L. Zingales (1997): "Do Investment-Cash Flow Sensitivities Provide Useful Measures of Financing Constraints?" In: *The Quarterly Journal of Economics* 112 (1), pp. 169–215. DOI: 10.1162/003355397555163.

– (2000): "Investment-Cash Flow Sensitivities Are Not Valid Measures of Financing Constraints." In: *The Quarterly Journal of Economics* 115 (2), pp. 707–712. DOI: 10.1162/003355300554782.

Kayhan, A. and S. Titman (2007): "Firms' histories and their capital structures." In: *Journal of Financial Economics* 83 (1), pp. 1–32. DOI: 10.1016/j.jfineco.2005.10.007.

Kellogg, R. (2014): "The Effect of Uncertainty on Investment: Evidence from Texas Oil Drilling." In: *American Economic Review* 104 (6), pp. 1698–1734. DOI: 10.1257/aer.104.6.1698.

Kilian, L. (2016): "The Impact of the Shale Oil Revolution on U.S. Oil and Gasoline Prices." In: *Review of Environmental Economics and Policy* 10 (2), pp. 185–205. DOI: 10.1093/reep/rew001.

Lehn, K. and P. Zhu (2016): *Debt, Investment and Production in the U.S. Oil Industry: An Analysis of the 2014 Oil Price Shock*. SSRN Scholarly Paper ID 2817123. Rochester, NY: Social Science Research Network. DOI: 10.2139/ssrn.2817123.

Loo, M. van der (2014): "stringdist: an R Package for Approximate String Matching." In: *The R Journal* 6 (1), pp. 111–122.

Love, I. and L. Zicchino (2006): "Financial development and dynamic investment behavior: Evidence from panel VAR." In: *The Quarterly Review of Economics and Finance* 46 (2), pp. 190–210. DOI: 10.1016/j.qref.2005.11.007.

Moel, A. and P. Tufano (2002): "When Are Real Options Exercised? An Empirical Study of Mine Closings." In: *The Review of Financial Studies* 15 (1), p. 35. DOI: 10.1093/rfs/15.1.35.

Nickell, S. (1981): "Biases in Dynamic Models with Fixed Effects." In: *Econometrica* 49 (6), pp. 1417–1426.

Occupational Safety & Health Administration, U.S. Department of Labor (1987): *SIC description for 1311*. Accessed: 2018-10-05.

Pesaran, H. and Y. Shin (1998): "Generalized impulse response analysis in linear multivariate models." In: *Economics Letters* 58 (1), pp. 17–29. DOI: 10.1016/S0165-1765(97)00214-0.

Phillips, G. M. (1995): "Increased debt and industry product markets an empirical analysis." In: *Journal of Financial Economics* 37 (2), pp. 189–238. DOI: 10.1016/0304-405X(94)00785-Y.

Phillips, R. F. (2019): "Quantifying the Advantages of Forward Orthogonal Deviations for Long Time Series." In: *Computational Economics*. DOI: 10.1007/s10614-019-09907-w.

R Core Team (2018): *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria.

Roberts, M. R. and T. M. Whited (2013): "Chapter 7 - Endogeneity in Empirical Corporate Finance." In: *Handbook of the Economics of Finance*. Ed. by G. M. Constantinides, M. Harris, and R. M. Stulz. Vol. 2. Handbook of the Economics of Finance. Elsevier, pp. 493–572. DOI: 10.1016/B978-0-44-453594-8.00007-0.

Rogers, H. (2011): "Shale gas - the unfolding story." In: *Oxford Review of Economic Policy* 27 (1), p. 117. DOI: 10.1093/oxrep/grr004.

Roll, K. H. and R. E. Dahl (2017): "Survival of the fittest: US oil productivity during business cycles." In: *Heading towards Sustainable Energy Systems: Evolution or Revolution?* IAEE.

Roodman, D. (2009): "A Note on the Theme of Too Many Instruments." In: *Oxford Bulletin of Economics and Statistics* 71 (1), pp. 135–158. DOI: 10.1111/j.1468-0084.2008.00542.x.

Sigmund, M. and R. Ferstl (2019): "Panel vector autoregression in R with the package panelvar." In: *The Quarterly Review of Economics and Finance*. DOI: 10.1016/j.qref.2019.01.001.

The Economist (2015): "After OPEC - American shale firms are now the oil market's swing producers." In: *The Economist*.

Thompson, A. C. (2001): "The Hotelling Principle, backwardation of futures prices and the values of developed petroleum reserves - the production constraint hypothesis." In: *Resource and Energy Economics* 23 (2), pp. 133–156. DOI: 10.1016/S0928-7655(00)00040-3.

Titman, S. and R. Wessels (1988): "The Determinants of Capital Structure Choice." In: *The Journal of Finance* 43 (1), pp. 1–19. DOI: 10.1111/j.1540-6261.1988.tb02585.x.

# Online Appendix
## Debt and the Oil Industry – Analysis on the Firm and Production Level

Johannes Lips *

July 1, 2019

This is the online appendix to the paper "Debt and the Oil Industry – Analysis on the Firm and Production Level" by Johannes Lips

# Appendices

## A. Leverage Quartiles of Companies

To analyze the impact leverage might have on the production, the companies are categorized into quartiles according to their leverage just prior to the price decline in the third quarter of 2008 and the fourth quarter in 2014. This means that companies with a lower leverage, relative to all other companies, are in the 1st quartile and the companies with the highest relative leverage end up in the 4th quartile.

| | 2008 Q2 | | | 2014 Q3 | | |
|---|---|---|---|---|---|---|
| Leverage Percentile | No. | Assets | Debt | No. | Assets | Debt |
| $1^{st}$ *Quartile* | 33 | 3094 | 493 | 34 | 5872 | 948 |
| $2^{nd}$ *Quartile* | 36 | 11 869 | 2494 | 37 | 12 895 | 2749 |
| $3^{rd}$ *Quartile* | 35 | 5018 | 1380 | 37 | 4279 | 1328 |
| $4^{th}$ *Quartile* | 35 | 2876 | 1221 | 37 | 2002 | 885 |
| Non-calculable Leverage | 5 | 1190 | 343 | 6 | 1327 | 397 |

Table 1: Comparison of the number of companies for each leverage group prior to price declines in 2008 Q2 and 2014 Q3 and their average value of total assets and debt in million US dollar.

To investigate if the adoption of new technologies is affected by a companies' leverage, the share of oil and gas production from conventional and unconventional for the four leverage quartile and its development over time is depicted in Figure 1 and 2. It can be seen that irrespective of the leverage quartile a company was in before the oil price decline in 2008, the adoption of new production technologies and thus the production from unconventional sources

*Department of Economics, Justus-Liebig University of Giessen, Chair of Statistics and Econometrics (VWL VII), Licher Str. 64, 35394 Giessen, Germany. E-mail: Johannes.Lips@wirtschaft.uni-giessen.de

| | Leverage Quartile 2014 | | | | |
|---|---|---|---|---|---|
| Leverage Quartile 2008 | $1^{st}Quartile$ | $2^{nd}Quartile$ | $3^{rd}Quartile$ | $4^{th}Quartile$ | Non-calculable leverage 2014 |
| $1^{st}Quartile$ | 11 | 4 | 3 | 4 | 13 |
| $2^{nd}Quartile$ | 4 | 10 | 9 | 4 | 9 |
| $3^{rd}Quartile$ | 5 | 8 | 9 | 3 | 11 |
| $4^{th}Quartile$ | – | 1 | 5 | 9 | 20 |
| Non-calculable leverage 2008 | 17 | 14 | 11 | 18 | 137 |

Table 2: Companies' transition from leverage quartiles in 2008 to 2014.

increases with a similar trend and pattern. This indicates that higher leverage did not act as a constraint on the companies and their adoption of new technologies. On the contrary, it appears to be the case that companies which in 2008 were in the three highest leverage quartiles more strongly increased the share of production from unconventional sources. This is also evident, when looking at the growth rates of the production for each leverage group. The production of oil from unconventional sources increased from the third quarter of 2008 to the first quarter of 2016 by 239% for the highest leverage quartile and only by 126% for the lowest quartile.[1] In case of natural gas the differences between the leverage groups are less pronounced and vary between 28% for the third leverage group and 102% for the highest leverage group.[2] The difference between the two fossil fuels is mainly due to a much higher initial production from unconventional sources in case of natural gas already in 2008. Across all leverage groups, the production from conventional sources decreased substantially.

To analyze the relationship between the adoption of new technologies and the companies' leverage quartile, the movements between the leverage quartiles from 2008 to 2014 are categorized into upward, downward and no movement. In Figures 3 and 4 the share of oil and gas production from unconventional sources is displayed and it can be observed that the adoption of new technologies is not associated with companies moving into a higher leverage quartile. Rather it can be seen that the share of unconventional oil production increased more for companies which moved into a lower leverage quartile in 2014.

The movement into a lower leverage group could be seen as an indicator that especially the possibility of unconventional production techniques and their considerably lower upfront investment volumes allowed the increase of production capacity with lower investment volumes. Although it has to be considered that before the oil price drop in 2014 the asset valuation of companies might be relatively high as well.

[1] The growth rate for the second and third quartile are 173% and 132%, respectively.

[2] The growth rate for the first and second leverage quartile is 81% and 86%, respectively.

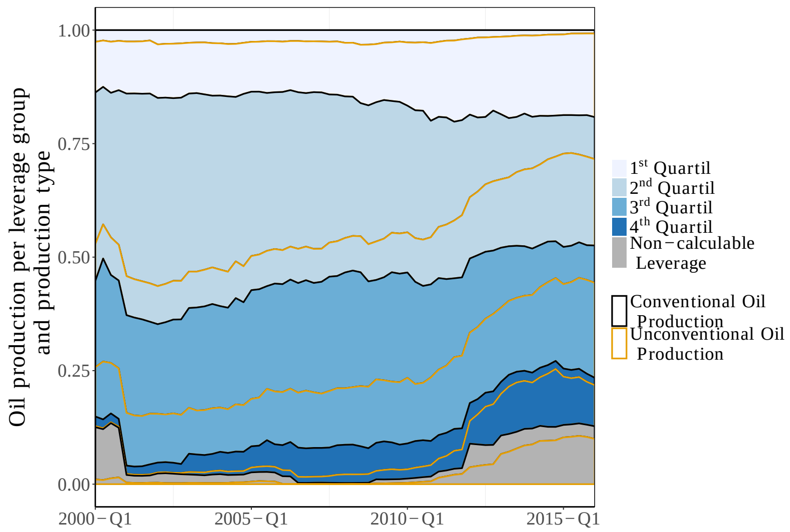



Figure 1: Total oil production differentiated by production type and leverage quartile of the companies in 2008. Yellow line separates the production types with conventional share above and unconventional share below.

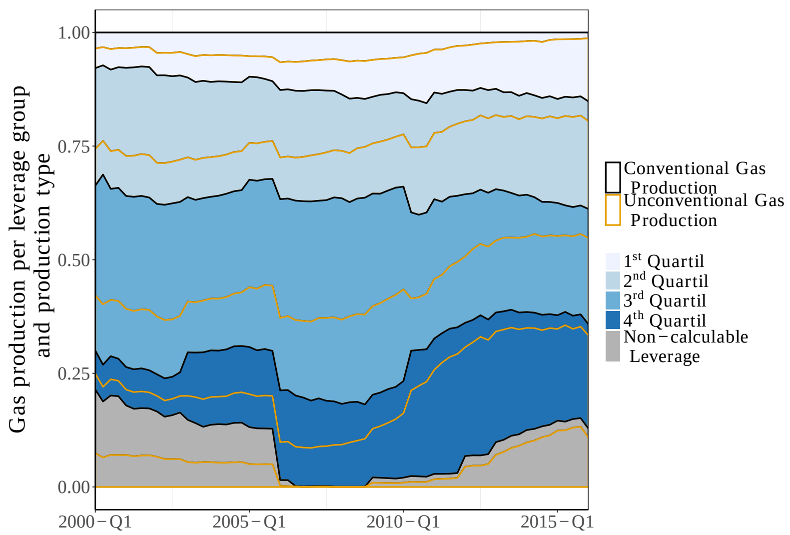



Figure 2: Total gas production differentiated by production type and leverage quartile of the companies in 2008. Yellow line separates the production types with conventional share above and unconventional share below.

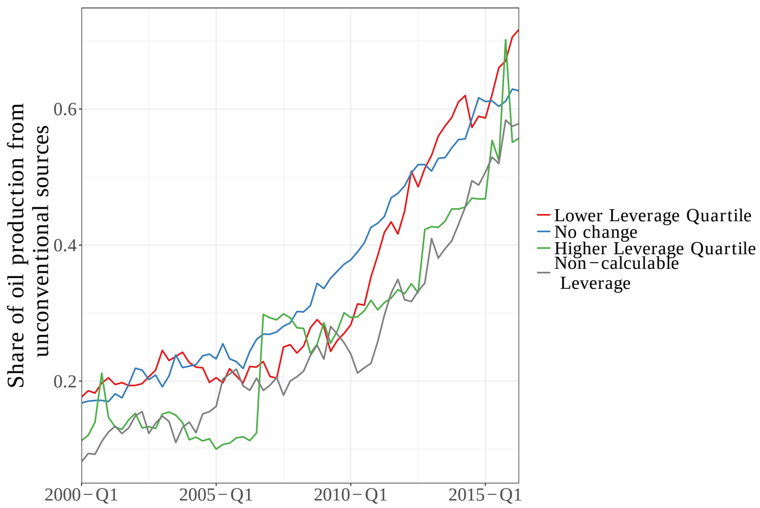



Figure 3: Share of oil production from unconventional sources differentiated by companies' leverage transition from 2008 to 2014.

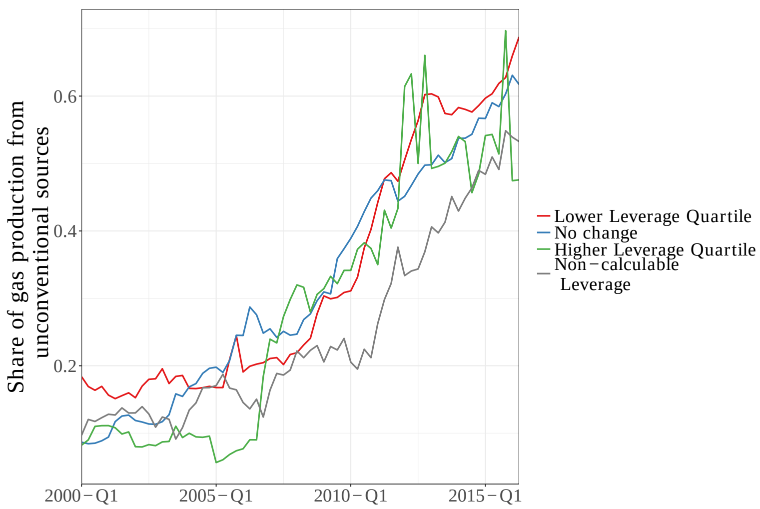



Figure 4: Share of gas production from unconventional sources differentiated by companies' leverage transition from 2008 to 2014.

## B. Development of differences between spot and future markets

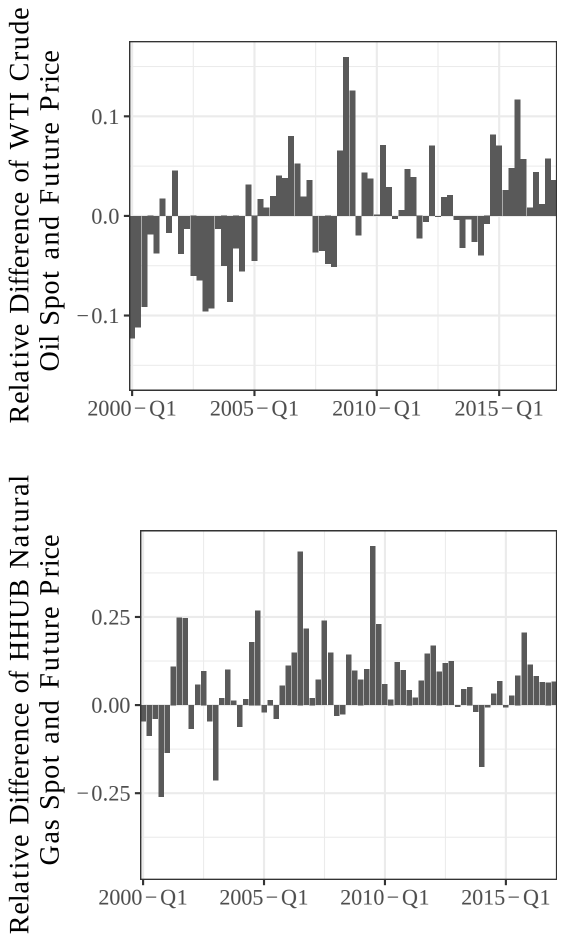



Figure 5: Relative difference of WTI crude oil and Henry Hub natural gas spot and future prices. Positive differences indicate periods of contango and negative differences periods of backwardation.
Data source: WTI price time series (EIA 2017b) and Henry Hub Natural Gas price time series (EIA 2017a)

# C. Development of the share of gas production from unconventional sources for different leverage groups

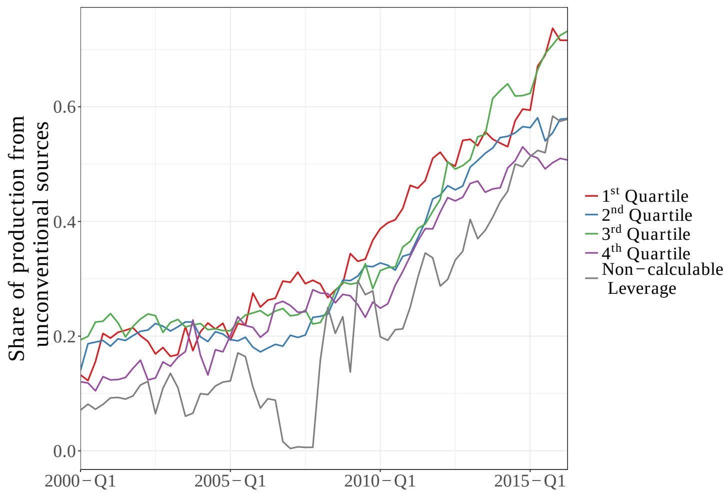



Figure 6: Share of oil production from unconventional sources, based on the leverage quartile of the companies in 2008
Source: Own calculations

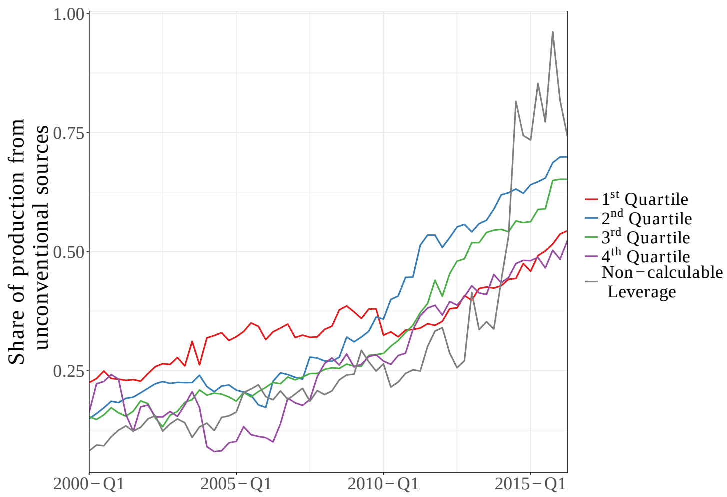



Figure 7: Share of oil production from unconventional sources, based on the leverage quartile of the companies in 2014
Source: Own calculations

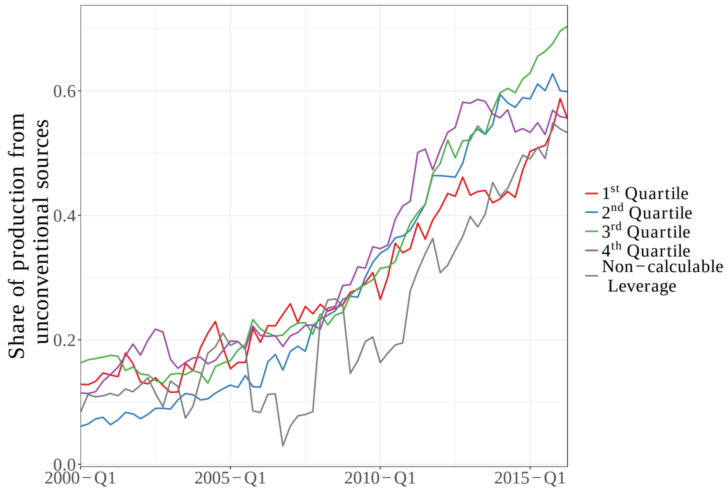



Figure 8: Share of gas production from unconventional sources, based on the leverage quartile of the companies in 2008
Source: Own calculations

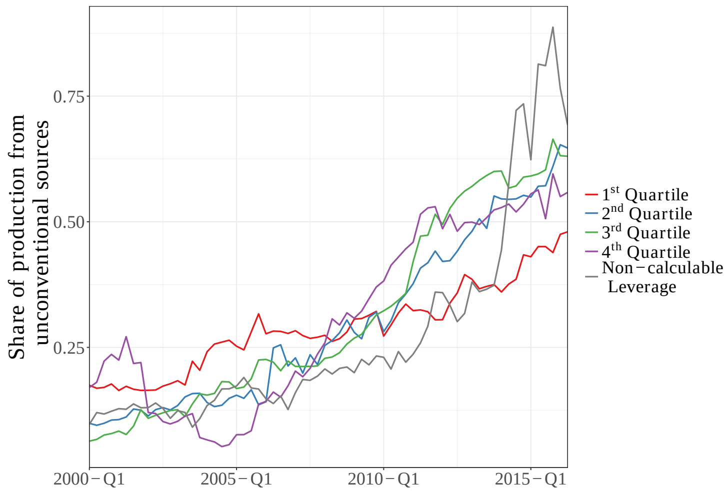



Figure 9: Share of gas production from unconventional sources, based on the leverage quartile of the companies in 2014
Source: Own calculations

# D. Debt-to-EBITDA ratio categorization of companies

| | 2008 Q2 | | | 2014 Q3 | | |
|---|---|---|---|---|---|---|
| Leverage Percentile | No. | Assets | Debt | No. | Assets | Debt |
| $1^{st} Quartile$ | 32 | 2752 | 987 | 33 | 551 | 196 |
| $2^{nd} Quartile$ | 32 | 5465 | 880 | 36 | 11 677 | 2448 |
| $3^{rd} Quartile$ | 33 | 8978 | 2034 | 36 | 5454 | 1462 |
| $4^{th} Quartile$ | 34 | 5827 | 1748 | 36 | 3862 | 1221 |
| Non-calculable Leverage | 13 | 1821 | 435 | 10 | 1882 | 477 |

Table 3: Comparison of the number of companies for each debt-to-EBITDA group prior to price declines in 2008 Q2 and 2014 Q3 and their average value of total assets and debt in million US dollar.

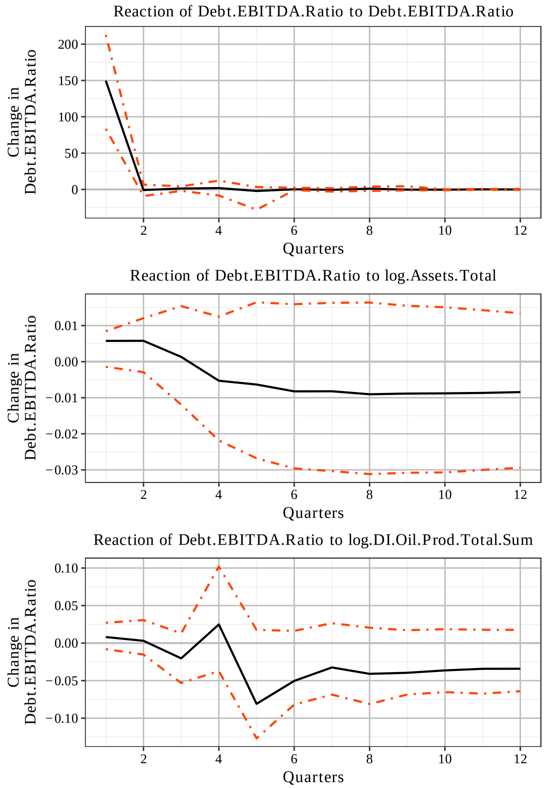
| | Leverage Quartile 2014 | | | | |
|---|---|---|---|---|---|
| Debt-to-EBITDA Ratio 2008 | $1^{st} Quartile$ | $2^{nd} Quartile$ | $3^{rd} Quartile$ | $4^{th} Quartile$ | Non-calculable ratio 2014 |
| $1^{st} Quartile$ | 4 | 4 | 5 | 3 | 17 |
| $2^{nd} Quartile$ | 6 | 6 | 3 | 5 | 13 |
| $3^{rd} Quartile$ | – | 7 | 9 | 8 | 9 |
| $4^{th} Quartile$ | 2 | 11 | 5 | 4 | 12 |
| Non-calculable ratio 2008 | 23 | 8 | 14 | 16 | 145 |

Table 4: Companies' transition from Debt-to-EBITDA ratio quartiles in 2008 to 2014.

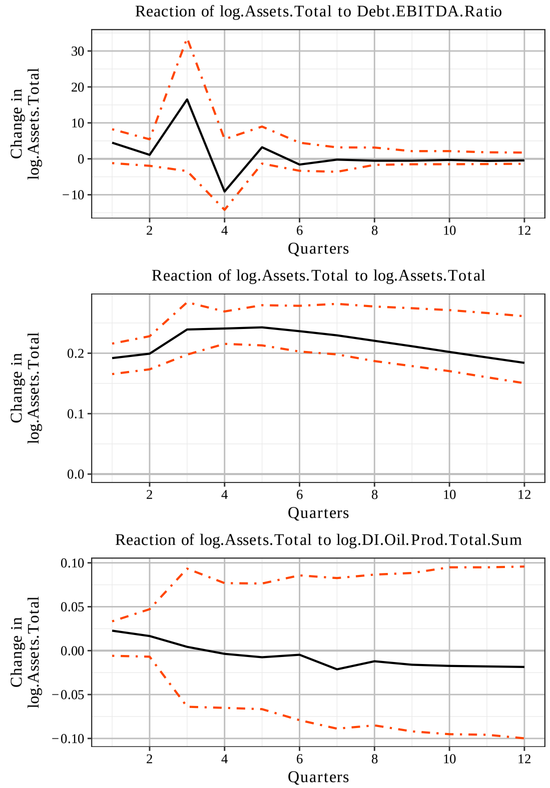
# E. Panel VAR Results – Debt-to-EBITDA Ratio

| | Dependent Variables | | |
|---|---|---|---|
| | Debt.EBITDA.Ratio | log.Assets.Total | log.DI.Oil.Prod.Total.Sum |
| Debt.EBITDA.Ratio$_{t-1}$ | −0.0058 | −0.0000 | −0.0000 |
| | (0.0265) | (0.0000) | (0.0001) |
| log.Assets.Total$_{t-1}$ | 6.0653 | 1.0380*** | 0.0149 |
| | (10.3509) | (0.0428) | (0.0639) |
| log.DI.Oil.Prod.Total.Sum$_{t-1}$ | −1.4714 | 0.0003 | 0.6094*** |
| | (4.0463) | (0.0025) | (0.0773) |
| Debt.EBITDA.Ratio$_{t-2}$ | 0.0052 | −0.0000 | −0.0002 |
| | (0.0082) | (0.0000) | (0.0001) |
| log.Assets.Total$_{t-2}$ | 79.4668 | 0.1701*** | −0.0516 |
| | (51.1615) | (0.0301) | (0.1941) |
| log.DI.Oil.Prod.Total.Sum$_{t-2}$ | 3.4038 | −0.0005 | 0.0794* |
| | (3.9367) | (0.0049) | (0.0416) |
| Debt.EBITDA.Ratio$_{t-3}$ | 0.0152 | −0.0000* | 0.0002 |
| | (0.0320) | (0.0000) | (0.0003) |
| log.Assets.Total$_{t-3}$ | −136.0628* | −0.2150 | −0.0326 |
| | (72.7385) | (0.1537) | (0.2811) |
| log.DI.Oil.Prod.Total.Sum$_{t-3}$ | −13.1325* | 0.0073 | 0.2063* |
| | (6.7019) | (0.0075) | (0.1184) |
| Debt.EBITDA.Ratio$_{t-4}$ | −0.0110 | 0.0000 | −0.0006*** |
| | (0.0162) | (0.0000) | (0.0001) |
| log.Assets.Total$_{t-4}$ | 50.2898* | −0.0234 | 0.0533 |
| | (27.5435) | (0.1087) | (0.1874) |
| log.DI.Oil.Prod.Total.Sum$_{t-4}$ | 12.7301 | −0.0046 | 0.0410 |
| | (9.7839) | (0.0067) | (0.1048) |
| log.Last.Quarterly.WTI.Spot.Price | 0.2157 | 0.0467*** | 0.1191*** |
| | (4.0823) | (0.0104) | (0.0381) |
| const | 9.9956 | 0.0308 | −0.5291** |
| | (15.2542) | (0.0438) | (0.2118) |
| Observations | 8373 | | |
| Number of Groups | 330 | | |
| Avg. Obs. Group | 25.37 | | |
| Min. Obs. Group | 1 | | |
| Max. Obs. Group | 66 | | |
| *Note:* | | | *p<0.1; **p<0.05; ***p<0.01 |
| | | | Corrected standard errors are reported in parentheses. |
| | | | Variable transformation: Forward Orthogonal Deviation |

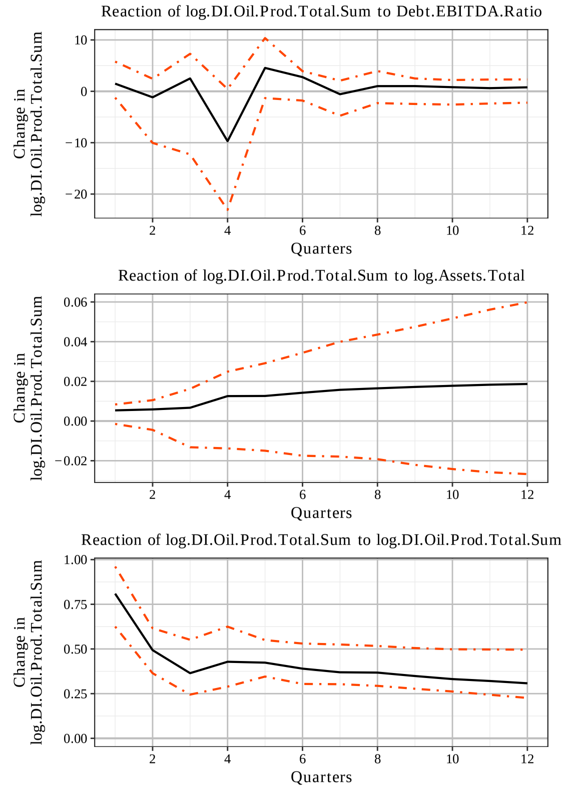
Table 5: Results of the panel vector autoregressive (VAR) approach for the oil production.




(a) Impact on the debt to EBITDA ratio.




(b) Impact on total assets.




(c) Impact on oil production.

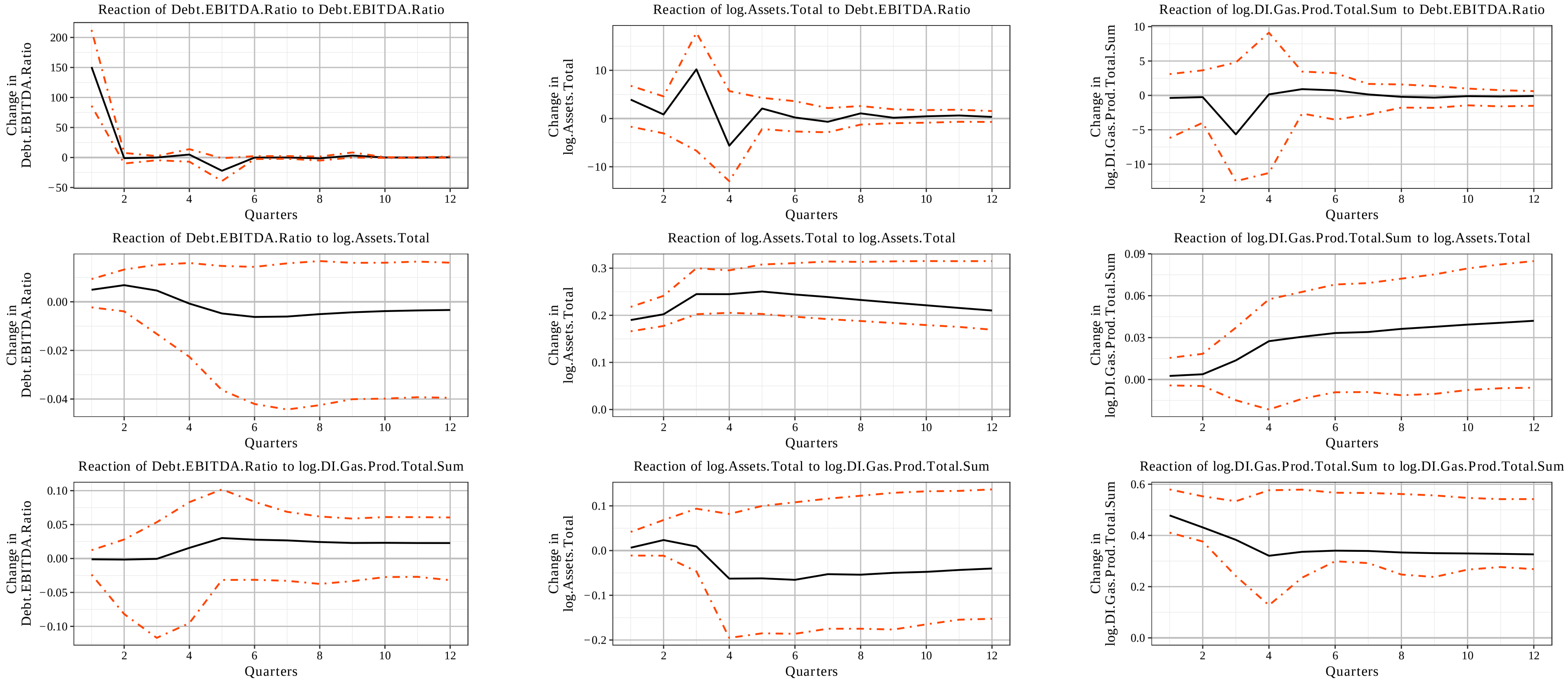
Figure 10: Generalized impulse response function (IRF) for the whole sample. Dashed red lines indicate the bootstrapped 95% confidence intervals based on 500 iterations.
Source: Own calculations

# F. Panel VAR Results – Gas Production

| | Dependent Variables | | |
|---|---|---|---|
| | log.Assets.Total | Debt.EBITDA.Ratio | log.DI.Gas.Prod.Total.Sum |
| log.Assets.Total$_{t-1}$ | 1.0641*** | 4.6058 | 0.0935** |
| | (0.0477) | (7.7641) | (0.0439) |
| Debt.EBITDA.Ratio$_{t-1}$ | 0.0000 | −0.0075 | −0.0000 |
| | (0.0000) | (0.0234) | (0.0001) |
| log.DI.Gas.Prod.Total.Sum$_{t-1}$ | 0.0021 | −0.5538 | 0.9017*** |
| | (0.0070) | (3.0848) | (0.0416) |
| log.Assets.Total$_{t-2}$ | 0.1565*** | 49.3703 | −0.1628 |
| | (0.0323) | (30.8471) | (0.1486) |
| Debt.EBITDA.Ratio$_{t-2}$ | −0.0000 | −0.0020 | 0.0000 |
| | (0.0000) | (0.0060) | (0.0001) |
| log.DI.Gas.Prod.Total.Sum$_{t-2}$ | 0.0175 | −11.6515 | −0.0127 |
| | (0.0218) | (7.1595) | (0.1157) |
| log.Assets.Total$_{t-3}$ | −0.2520* | −87.3027 | −0.3199 |
| | (0.1325) | (57.1672) | (0.2143) |
| Debt.EBITDA.Ratio$_{t-3}$ | −0.0000 | 0.0330 | 0.0001 |
| | (0.0000) | (0.0294) | (0.0001) |
| log.DI.Gas.Prod.Total.Sum$_{t-3}$ | 0.0097 | 11.1623 | −0.0396 |
| | (0.0210) | (11.2449) | (0.0644) |
| log.Assets.Total$_{t-4}$ | 0.0148 | 35.9993 | 0.3975 |
| | (0.0846) | (30.9461) | (0.2463) |
| Debt.EBITDA.Ratio$_{t-4}$ | −0.0000 | −0.1446*** | 0.0001 |
| | (0.0001) | (0.0427) | (0.0002) |
| log.DI.Gas.Prod.Total.Sum$_{t-4}$ | −0.0242 | 0.2333 | 0.1438* |
| | (0.0147) | (12.2327) | (0.0819) |
| Last.Quarterly.WTI.Spot.Price | 0.0006*** | −0.0043 | 0.0008* |
| | (0.0001) | (0.0572) | (0.0005) |
| const | 0.0510 | −4.7637 | −0.0718 |
| | (0.0315) | (9.6285) | (0.0521) |
| Observations | 8373 | | |
| Number of Groups | 330 | | |
| Avg. Obs. Group | 25.37 | | |
| Min. Obs. Group | 1 | | |
| Max. Obs. Group | 66 | | |

*Note:* *p<0.1; **p<0.05; ***p<0.01
Corrected standard errors are reported in parentheses.
Variable transformation: Forward Orthogonal Deviation

Table 6: Results of the panel VAR approach for the natural gas production.




(a) Impact on the debt to EBITDA ratio. (b) Impact on total assets. (c) Impact on gas production.

Figure 11: Generalized IRF for the whole sample. Dashed red lines indicate the bootstrapped 95% confidence intervals based on 500 iterations.
Source: Own calculations

# G. Empirical Results for Subsamples

## High and Low Leverage Subsample

To shed further light on the potential determinants of production decisions, the sample is divided into subsamples and it is analyzed how determinants vary across subsamples. Specifically, the analysis focuses on companies' variations in (1) leverage, (2) share of unconventional production.

Based on the categorization the estimation of the panel Vector Autoregressive (VAR) is performed on the 50% companies comprising the higher leverage group and on the 50% being in the group with a lower leverage. This means, that leverage in this analysis is a relative measure always in relation to the leverage across the sample in each quarter.

Table 7 shows some basic descriptive statistics for the two subsamples. It can be seen that the leverage and the Debt-to-earnings before interest, taxes, depreciation, and amortization (EBITDA) ratio are, as expected, higher for the high leverage subsample. In nearly all cases the distribution of the variables is heavily right-skewed, since the median value, in most cases, is considerably smaller than the mean value. The two subsamples are roughly the same size, since on average there are 30 companies in the lower and 33 in the higher subsample, respectively.

| | Mean | | Median | |
|---|---|---|---|---|
| | Low | High | Low | High |
| Leverage | 0.14 | 0.57 | 0.15 | 0.41 |
| Assets | 6473.25 | 3423.78 | 379.66 | 906.68 |
| Debt | 1226.47 | 1283.63 | 38.54 | 384.68 |
| Debt-to-EBITDA ratio | Inf | 6.93 | 2.05 | 7.25 |
| Quarterly Oil Production | 1.53 | 0.93 | 0.03 | 0.11 |
| Quarterly Oil Production | 13 104.37 | 16 897.92 | 318.31 | 1877.03 |
| Avg. Number of Companies | 30 | 33 | | |

Table 7: Descriptive statistics for the variables used in the analysis and differentiated between the two subsamples with low and high levels of leverage. Oil Production is measured in million Barrels (mmBbls) per quarter, gas production in million British thermal units (mmBtus) and the financial data is reported in million US dollar.

## Results Subsamples Leverage

For both subsamples the same panel VAR approach is estimated and the results are shown in Table 8, for the low and in Table 8 for the high leverage subsample. The estimation results are based on 3960 (4413) company quarter observations, which are made up of 253 (228) companies in the low (high) leverage subsample. There are some notable differences between the results for the two subsamples. Especially the impact of the leverage ratio on the total assets is much more pronounced for the sample with a relatively high level of leverage. Additionally, it is interesting, that in the high leverage subsample the price elasticity of the oil production, with a value of 0.1587 is even greater than for the whole subsample and in contrast to the low leverage subsample it is highly significant.

The generalized IRF for the two subsamples are shown side by side in Figures 12, 13 and 14. There are no substantial differences identifiable between the two subsamples, only in certain cases some minor differences in the reaction to shocks is discernible.

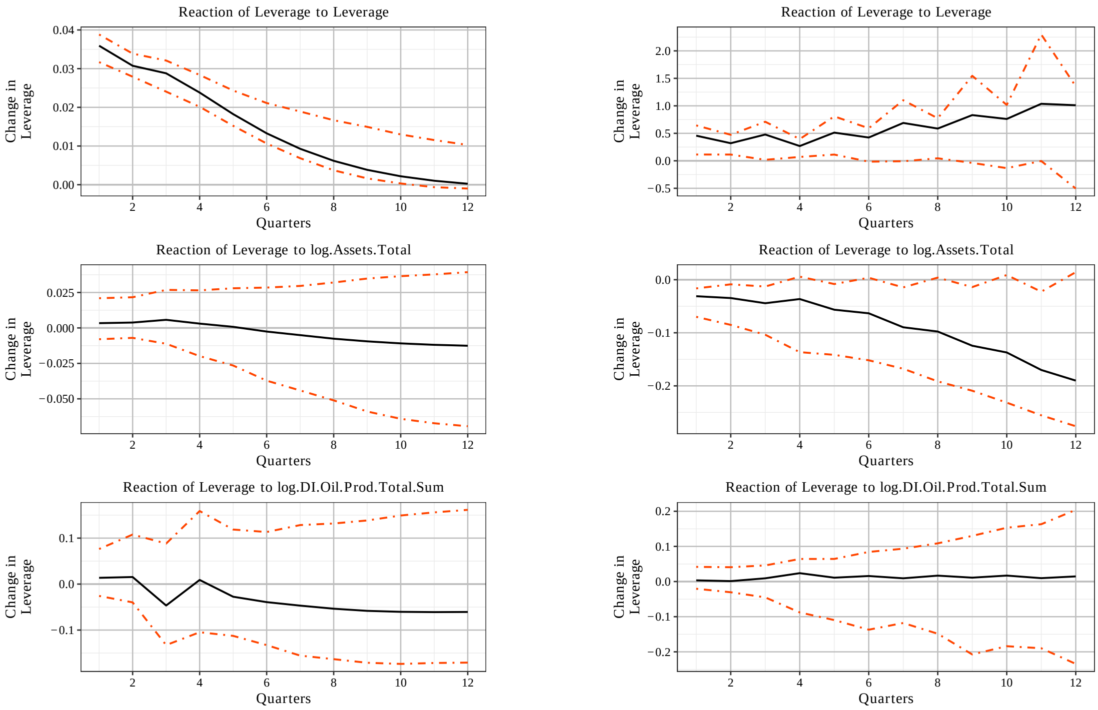
| | Dependent Variables | | |
|---|---|---|---|
| | Leverage | log.Assets.Total | log.DI.Oil.Prod.Total.Sum |
| $Leverage_{t-1}$ | 0.8562*** | 0.0060 | 0.1667 |
| | (0.0307) | (0.1171) | (0.4903) |
| $log.Assets.Total_{t-1}$ | −0.0036 | 1.0492*** | 0.0345 |
| | (0.0064) | (0.0450) | (0.0960) |
| $log.DI.Oil.Prod.Total.Sum_{t-1}$ | 0.0000 | 0.0041 | 0.6773*** |
| | (0.0012) | (0.0026) | (0.0723) |
| $Leverage_{t-2}$ | 0.0701 | 0.0349 | −1.7511 |
| | (0.0510) | (0.1509) | (1.4410) |
| $log.Assets.Total_{t-2}$ | −0.0010 | 0.0718 | −0.3447 |
| | (0.0125) | (0.0922) | (0.3025) |
| $log.DI.Oil.Prod.Total.Sum_{t-2}$ | −0.0022* | −0.0010 | 0.1321** |
| | (0.0013) | (0.0042) | (0.0670) |
| $Leverage_{t-3}$ | −0.0854 | −0.1087 | 2.4115 |
| | (0.0745) | (0.2464) | (1.8750) |
| $log.Assets.Total_{t-3}$ | 0.0267* | −0.1262 | 0.1643 |
| | (0.0154) | (0.0795) | (0.4145) |
| $log.DI.Oil.Prod.Total.Sum_{t-3}$ | 0.0031** | 0.0051 | 0.1110 |
| | (0.0015) | (0.0050) | (0.1266) |
| $Leverage_{t-4}$ | −0.0482 | −0.0056 | −1.5730 |
| | (0.0488) | (0.2109) | (1.5486) |
| $log.Assets.Total_{t-4}$ | −0.0161 | −0.0210 | 0.1693 |
| | (0.0099) | (0.0546) | (0.3805) |
| $log.DI.Oil.Prod.Total.Sum_{t-4}$ | 0.0007 | −0.0040 | 0.0248 |
| | (0.0011) | (0.0046) | (0.0549) |
| log.Last.Quarterly.WTI.Spot.Price | −0.0111*** | 0.0366*** | 0.0521 |
| | (0.0022) | (0.0129) | (0.0672) |
| const | 0.0419*** | 0.0572 | −0.3778 |
| | (0.0102) | (0.0440) | (0.2611) |
| Observations | 4031 | | |
| Number of Groups | 261 | | |
| Avg. Obs. Group | 15.44 | | |
| Min. Obs. Group | 1 | | |
| Max. Obs. Group | 66 | | |

*Note:* *p<0.1; **p<0.05; ***p<0.01
Corrected standard errors are reported in parentheses.
Variable transformation: Forward Orthogonal Deviation

Table 8: Results of the panel VAR approach for the subsample with a relatively low leverage.

| | Dependent Variables | | |
|---|---|---|---|
| | Leverage | log.Assets.Total | log.DI.Oil.Prod.Total.Sum |
| $\text{Leverage}_{t-1}$ | 0.6682*** | 0.0023 | 0.0057 |
| | (0.1797) | (0.0079) | (0.0161) |
| $\text{log.Assets.Total}_{t-1}$ | −0.4663*** | 1.1448*** | 0.1015 |
| | (0.1649) | (0.0236) | (0.1327) |
| $\text{log.DI.Oil.Prod.Total.Sum}_{t-1}$ | 0.0046 | −0.0083** | 0.5284*** |
| | (0.0064) | (0.0038) | (0.1374) |
| $\text{Leverage}_{t-2}$ | 0.5434 | −0.0051 | 0.0381 |
| | (0.6353) | (0.0122) | (0.0270) |
| $\text{log.Assets.Total}_{t-2}$ | 0.0117 | 0.1006* | 0.2498 |
| | (0.5689) | (0.0535) | (0.2840) |
| $\text{log.DI.Oil.Prod.Total.Sum}_{t-2}$ | −0.0068 | −0.0042 | 0.1678** |
| | (0.0122) | (0.0051) | (0.0848) |
| $\text{Leverage}_{t-3}$ | −0.4586 | 0.0237** | 0.0224 |
| | (0.5078) | (0.0116) | (0.0222) |
| $\text{log.Assets.Total}_{t-3}$ | 1.1371 | −0.2414*** | −0.1868 |
| | (1.0425) | (0.0899) | (0.2601) |
| $\text{log.DI.Oil.Prod.Total.Sum}_{t-3}$ | 0.0417 | 0.0107 | 0.2763* |
| | (0.0284) | (0.0088) | (0.1426) |
| $\text{Leverage}_{t-4}$ | 0.4880*** | −0.0545** | −0.0562* |
| | (0.0855) | (0.0249) | (0.0304) |
| $\text{log.Assets.Total}_{t-4}$ | −0.6444 | −0.0222 | −0.1178 |
| | (0.4993) | (0.0600) | (0.2036) |
| $\text{log.DI.Oil.Prod.Total.Sum}_{t-4}$ | −0.0186 | −0.0024 | −0.0073 |
| | (0.0313) | (0.0087) | (0.1763) |
| log.Last.Quarterly.WTI.Spot.Price | −0.0297 | 0.0268*** | 0.1587*** |
| | (0.0212) | (0.0083) | (0.0523) |
| const | −0.2024 | 0.0274 | −1.0860*** |
| | (0.1424) | (0.0483) | (0.3661) |
| Observations | 4341 | | |
| Number of Groups | 262 | | |
| Avg. Obs. Group | 16.57 | | |
| Min. Obs. Group | 1 | | |
| Max. Obs. Group | 66 | | |
| *Note:* | *p<0.1; **p<0.05; ***p<0.01 | | |
| | Corrected standard errors are reported in parentheses. | | |
| | Variable transformation: Forward Orthogonal Deviation | | |

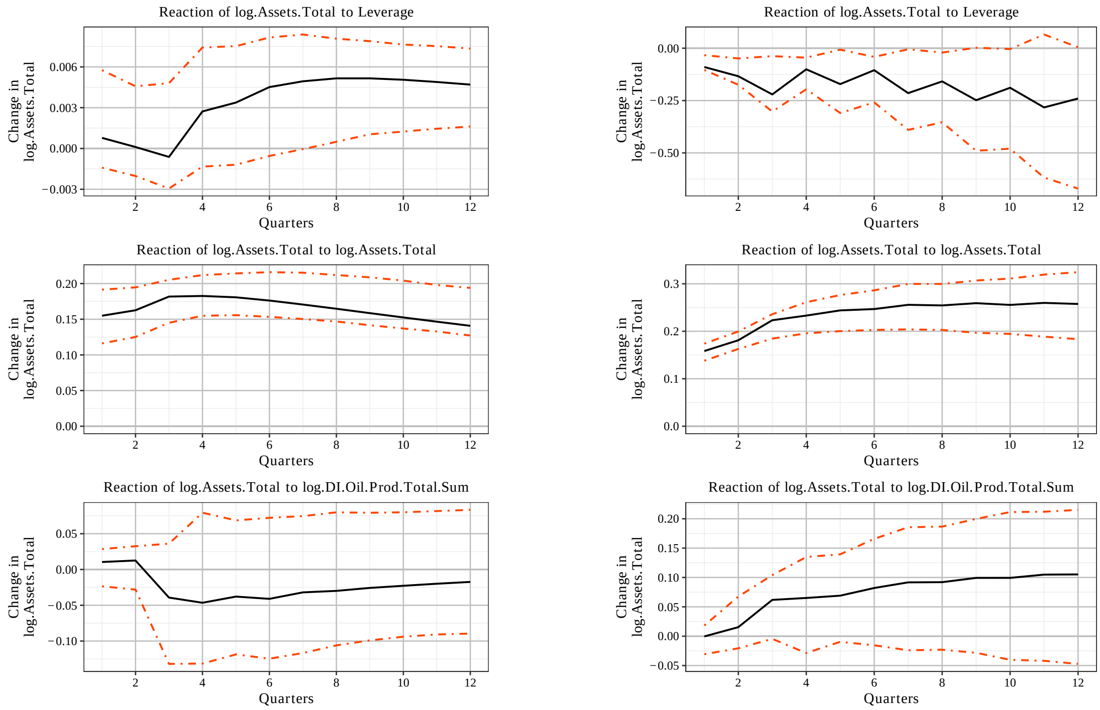
Table 9: Results of the panel VAR approach for the subsample with a relatively high leverage.




(a) Subsample of companies with a relatively **low** leverage.

(b) Subsample of companies with a relatively **high** leverage.

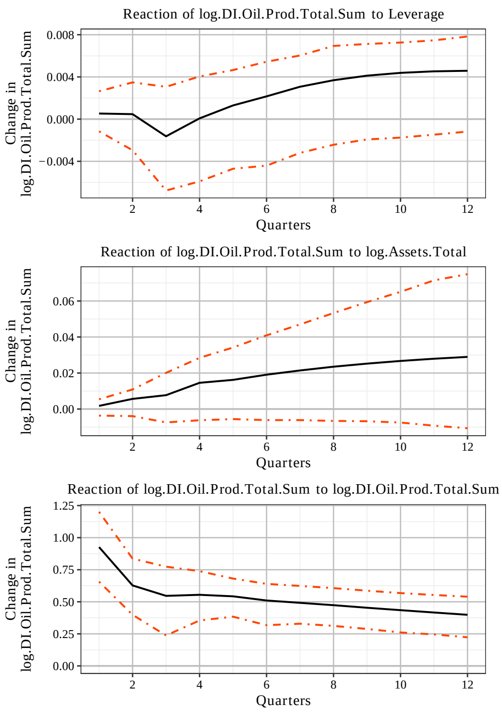
Figure 12: Generalized IRF for the impact on the leverage ratio. Dashed red lines indicate the bootstrapped 95% confidence intervals based on 500 iterations.
Source: Own calculations



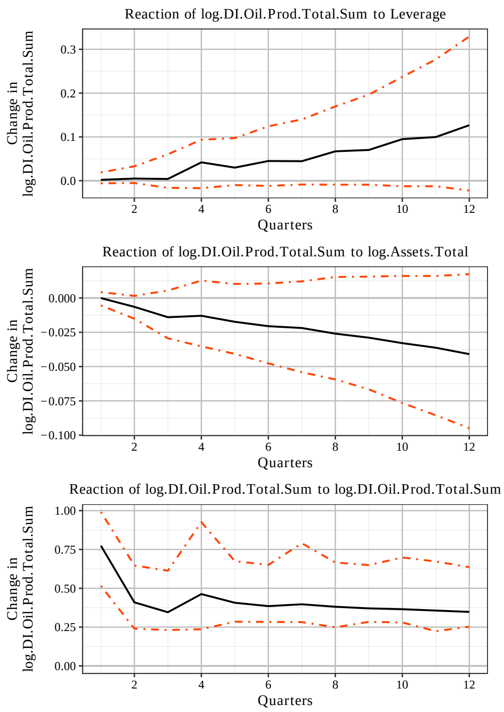

(a) Subsample of companies with a relatively **low** leverage.

(b) Subsample of companies with a relatively **high** leverage.

Figure 13: Generalized IRF for the impact on the assets. Dashed red lines indicate the bootstrapped 95% confidence intervals based on 500 iterations.
Source: Own calculations




(a) Subsample of companies with a relatively **low** leverage.




(b) Subsample of companies with a relatively **high** leverage.

Figure 14: Generalized IRF for the impact on the assets. Dashed red lines indicate the bootstrapped 95% confidence intervals based on 500 iterations.
Source: Own calculations

## High and Low Unconventional Production Subsample

The second subsample is based on a company's share of oil production from unconventional sources and thus should shed some light on different reactions based on the production technology. Therefore, the companies are clustered into two groups based into which of the four percentiles they fall. Companies in the lower two quartiles are assigned into the subsample with a relatively lower share of production from unconventional sources, the companies in the two upper quartiles comprise the subsample with a higher share accordingly.

In Table 10 some descriptive statistics for the two subsamples are provided. The main difference in companies' characteristics based on these statistics is that companies with a relatively high share of production from unconventional sources are considerably larger than companies in the lower subsample. These companies have on average more assets and a much larger oil production, nevertheless the average and median values of the leverage ratio are pretty similar across subsamples.

| | Mean | | Median | |
|---|---|---|---|---|
| | Low | High | Low | High |
| Leverage | 0.38 | 0.34 | 0.27 | 0.28 |
| Assets | 3769.42 | 5899.81 | 246.67 | 1313.83 |
| Debt | 1002.79 | 1483.42 | 57.85 | 394.28 |
| Debt-to-EBITDA ratio | Inf | 7.64 | 3.37 | 4.59 |
| Quarterly Oil Production | 0.57 | 1.80 | 0.00 | 0.35 |
| Quarterly Oil Production | 4287.24 | 24 748.06 | 61.47 | 5099.78 |
| Avg. Number of Companies | 57 | 33 | | |

Table 10: Descriptive statistics for the variables used in the analysis and differentiated between the two subsamples with low and high levels of production from unconventional oil sources. Oil Production is measured in mmBbls per quarter, gas production in mmBtus and the financial data is reported in million US dollar.

## Results Subsamples Unconventional Production

The results for the generalized method of moments (GMM) estimation of the panel VAR are displayed in Tables 11 and 12 and they provide some interesting insights. In general the results confirm previous results that the three endogenous variables are mainly affected by lagged values of themselves. Nevertheless, there is an interesting difference in the results for the two subsamples, namely the coefficient for the impact of the price of West Texas Intermediate (WTI) on the oil production differs considerably. In the low unconventional production sample the impact of the oil price is not statistically significant and thus it seems that for these companies the oil price does not affect their production decisions. In contrast the coefficient for the sample with a higher share from unconventional production the coefficient is highly significant and a value of 0.0929 shows a relatively high price elasticity of the oil production. This result lends some support for the hypothesis that shale oil producers are the new swing producers for the world oil market, since they are able to dynamically adjust their production in response to changes in the oil price.

The generalized IRF for the two subsamples are shown side by side in Figures 15, 16 and 17. Generally there are no substantial differences between the reaction of the variables for the two subsamples.

| | Dependent Variables | | |
|---|---|---|---|
| | Leverage | log.Assets.Total | log.DI.Oil.Prod.Total.Sum |
| $\text{Leverage}_{t-1}$ | 0.5134*** | 0.0243 | 0.0398 |
| | (0.0961) | (0.0365) | (0.0704) |
| $\text{log.Assets.Total}_{t-1}$ | −0.0621 | 1.0763*** | 0.0263 |
| | (0.0615) | (0.0575) | (0.1116) |
| $\text{log.DI.Oil.Prod.Total.Sum}_{t-1}$ | 0.0028 | −0.0074** | 0.5183*** |
| | (0.0047) | (0.0036) | (0.1021) |
| $\text{Leverage}_{t-2}$ | 1.1964*** | −0.0091 | −0.0077 |
| | (0.1365) | (0.0143) | (0.0336) |
| $\text{log.Assets.Total}_{t-2}$ | 0.4468* | 0.1266* | −0.1010 |
| | (0.2554) | (0.0700) | (0.1769) |
| $\text{log.DI.Oil.Prod.Total.Sum}_{t-2}$ | 0.0019 | −0.0028 | 0.1086 |
| | (0.0074) | (0.0049) | (0.0686) |
| $\text{Leverage}_{t-3}$ | −1.1439*** | −0.0593 | −0.1079 |
| | (0.1358) | (0.0686) | (0.1454) |
| $\text{log.Assets.Total}_{t-3}$ | −0.5989 | −0.3822*** | −0.2605 |
| | (0.4013) | (0.1299) | (0.2288) |
| $\text{log.DI.Oil.Prod.Total.Sum}_{t-3}$ | 0.0141 | −0.0072 | 0.1450 |
| | (0.0209) | (0.0080) | (0.1346) |
| $\text{Leverage}_{t-4}$ | 0.5292*** | 0.0155 | 0.0651 |
| | (0.0755) | (0.0425) | (0.0972) |
| $\text{log.Assets.Total}_{t-4}$ | 0.2548 | 0.1332* | 0.3387 |
| | (0.1931) | (0.0765) | (0.2155) |
| $\text{log.DI.Oil.Prod.Total.Sum}_{t-4}$ | −0.0038 | 0.0070 | 0.1043 |
| | (0.0114) | (0.0072) | (0.1255) |
| log.Last.Quarterly.WTI.Spot.Price | −0.0618 | 0.0726*** | 0.0175 |
| | (0.0430) | (0.0185) | (0.0599) |
| const | 0.0702 | −0.0547 | −0.6019 |
| | (0.0603) | (0.0653) | (0.5035) |
| Observations | 3960 | | |
| Number of Groups | 253 | | |
| Avg. Obs. Group | 15.65 | | |
| Min. Obs. Group | 1 | | |
| Max. Obs. Group | 64 | | |

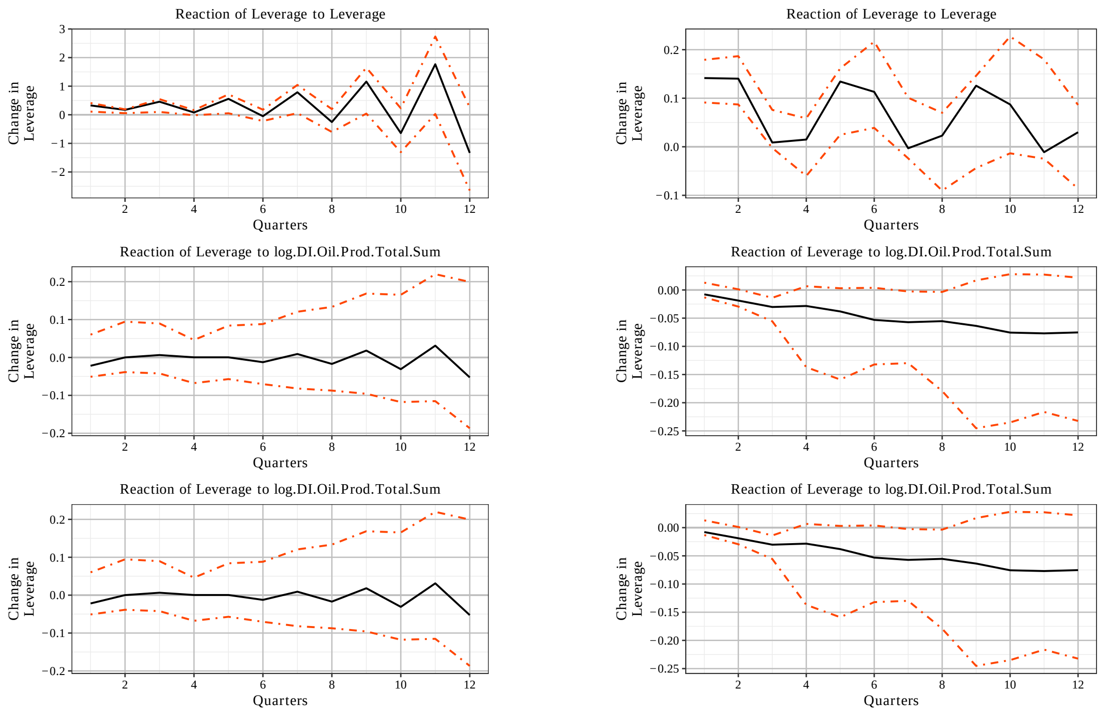
*Note:* *p<0.1; **p<0.05; ***p<0.01
Corrected standard errors are reported in parentheses.
Variable transformation: Forward Orthogonal Deviation

Table 11: Results of the panel VAR approach for the subsample with low production from unconventional oil sources.

| | Dependent Variables | | |
|---|---|---|---|
| | Leverage | log.Assets.Total | log.DI.Oil.Prod.Total.Sum |
| $\text{Leverage}_{t-1}$ | 0.9834*** | −0.0072 | −0.0836 |
| | (0.0357) | (0.0381) | (0.0525) |
| $\text{log.Assets.Total}_{t-1}$ | −0.0285 | 0.9670*** | −0.0079 |
| | (0.0400) | (0.0784) | (0.0334) |
| $\text{log.DI.Oil.Prod.Total.Sum}_{t-1}$ | 0.0108* | 0.0020 | 0.9333*** |
| | (0.0065) | (0.0109) | (0.1361) |
| $\text{Leverage}_{t-2}$ | −0.9516*** | −0.0242 | 0.0510 |
| | (0.0481) | (0.0314) | (0.0377) |
| $\text{log.Assets.Total}_{t-2}$ | −0.1075 | 0.1148** | 0.1351** |
| | (0.0951) | (0.0563) | (0.0646) |
| $\text{log.DI.Oil.Prod.Total.Sum}_{t-2}$ | 0.0006 | −0.0020 | 0.3628** |
| | (0.0106) | (0.0193) | (0.1502) |
| $\text{Leverage}_{t-3}$ | 1.0030*** | 0.0191 | 0.0020 |
| | (0.0744) | (0.0577) | (0.1356) |
| $\text{log.Assets.Total}_{t-3}$ | 0.1897*** | −0.0044 | −0.0705 |
| | (0.0709) | (0.0638) | (0.0758) |
| $\text{log.DI.Oil.Prod.Total.Sum}_{t-3}$ | −0.0269 | 0.0276 | −0.2792*** |
| | (0.0267) | (0.0196) | (0.0687) |
| $\text{Leverage}_{t-4}$ | −0.0968 | −0.0306 | −0.0396 |
| | (0.1395) | (0.0897) | (0.2278) |
| $\text{log.Assets.Total}_{t-4}$ | −0.0488 | −0.1110*** | −0.0859 |
| | (0.0366) | (0.0390) | (0.0660) |
| $\text{log.DI.Oil.Prod.Total.Sum}_{t-4}$ | 0.0197 | −0.0233 | −0.0033 |
| | (0.0214) | (0.0184) | (0.0907) |
| log.Last.Quarterly.WTI.Spot.Price | −0.0071 | 0.0383*** | 0.0929*** |
| | (0.0065) | (0.0096) | (0.0239) |
| const | 0.0320 | 0.1310** | −0.1464 |
| | (0.0417) | (0.0513) | (0.1134) |
| Observations | 4413 | | |
| Number of Groups | 228 | | |
| Avg. Obs. Group | 19.36 | | |
| Min. Obs. Group | 1 | | |
| Max. Obs. Group | 66 | | |

*Note:* *p<0.1; **p<0.05; ***p<0.01
Corrected standard errors are reported in parentheses.
Variable transformation: Forward Orthogonal Deviation

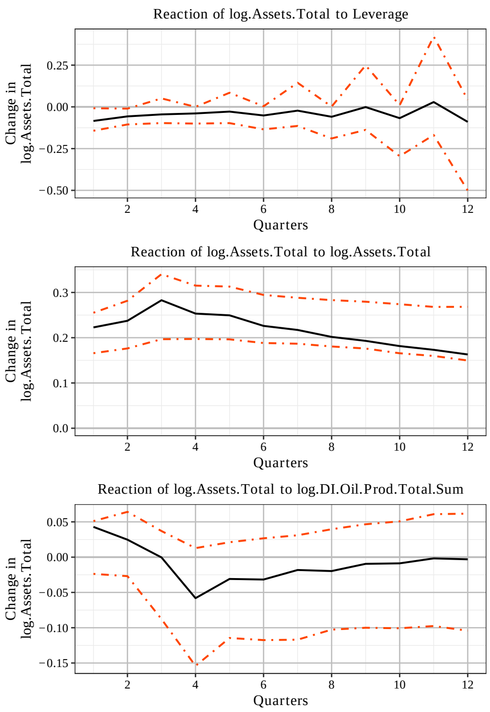
Table 12: Results of the panel VAR approach for the subsample with high production from unconventional oil sources.



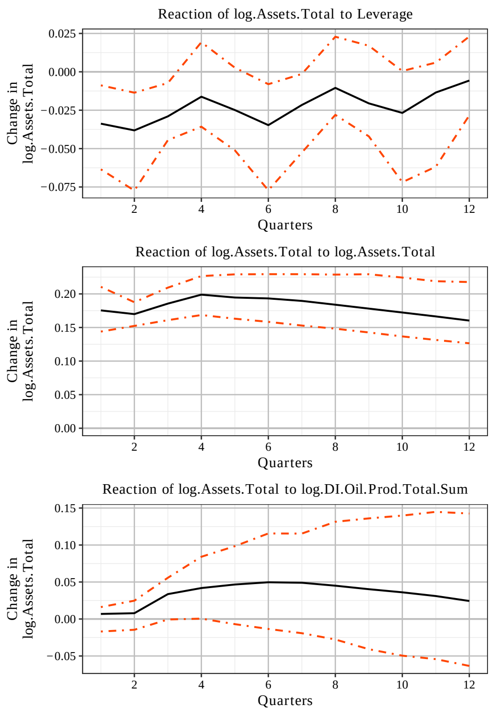

(a) Subsample with a **low** share of production from unconventional sources.

(b) Subsample with a **high** share of production from unconventional sources.

Figure 15: Generalized IRF for the impact on the leverage ratio. Dashed red lines indicate the bootstrapped 95% confidence intervals based on 500 iterations.
Source: Own calculations

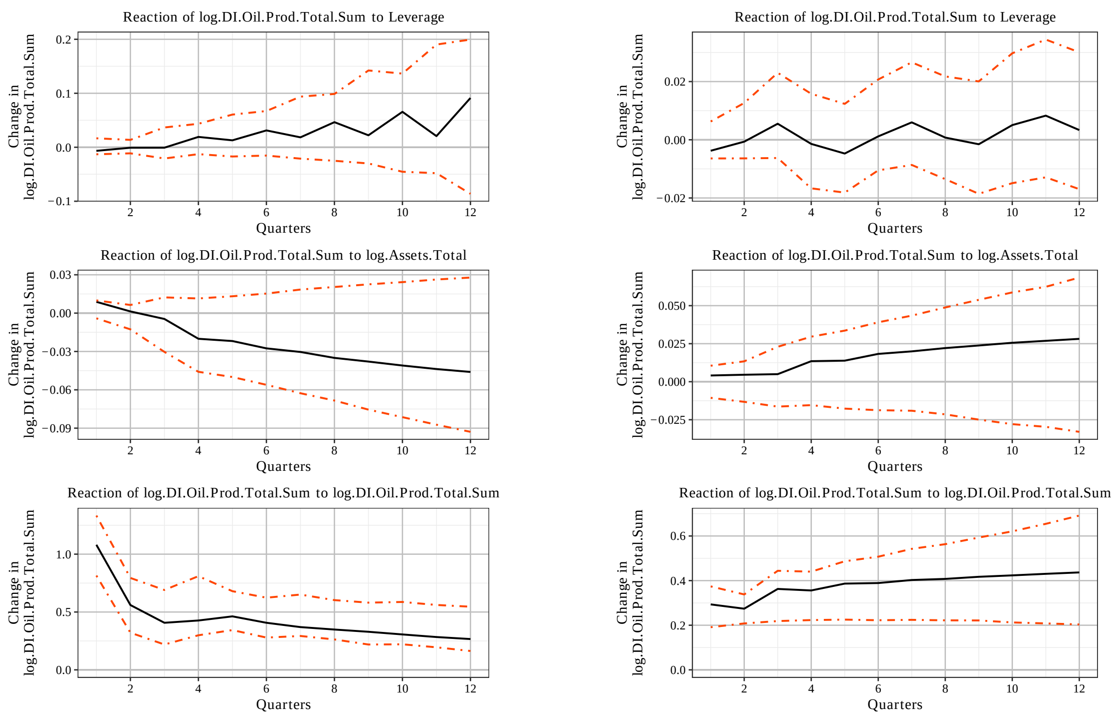



(a) Subsample with a **low** share of production from unconventional sources.




(b) Subsample with a **high** share of production from unconventional sources.

Figure 16: Generalized IRF for the impact on the assets. Dashed red lines indicate the bootstrapped 95% confidence intervals based on 500 iterations.
Source: Own calculations




(a) Subsample with a **low** share of production from unconventional sources.

(b) Subsample with a **high** share of production from unconventional sources.

Figure 17: Generalized IRF for the impact on the oil production. Dashed red lines indicate the bootstrapped 95% confidence intervals based on 500 iterations.
Source: Own calculations

# References

EIA (2017a): *Henry Hub Natural Gas Spot and Futures Price*. URL: `https://www.eia.gov/dnav/ng/ng_pri_fut_s1_d.htm`.

– (2017b): *WTI Spot and Futures Prices*. URL: `https://www.eia.gov/petroleum/data.php#prices`.

# Chapter 4 Oil Price Shocks and Cost of Debt – Evidence from Oil Firms

**Reference for this Paper:**

Funk, C., K. Kempa and J. Lips (2019): "Oil Price Shocks and Cost of Debt – Evidence from Oil Firms". In: *Working Paper*

**Conferences with Review Process:**

- 11th CFE-CMStatistics 2018, Pisa, December 2018. (Presentation by Christoph Funk)
- 8th Energy Finance Christmas Workshop (EFC18), Bolzano, December 2018. (Presentation by Christoph Funk)
- 7th International Symposium on Environment and Energy Finance Issues (ISEFI), Paris, May 2019. (Presentation by Johannes Lips)
- 16th IAEE European Conference, Ljubljana, August 2019. (Presentation by Christoph Funk)

# Oil Price Shocks and Cost of Debt – Evidence from Oil Firms *

Christoph Funk[†‡] Karol Kempa[§] Johannes Lips[†¶]

This version: November 2019

**Abstract**

This paper empirically analyses the relationship between oil price and the costs of debt of firms in the United States (US) oil industry, which we differentiate along the oil industry's value chain. In particular, we analyse how oil price, its volatility, specific shocks affect an oil firm's borrowing decisions and creditworthiness perceived by banks and capital markets. We collect data on (i) individual syndicated loans taken, (ii) bonds issued, and (iii) bond trades on the secondary markets and combine this data with information from these firms' corporate financial statements. We analyse how a firm's characteristics, e.g. firm size, profitability, leverage/indebtedness, affect the credit spread of loans and bonds, i.e. the cost of debt. In addition to these firm characteristics, we investigate the impact of the oil price, its volatility, and oil price shocks – considering both the 2008 and the 2014 oil price shock – on the firms' costs of debt. Overall, we find some evidence for an effect of the oil price itself, which negatively affects the credit spreads of corporate bonds on secondary markets. In particular, we find that oil price volatility plays an important role for the costs of debt. Higher volatility means higher uncertainty in the oil market and, as a consequence, banks and capital markets charge higher prices for debt raised by oil firms. The 2008 and 2014 oil price shocks seem to have mainly led to credit rationing.

**Keywords:** Corporate Finance, Debt, Energy Economics, Leverage, Oil Industry

**JEL classification:** C33, C58, G01, G30, Q40

*We are grateful to Peter Winker, the participants of the 8th Energy & Finance Christmas Workshop, Bolzano 2018, the 12th International Conference on Computational and Financial Econometrics, Pisa 2018 and the 7th International Symposium on Environment and Energy Finance Issue, Paris 2019 for helpful comments and support.
[†]Department of Economics, Justus-Liebig University of Giessen, Licher Str. 64, Giessen, Germany
[‡]Department of Actuarial Studies and Business Analytics, Macquarie University, Sidney, Australia
[§]Department of Economics, Frankfurt School of Finance & Management, Frankfurt/Main, Germany
[¶]Corresponding author: Johannes.Lips@wirtschaft.uni-giessen.de

# 1. Introduction

The effects of oil price shocks on the (world) economy have been extensively studied over the last decade (Hamilton 2009; Kilian 2009; Kilian and Vigfusson 2011; Ravazzolo and Rothman 2013). Most of these studies focus on the effect of oil prices on macroeconomic aggregates such as, e.g. real GDP or real consumption spending. Yet, comparably little is known of the effects of these shocks on a firm level basis in the oil industry, in particular with respect to the cost of debt. This is, however, of high relevance in the sector as debt is the main external source of finance used for financing real investments and maintaining flexibility in operations (Valta 2012). This is particularly relevant for the capital-intensive oil industry. At the same time, oil price fluctuations and shocks are likely to have a high impact on firms' risks of default and thus on their costs of debt. This paper aims to fill this gap by empirically analysing the impact of oil prices on the spreads of bank loans and corporate bonds of oil firms in the US.

Falling oil prices and shocks have several implications for the oil industry. First, oil firm's revenues decrease, which might also increase uncertainty around future oil prices and earnings. The uncertainty might further increase when price volatility is high. Second, assets backing a firm's debt might devaluate. The effects on firm production and investment, however, are not clear. Firms with relatively high costs might reduce production or completely shut down (Sengupta et al. 2017). At the same time, firms with not fully used capacities and efficiencies might increase production to make up for the downward pressure of low prices on revenues (Cakir Melek 2015). Moreover, highly leveraged oil producers are likely to cut down their investment and increase production to generate a cash flow that is needed to fulfil their debt obligations, which could be absorbed in the aftermath of the 2014 oil price shock (Lehn and Zhu 2016; Lips 2019). Rodziewicz (2018) shows that investments in the energy sector fell dramatically and were a drag on the aggregate investment in the US. The higher risk to default is also likely to lead to increasing financing costs for oil firms (Domanski et al. 2015).

In spite of these channels through which the oil price might affect a firm's risk of default and thus its costs of debt, there is almost no empirical evidence on this issue. Debt financing and its cost is, however, essential for firms in the US oil industry and its importance is particularly pronounced in more recent years. Azar (2017) argues that access to relatively cheap debt in the aftermath of the financial crisis was a key fact enabling investments in new technologies to participate in the 'shale oil revolution'. In this time period, the indebtedness of US oil firms increased substantially. We fill this gap by empirically examining the relationship between oil price, its volatility and specific shocks and the costs of debt of US oil firms.

We investigate whether oil prices, in addition to directly affecting oil firms' sales revenues, have an impact on the price a firm has to pay to raise new debt. In general, firms can choose to raise debt through bank loans or on the capital markets by issuing bonds. In our analysis, we consider both forms of debt financing, which is an extension of previous studies, e.g. Sengupta et al. (2017). Hence, we capture both the corporate bond and the banking market that are both frequently used by US oil firms. This allows us to compare whether banks and the capital market evaluate the effect of oil prices on the creditworthiness of oil firms differently. Furthermore, we can explicitly check, whether certain effects might be driven by specifics of the banking sector or debt markets.

We collect data on (i) syndicated loans taken and (ii) bonds issued by US oil firms as well as (iii) bonds traded on the secondary market. We then combine this data with information from these firms' corporate financial statements. Thus, we can analyse how a firm's (financial) characteristics, e.g. firm size, profitability, or leverage/indebtedness, affect the credit spread of loans and bonds, i.e. the cost of debt. In addition to these firm characteristics proposed in the financial economics literature (Chava et al. 2009; Chen et al. 2007; Dennis et al. 2000; Goss and G. S. Roberts 2011; Valta 2012), we consider the oil price, oil price volatility, and shocks. In our analysis, we distinguish between the 2008 and the 2014 oil price shocks' effects on the firms' costs of debt, as they differ in various aspects. Compared to the 2014 shock, the 2008 oil prices shock coincides with a financial crisis and recession. We further differentiate firms along the oil industry's value chain. Finally, we also control for the macroeconomic environment.

Overall, we find that, even after controlling for loan/bond and firm characteristics as well as the macroeconomic environment, the oil price and its volatility, have an effect on a firm's cost of debt. Thus, in addition to directly affecting oil firms' sales revenues, particularly oil price volatility positively affects the price a firm has to pay to raise new debt. Both banks and the bond market seem to consider a high price volatility a risk that increases uncertainty and thus reduces the creditworthiness of oil firms. As a result, banks and the capital market demand higher credit spreads. For credit spreads of bonds traded on the secondary market, we further find a strong negative effect of the oil price. We further find some evidence that oil price shocks rather led to credit rationing, i.e. firms not receiving debt, rather than higher costs of debt.

The remainder of the paper is organised as follows. Section 2 reviews the literature on the determinants of firms' costs of debt and presents specifics of debt in the oil industry. In Section 3, we describe the data set in detail. Section 4 first provides an exploratory data analysis and then presents the estimation approaches and their results. Section 5 concludes.

# 2. Debt Financing and the Oil Industry

## 2.1. Determinants of Sources and Costs of Debt

Firms have two main ways of obtaining external debt, namely through issuing corporate bonds or taking (syndicated) bank loans. Numerous previous studies analysed both drivers for choosing between both types of debt and, closer related to this paper, determinants of the characteristics of these bonds and loans. In a perfect capital market, all firms would be able to obtain funding for all investments with a positive net present value. In a market with frictions, such as information asymmetry, potential lenders are not able to evaluate a firm's quality, at least not without incurring costs, which might lead to credit rationing (Stiglitz and Weiss 1981).

Becker and Ivashina (2014) investigate firms' decisions to substitute between loans and bonds to raise external funds. Their results of the firm-level analysis indicate that substitution of loans by bonds is particularly pronounced in times of tight lending markets, poor performance of the banking sector and tight monetary policy. In interpreting these results one has to keep in mind that the sample of firms is comprised of large firms, which can easier tap public debt markets. This is also highlighted by Lemmon and M. R. Roberts (2010) who, by analysing an exogenous shock to the supply of below-investment-grade credit after 1989, only observe a limited substitution away from bank debt to alternative sources of capital. Boneva and Linton (2017) investigate how the costs of funding in the corporate bond markets affect issuance decisions and especially focus on the importance of the transmission mechanism of monetary policy. The authors identify that the negative relationship between corporate bond yields and their issuance is driven by firms with low credit ratings. This effect is particularly strong in the aftermath of the financial crisis, which might indicate that it was probably more difficult for firms to obtain loans and the bond market offered a viable alternative for firms to fulfil their financing needs (Farrant et al. 2013).

Greenwood et al. (2010) explain the time variation in the maturity of corporate debt with changes in the maturity structure of government debt arguing that firms absorb supply shocks initially caused by the maturity choices of the government. This theoretical consideration is empirically tested in Greenwood and Vayanos (2014) and Badoer and James (2016), where the latter is especially interested in the very long-term corporate borrowing of 20 years or more. In a closely related paper, Graham et al. (2014) analyse the relationship between the US government's fiscal policy and corporate financing decisions. They also highlight the empirical observation that firms do not switch between different sources of funding, which is consistent with, e.g. Faulkender and Petersen (2006) and Leary (2009) and indicates a segmentation of financial markets.

Faulkender and Petersen (2006) provide evidence for restrictions for the substitution between private and public debt in the presence of market frictions. The authors show that access to public debt markets, proxied by the presence of a credit rating, typically have a higher leverage compared to firms that have to borrow from banks. Furthermore, firms with a credit rating seem to mainly issue public debt. However, ratings seem to also matter for private debt. According to Sufi (2009) the introduction of ratings for syndicated loans led to increased debt issuance and investment by riskier borrowers. This suggests that these ratings were effective in reducing the informational frictions that generate segmentation as noted by Faulkender and Petersen (2006).

Several articles focus on the role of different lender types in loan syndicates. In the presence of information asymmetries, lead banks in loan syndicates retain a higher share of the loan (Sufi 2007). Ivashina (2007, 2009) empirically show that, controlling for borrowers' characteristics, a higher retained share of the lead bank seems to work as a signalling device mitigating the information asymmetry problem and consequently lowering the loan spread. Lim et al. (2014) find that the inclusion of non-bank institutions in the syndicate raises the loan spread. The authors argue that these institutions have a higher required rate of return than banks. Additionally, non-bank premiums are larger if the borrowing firms are facing financial constraints and the capital supply by banks is curtailed.

There is also evidence for the effect of equity risk on a firm's cost of debt. Campbell and Taksler (2003) identify a firm's equity volatility as a major factor, which explains a sizeable share of the cross-sectional variation of corporate bond yields. This relationship is based on theoretical explanation of Merton (1974) that bond owners have issued put options to the equity holders and thus, both idiosyncratic and market volatility affects the value of this put option. Campbell and Taksler (2003) can provide evidence for the link between equity risk and a widening spread of corporate bond relative to Treasury bonds. Chen et al. (2007) stress the importance of bond volatility for corporate yield spreads. The authors find that a higher liquidity results in a lower spread, even when controlling for bond-specific, firm-specific, and macroeconomic variables affecting the default risk of the issuing firm. Dick-Nielsen et al. (2012) show that the effect of liquidity on corporate bond spreads increased substantially during the subprime crisis.

## 2.2. Specifics of the Oil Industry

In the past two decades, the US oil market has experienced substantial changes and dramatic events. After a long lasting decline since the 1980s, the oil production increased sharply after the financial crisis in 2008 (EIA 2017), which is mainly driven by the so called 'shale oil revolution' (Baumeister and Kilian 2016). The US Energy Information Administration (EIA) estimates that, as of 2016, 48% of total US oil production is attributable to shale oil. Connected with this, Domanski et al. (2015) describe an interesting phenomenon: a contemporaneous increase in debt-driven investments in the oil sector since the shale oil revolution started. This growth in debt was driven by a macroeconomic environment with low interest rates and investors searching for profitable investments after the financial crisis. This was also reinforced by the impact the Federal Reserve's Quantitative Easing (QE) had on the corporate bond market (Krishnamurthy and Vissing-Jorgensen 2011).

Particularly between January 2011 and June 2014, the crude oil price performed a comparatively stable sideways shift closely followed by a major decline until January 2015 (Figure 1). On the one hand, this event might be a supply shock based on the US shale oil boom and the refusal of the Organization of the Petroleum Exporting Countries (OPEC) to reduce supply quantities. On the other hand, there exists evidence that the price depreciation was driven by a demand shock based on a slowdown of the world economy (Baumeister and Kilian 2016). In addition, the appreciation of the US dollar could have lowered the demand even further as dollar denominated crude oil imports became more expensive (Baffes et al. 2015). Interestingly, supply increased even further despite the oil price decline in 2014. This seems to be counter intuitive, as one would expect a cut in production as observed after the oil price shock in the aftermath of the financial crisis in 2008.

The oil price, its volatility, and oil price shocks are likely to affect firms' cost of debt. In general, the cost of debt should depend on the default risk of a firm and the costs the bank incurs in the case of a default (Valta 2012). In the case of imperfect contracts and transaction

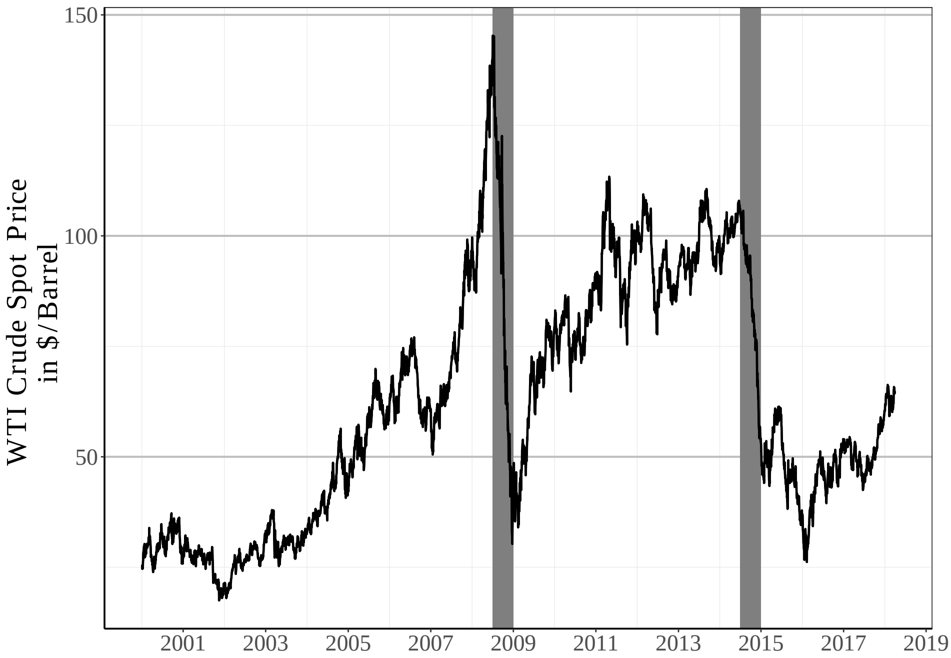



Figure 1: Development of WTI crude oil spot price.

costs exist, liquidation values play a major role for financial contracts between borrowers and lenders (Bolton and Scharfstein 1996; Hart and Moore 1994). This value is important for the interest rate on a loan, since creditors have the right to possess the assets, when a firm defaults on its debt payments: the lower the liquidation value, the higher the lender's costs when a firm defaults. Thus, one channel through which an oil price shock might affect the cost of debt is its effect on the liquidation value of a firm's assets. As assets of oil firms are very specific, adverse oil price shocks directly affect their liquidation value and consequently the cost of debt.

These theoretical considerations are supported by empirical evidence. Benmelech et al. (2005) analyse empirically how liquidation values affect the optimal debt policy. Since creditors anticipate how costly it will be for them to seize and liquidate the assets of a firm. The authors find that a higher asset liquidation value leads to lower costs of debt. In light of the limited redeployability in the oil and gas extraction industry, as noted by Kim and Kung (2016), an oil price shock affecting the whole upstream industry should increase the cost of debt for firms in this industry segment. The authors also find that uncertainty affects the investment decisions of firms differently depending on their assets' redeployability. However, the authors cannot disentangle possible interactions between the supply and demand for external financing, because they are only looking at the investment outcome. In cross-section and time-series tests, Ortiz-Molina and Phillips (2014) find that real asset illiquidity affects the cost of capital. Firms with less liquid real assets face higher cost of capital than firms with more liquid assets and the cost of capital also increases in periods with particularly illiquid real assets.

Benmelech and Bergman (2011) show that bankruptcies of industry peers lead to a contagion effect, by reducing the value of the collateral of other firms and thereby increasing their cost of debt financing. This is likely to be more pronounced, when only focussing on one part of the oil industry's value chain. As most assets are quite distinct for each part of the oil industry's value chain, distress should be relatively contained within each part of the chain.

In addition to affecting the asset value and thus the lenders' costs in the case of default, oil price changes also affect an oil firm's probability of default by affecting, e.g. firms' profits. Kinda et al. (2016) investigate the impact of commodity price shocks on financial sector fragility and find that negative price shocks increase the number of non-performing loans. The authors, however, do not analyse the effect of commodity price shocks on the loan spreads of individual loans. Such an analysis was conducted by Sengupta et al. (2017), who analyse the effect of oil price shocks on credit spreads of syndicated loans to US oil firms and found increasing loan spreads for upstream and support services firms in the aftermath of the 2014 oil price shock.

Overall, theory and previous evidence suggest that oil prices are likely to be an important determinant of the oil firms' cost of debt. In the presence of information asymmetries in financial markets, a firm's decision and ability to raise private or public debt depends on its characteristics. Larger and listed firms are usually less informationally opaque and thus more likely to be able to tap capital markets in order to raise debt. Smaller and riskier firms are more likely to borrow from banks, who specialise on screening potential lenders prior to approving credits. By analysing both loans received and bonds issued by oil firms, we provide a first comprehensive empirical examination of the determinants of oil firms' costs of debt.

# 3. Data Set and Variables

The data used in the empirical analysis can be split up into four categories: (i) characteristics of borrowing firms, (ii) bond and bank loan data, (iii) oil price data, and (iv) data on the macroeconomic conditions. This subsection presents all data sources and describes the construction of the variables used in the empirical analysis. Table 8 in the Appendix contains the definitions of the variables used in the analysis.

## 3.1. Oil Firms

In order to analyse the whole value chain of the oil industry in the US, we identify oil firms and classify them along the value chain: upstream, midstream, downstream and support services. Given the different features of firms along the value chain, it is to be assumed that the oil price might not affect all firms in the same way. Upstream firms are responsible for the exploration and production of oil and gas resources. They identify possible oil field, drill wells and extract the resources from underground. As such, one can expect that upstream firms' revenues and their credit conditions to be most affected by changes in the crude oil price. Midstream firms link upstream with downstream entities by transporting the oil resources through pipelines and gathering systems. Their revenue streams are fee-based and typically tied to long-term contracts which makes midstream firms less affected by oil price volatility in the short run (Sengupta et al. 2017). Downstream firms engage in the refining of crude oil and the marketing, distributing and selling of processed petroleum products (Rodziewicz 2018). Refiners face a double risk between the raw materials market and petroleum products market which largely determines their profitability. Thus, downstream firms engage in hedging strategies to reduce their sensitivity to oil price fluctuations (Ji and Fan 2011). Support services engage with all firms of the supply chain. Although these firms derive mostly stable cash flows from mid- and upstream firms, they are closely connected to upstream entities as they provide oilfield services, equipment and drilling site preparation (Rodziewicz 2018). Similar to the upstream industry, we expect the credit availability and profitability to react sensitively to changes in oil price.

The firms were selected based on their Standard Industrial Classification (SIC) and North American Industry Classification System (NAICS) classification. First, we identified all relevant industry classifications and assigned each of them to one of the four categories along the value chain. The resulting 31 SIC and 22 NAICS codes were used to identify North American oil firms[1] in the Compustat – Capital IQ database.[2] The Compustat – Capital IQ database contains detailed quarterly financial data of publicly listed firms, which we use to control for firms' financial situation and risk of default. We use the following firm-level control variables based on the literature (Chava et al. 2009; Chen et al. 2007; Dennis et al. 2000; Goss and G. S. Roberts 2011; Valta 2012). We use the natural logarithm of a firm's total assets, *log(Total Assets)*, as a measure for firm size. We further control for a firm's *Profitability*, which is the EBITDA relative to total assets. The variable *Leverage* is the sum of short-term and long-term debt divided by total assets.

Table 1: Summary Statistics - Full Sample for all Firms. All monetary variables are in million US dollars.

| Full Sample | n | $\varnothing T$ | mean | sd | min | Q0.25 | Q0.5 | Q0.75 | max |
|---|---|---|---|---|---|---|---|---|---|
| Total Assets | 1677 | 30.96 | 6747.00 | 27234.33 | -12.67 | 56.96 | 549.29 | 3277.18 | 419648.00 |
| Total Debt | 1677 | 30.96 | 1706.44 | 5736.95 | 0.00 | 3.70 | 126.078 | 1056.40 | 138237.00 |
| Leverage | 1677 | 30.96 | 0.86 | 14.61 | 0.00 | 0.11 | 0.27 | 0.40 | 1302.00 |
| Profitability | 1653 | 28.76 | -0.25 | 12.76 | -2180.75 | 0.00 | 0.03 | 0.05 | 511.00 |
| Loan Credit Spread | 591 | 5.55 | 184.80 | 139.80 | 12.50 | 100.00 | 150.00 | 239.21 | 1325.00 |
| Loan Amount | 592 | 5.58 | 822.53 | 1375.37 | 2.00 | 195.00 | 400.00 | 909.41 | 29762.75 |
| Loan Maturity | 592 | 5.58 | 43.21 | 21.48 | 1.00 | 26.00 | 48.00 | 60.00 | 324.00 |
| Bond Credit Spread | 234 | 30.49 | 2.63 | 2.26 | -10.66 | 1.30 | 2.10 | 3.70 | 11.55 |
| Bond Amount | 227 | 5.14 | 772.93 | 989.27 | 0.00 | 300.00 | 500.00 | 800.00 | 11000.00 |
| Bond Maturity | 235 | 30.56 | 191.17 | 99.82 | 37.50 | 112.56 | 177.00 | 252.00 | 779.00 |

[1] According to the data descriptions this covers North American (US and Canada) firms, which were publicly held and were active in the US over the period analysed.

[2] Please see the appendix for a complete list of the SIC and NAICS codes and their assigned industry classification in the supply chain.

## 3.2. Bank Loans and Bonds

The data on syndicated loans is obtained from the Thomson Reuters' Dealscan database. This database contains a comprehensive overview on the characteristics of syndicated loans, like pricing, contract details and additional terms and conditions. Besides information on the loan characteristics, the database also includes data on the different lenders participating in the syndication of a loan. Chava and M. R. Roberts (2008) provide a detailed introduction to the Dealscan data and also emphasise the good coverage of the US syndicated loan market. The costs of bank loans is measured using the log of the *Loan Credit Spread*, i.e. the Dealscan variable all-in-drawn spread (Chava 2014; Chava et al. 2009; Sengupta et al. 2017; Valta 2012). This variable measures the amount the borrowing firm pays in basis points over the London Interbank Offering Rate (LIBOR) or equivalent for each dollar drawn including annual fees paid to the syndicate.[3]

Since syndicated loans represent only a certain and relatively small part of corporate debt, we additionally use bond data obtained from the enhanced version of the Trade Reporting and Compliance Engine (TRACE) database provided by the Financial Industry Regulatory Authority (FINRA).[4] This database was introduced in 2001 to enhance the transparency in corporate bond markets. The initial phase of TRACE was implemented in July 2002, which is also the earliest date for which transaction data is available. All members of the FINRA are obliged to report their over-the-counter (OTC) transactions of fixed-income securities. We use the bond data in two different ways. The monthly data on the constant maturity yields of the US Treasury securities, which we use to calculate bond credit spreads, were obtained from the Federal Reserve Economic Data (FRED) database of the Federal Reserve Bank of St. Louis[5]. Besides data from the FRED database the TRACE data is enhanced with base data from Bloomberg, which has additional information on the issuance of the loan. For example this data provides information on the initial rating, the coupon rate, the maturity and the issued amount of the loan.

In a first approach, we only consider the costs of debt at the time of issuance. We calculate the *Bond Credit Spread at Issuance* as the difference between the coupon of the bond and the yield of a US Treasury bond with a similar or the closest matching maturity available. This approach makes the bond analysis very comparable to bank loans, where the price is also determined once at the time the loan is issued.

In a second approach, we exploit the fact that bonds, after their issuance, can be traded on secondary markets. These trades contain further information on the costs of debt and allow to track specific bonds over time. The TRACE enhanced database contains detailed information of all fixed-income transactions on the secondary market. It is possible to see the price, volume

[3]See Berg et al. (2016) for a discussion of the importance of fees for the costs of borrowing in the case of syndicated loans.

[4]The different data sets were combined using the linking suite provided by Wharton Research Data Services (WRDS)

[5]Board of Governors of the Federal Reserve System (2019) via the website: https://fred.stlouisfed.org/categories/115

and yield of the traded security. Thus, the information in this data set allows to track individual bonds that are traded over time and thus observe development of the cost of debt over the horizon of our analysis. In order to do that, it is necessary to clean the data and remove reporting errors. We follow the procedures described in Dick-Nielsen (2009, 2014) and remove all trades, which are subsequently cancelled and only keep the information of the last modification in the database. Following the procedure in Rossi (2014) the return reversals in the price and yield time series are removed as well. This means a trade is eliminated if its price or yield is preceded and followed by a price change of more than 50%. Additionally, a filter is applied, which is also proposed by Rossi (2014) and follows earlier work by Brownlees and Gallo (2006) on high-frequency trading data. To remove outliers the smallest and largest 0.0001% of reported yields of individual trades were removed, which resulted in the removal of only 28 individual trades, which should not affect the overall results of the analysis.

To calculate the cost of debt based on secondary market bond trades, we apply the methodologies employed by Bessembinder et al. (2008) and Li and Richie (2016), i.e. we aggregate the TRACE data per quarter. The aggregation of price and yield data is done using a weighted average with the reported trade volume as weights. Until November 2008 the reported yield of a transaction was calculated by the reporting firms. Since then, the calculations are done by FINRA[6]. We calculate the *Bond Credit Spread on the Secondary Market* as the difference between the weighted bond yield from TRACE and US Treasury securities with the same time to maturity. In case an exact match is not possible, we use the closest maturity available. This is in line with Gilchrist and Zakrajšek (2012), who use secondary market credit spreads to calculate the cost of capital.

The loan- and bond-level data is matched to the firm-level data presented above. The matching between the Dealscan loan data and the Compustat firm-level data was facilitated by using the matching table provided by Chava and M. R. Roberts (2008). The bond data is matched with the firm-level data based on the Committee on Uniform Security Identification Procedures (CUSIP), which is used as unique identifier in both TRACE and Compustat.

## 3.3. Oil Prices

The oil price itself plays a crucial role for the revenue generated and thus, directly influences the creditworthiness of oil firms. Changes in the volatility can influence firms' investment decisions and also increase the perceived risk of investors in the oil industry, since future cash flows might become more uncertain. Gilchrist et al. (2014) are able to show that changes to uncertainty, measured as the standard deviation of the unforecastable daily excess stock return, have an impact on investment activity, mainly through changes of credit spreads. To control for the effect of the oil price and its volatility on the cost of debt, we use the average quarterly West Texas Intermediate (WTI) spot price (*Oil Price*) and the average quarterly oil price volatility

[6]For more details on the content and the data, please see http://www.finra.org/industry/trace/historic-file-layout for all data before 6th February 2012 and http://www.finra.org/industry/trace/historic-data-02062012 for all later dates.

(*Oil Volatility*), measured as the standard deviation. We further use dummy variables to model the two major adverse oil price shocks in 2008 and 2014. In December 2015, the US has lifted its ban on crude oil exports. Presumably, this might have an impact on firms in the value chain of the oil industry. To control for this, we include the log of quarterly US crude oil exports. The data is taken from the EIA[7].

### 3.4. Macroeconomic Environment

To capture the overall risk environment of the economy, we use three different interest rate spread variables. The first variable is the spread between the 3-Month LIBOR based on US dollars and the 3-Month Treasury Bill, commonly known as the TED spread (Federal Reserve Bank of St. Louis 2018a). This variable can be seen as an indicator for perceived credit risk in the overall economy and especially in the banking sector. The second variable is the spread between the corporate bond yield for US firms rated as AAA and Baa by Moody's with maturities as close as possible to 30 years. A widening of this credit spread is also an indicator for current or expected poorer economic conditions (Moody's 2018). To control for the overall state of the economy, we include the *Term Spread* as the difference between the 10-year Treasury yield and the T-bill yield (Federal Reserve Bank of St. Louis 2018b).

## 4. Empirical Analysis

### 4.1. Exploratory Data Analysis

The firms in our sample are not evenly distributed across the oil value chain. The most firms are active in the upstream sector, although the number of firms declined since 2013 (see Figure 8 in Appendix C). Figure 2 provides a comprehensive overview on the average firm size and indebtedness across the four different industry classifications. In the upper panel of the figure, it can be seen that the firms in the downstream industry are the largest in our sample in absolute terms, followed by midstream, upstream, and support services firms. In connection with the information provided in Figure 8, this means that there is a big difference in firm size between the different industry classifications. Although the number of upstream firms in our sample is far larger than downstream or midstream firms, they have less assets in total. The lower panel of Figure 2 depicts the development of the overall debt-to-asset ratio, i.e. the indebtedness, in the industry. The highest indebtedness can be observed for midstream firms, while the debt-to-asset ratios of upstream and support service firms increased the most between 2000 and 2017. This increase is most probably driven by increasing investment in shale oil projects. Firms active in the downstream part of the oil industry's value chain have the lowest debt-to-asset ratio, which, however, increased considerably following the financial crisis.

[7]Monthly volume of US Crude Oil Exports in thousand barrels per day was taken from the EIA website at: https://www.eia.gov/dnav/pet/hist/LeafHandler.ashx?n=pet&s=mcrexus2&f=m.

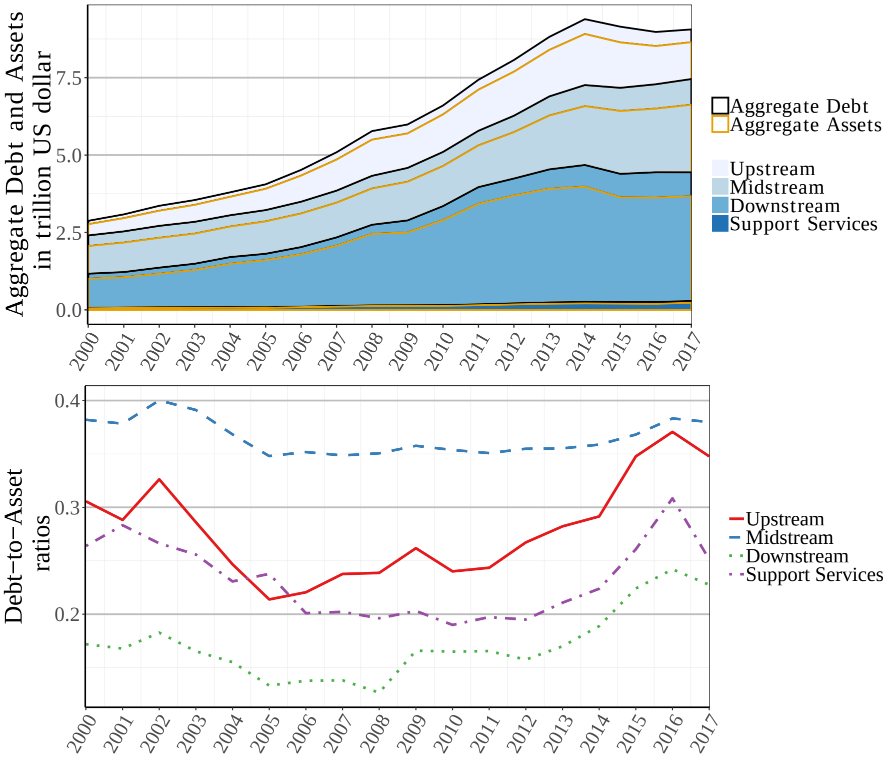



Figure 2: Development of aggregate assets and debt in each part of the value chain and the resulting debt-to-asset ratio.

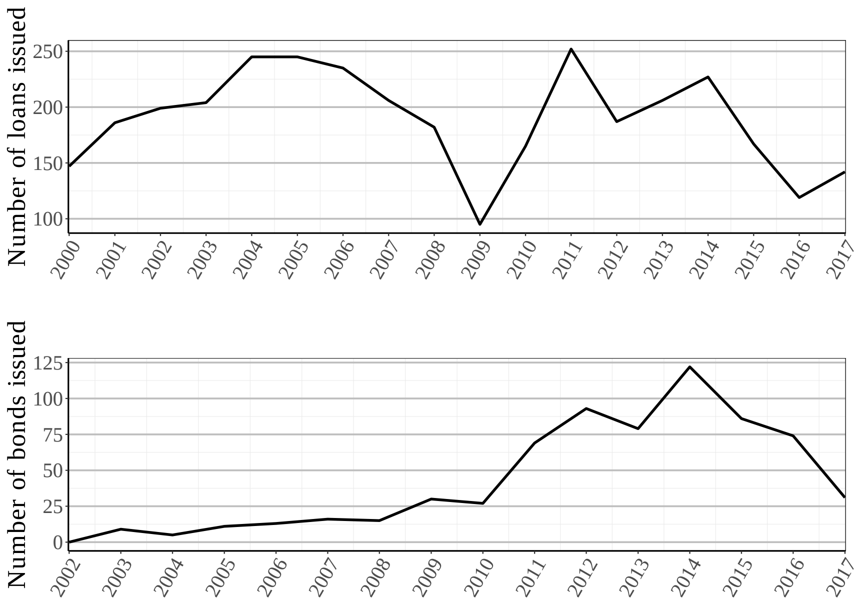



Figure 3: Number of bonds and loans issued per year for all firms in the sample.

Figure 3 depicts the number of loans and bonds issued and reveals distinct differences between the bond market and the market for syndicated loans. First of all, syndicated loans availability seem to have tightened after the oil price declines during the financial crisis in 2008 and after the adverse oil price shock in 2014. This holds especially true for upstream and midstream entities due to a significant decrease of the number of loans issued (see Figure 9 in Appendix C). A similar pattern can be observed in the development of average loan amounts. As depicted in Figure 12 in Appendix C, the average loan sizes across the value chain drastically dropped after the oil price shocks in 2008 and 2014, while they overall show positive trends across the sample period.

The opposite can be observed for the bond market. Corporate bonds played a negligible role before the financial crisis. After 2010, however, the number of bonds issued by oil companies increased notably peaking in 2014. The average volume of bonds issued behaved similarly in that time period (see Figure 13 in Appendix C). An explanation might be that firms substituted bank loans by public debt due to tightened credit conditions. The increase in bond issuance is particularly pronounced in the case of firms in the midstream and downstream sectors (see Figure 10 in Appendix C). This is not surprising for two reasons. First, as discussed in Section 2, the larger size of midstream and downstream firms allows them to borrow money on the corporate bond market, while smaller upstream and support services firms rely more on bank lending. Second, given that the banking sector was in distress after the financial crisis, firms' access to credit was affected by changes in the financial conditions of their banks (Popov and Udell 2012).

Figure 4 depicts the credit spreads of syndicated loans and corporate bonds at issuance and reveals further differences between private and public debt markets. It can be seen that the spread for bonds at issuance exhibits a much larger variation as the spread of syndicated loans. Over the whole time period under consideration, the spreads of the syndicated loans are much more concentrated around the mean and median, whereas the credit spread of bonds at issuance have a higher variance. Figures 11 and 14 in Appendix C disentangle credit spread development for bonds and syndicated across industry classifications. There is a general tendency of midstream firms paying the lowest spreads, followed by downstream, upstream and support services firms.

Figure 5 depicts the continuously calculated credit spread of bonds traded on the secondary market. These bond credit spreads seem to be affected more by the financial crisis and, in particular, the adverse oil price shock in 2014. This effect is particularly strong form downstream and support firms (see Figure 15 in Appendix C). This hints at the possibility of a selection effect in times of economic distress. In a difficult economic environment, some firms that were able to raise debt via bonds or, in particular, loans, might not receive debt in difficult times, as can be seen in the lower number of new loans during the financial crisis. This selection effect towards less risky firms might explain why the credit spreads of issued loans do not increase too substantially during the financial crisis. The credit spread on the secondary market, however, is based on bonds issued in the past and thus shows a stronger reaction to market shocks and recessions or oil price shocks.

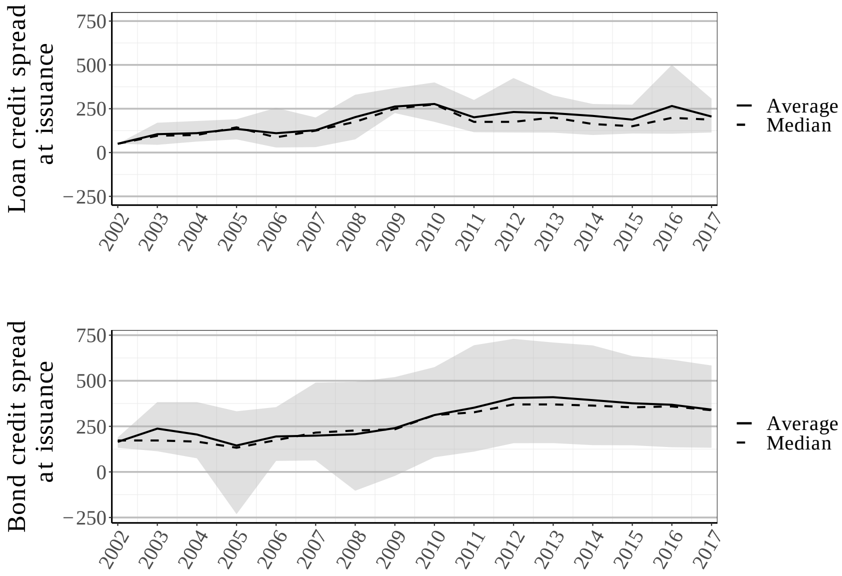



Figure 4: Loan and bond credit spreads at issuance. Shaded area indicates upper (90%) and lower (10%) quantile of the credit spread.

Overall, there seem to be similar patterns between oil firms' debt (decisions) and the oil price. This seems to be particularly strong in the case of bank loans. In periods of rather high oil prices, oil firms seem to borrow more, i.e. the number and the average size of loans are higher, at better conditions, i.e. credit spreads of loans are relatively low. Most of these patters can also be observed for corporate bonds, both at their issuance and for trades on the secondary market. One exception are the average volumes of newly issued bonds, which seem not to decline as pronounced as loan volumes in the aftermath of adverse oil price shocks.

Figure 6 plots our three endogenous variables, (a) the loan credit spread at issuance, (b) the credit spread of bonds at issuance, and (c) the bond credit spread on the secondary market for the full sample. For all three measures of cost of debt, we find a similar pattern: the median credit spreads for downstream and especially midstream firms are lower compared to the upstream and support services sectors. This is particularly pronounced for loans. One possible explanation for this could be that midstream entities operate with long-term contracts, which provides them with a stable cash flow rendering them less risky from the debt providers' perspective.

Figure 11 and 14 provide information about the median credit spread development for TRACE bonds and the spread for syndicated loans from the Dealscan database. There is a general tendency of midstream firms paying the lowest spreads, followed by downstream, upstream and support services firms. The spreads piked in both cases around the global financial crisis and oil price shock in 2008/09 and increased substantially after the oil price shock in 2014.

Finally, we analyse how all the main variables discussed above change across three different time periods. Our sample consists of 1,682 firms from 2000:Q2 to 2018:Q1. So we are covering

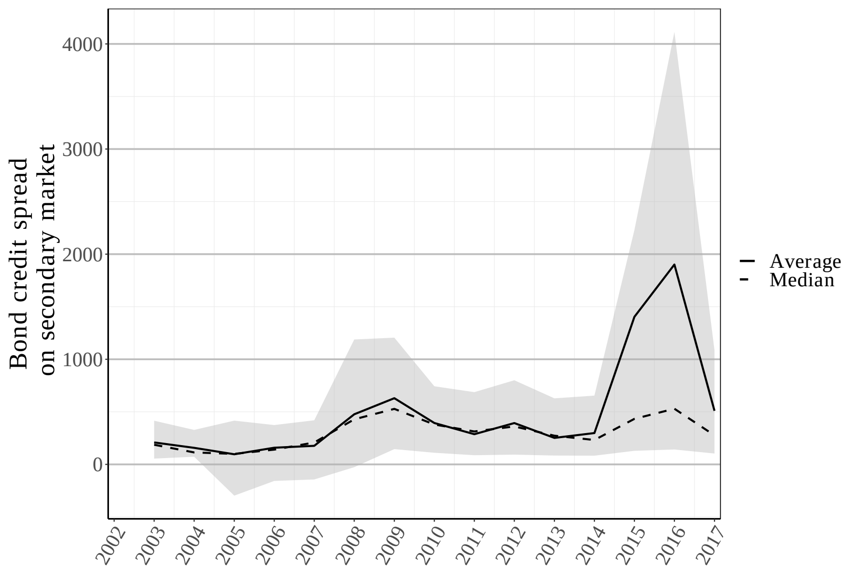
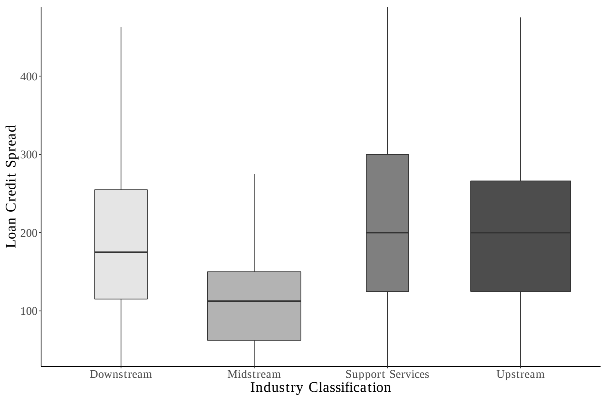



Figure 5: Development of the average and median bond credit spread on the secondary market. Shaded area indicates upper (90%) and lower (10%) quantile of the credit spread.

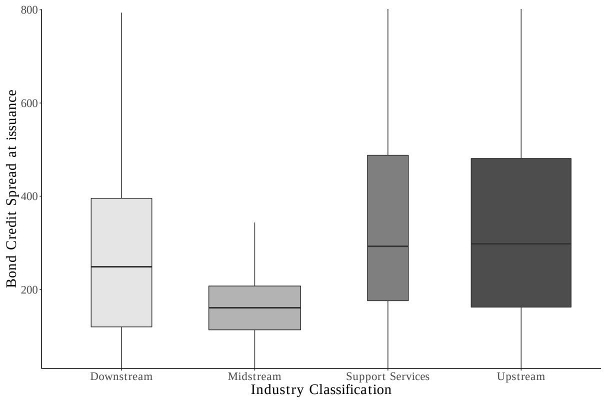
the three time periods, the pre-oil shock period from 2000:Q2 to 2008:Q3, the period in between the two oil price shocks from 2008:Q4 to 2014:Q3 and the post-oil shock period 2014:Q4 to 2018:Q1. In total, 296 firms of these are available over the entire horizon, while 604 firms have at least one Dealscan loan and 275 at least one TRACE bond. Of these 253 have both, so at least one syndicated loan and an issued bond. Only 22 firms do have only bonds, whereas in total 351 firms have only a syndicated loan.

Table 2 summarises the mean values of the main variables for these three time periods.[8] Columns 5 and 6 present the differences in the mean values of these periods. We find that the leverage ratio did not vary after 2008 and almost doubled after the shock in 2014. At the same time, the profitability increased after the global financial crisis and after 2014, however not significantly. The loan credit spread significantly increased after the shock in 2008 by 69 basis points but declined after 2014 by 25 basis points on average. At the same time, the average loan maturity increased by 5.5 months after 2008 and remained at this level throughout. In both periods, the need for financing as measured by the facility amount increased, however especially heavily and significantly after 2014. A similar picture emerges for the bond market. We find a significant increase in the bond credit spread after both 2008 and 2014. In addition, the bond amount increased substantially after 2014. On an aggregated level, we find that for the bonds that the average months to maturity decreased by 37 months after 2008 and 16 after 2014, respectively. While the simple comparison of means by a $t$-test suggests an increase after the financial crisis and an increase for the bond market after the oil price shock in 2014, we do not account for any differences between firms along the supply chain as well as for lender

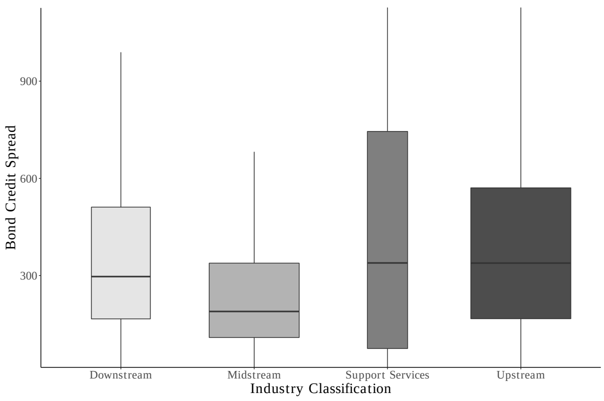
[8] We found similar results for the differences in the samples using only firms that have at least one bond traded on the secondary market or a syndicated loan which are summarised in Table 10 in the appendix.




(a) Loan Credit Spread




(b) Bond Credit Spread at issuance



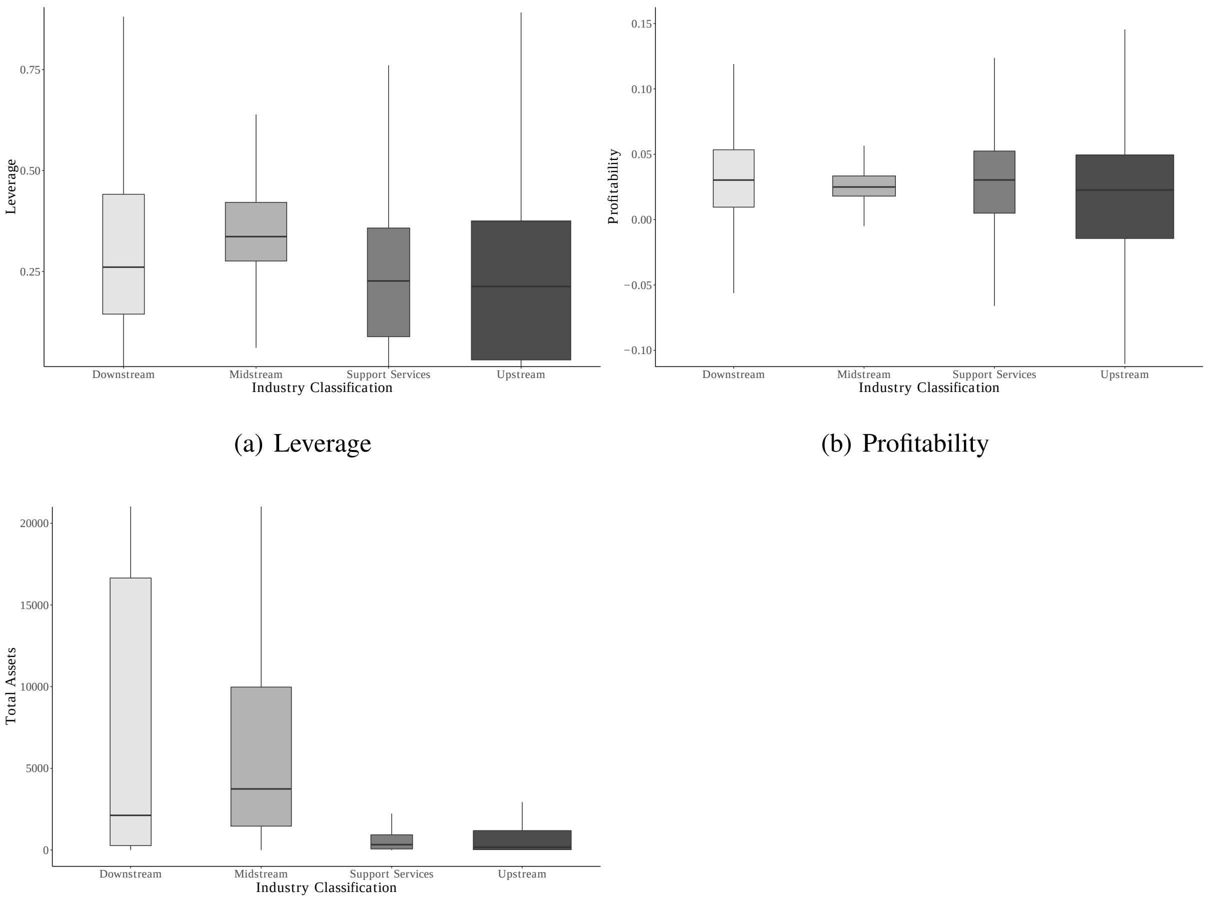

(c) Bond Credit Spread on the secondary market

Figure 6: Boxplots of the endogenous variables

characteristics that have an effect on the credit spreads. To take these firm characteristics into account, we are using different regression analyses described in the following subsections.

## 4.2. Estimation Approach

We employ two different empirical approaches to investigate the effect on oil prices on firms' costs of debt. The first approach is focusing on the costs of debt of individual loans or bonds at issuance. This approach is similar to the one employed by Sengupta et al. (2017) for syndicated loans, which we extend by also applying it to corporate bonds issued by oil firms. The second approach utilises not only the credit spreads at issuance but the fact that bonds are traded on the secondary market and includes this information on a firm's cost of debt after the issuance of the bond. One advantage is that it possible to observe costs of debt during phases when bond issuing and loan supply might be curtailed. This presents an opportunity to estimate and analyse a possible selection effect. This selection effect might be especially pronounced during phases, when credit markets dry up. Thus, this approach offers the opportunity to assess the extent of this selection effect.




(a) Leverage

(b) Profitability

(c) Total Assets

Figure 7: Boxplots of main firm characteristics

#### 4.2.1. Distributed Lag Model for Credit Spreads at Issuance

This first approach could be described as a variant of a distributed lag model, since only the lagged financial variables are included in the model. This is based on the assumption that credit spreads are based on the latest available financial statement of a firm. The initial approach on the loan and bond level is implemented estimating the following model of credit spreads:

$$\begin{aligned} Y_{i,j,t} = {} & \beta_0 + \beta_1 DEBT_{i,j,t} + \beta_2 FIRM_{i,t-1} + \beta_3 OIL_t + \beta_4 MACRO_t \\ & + \beta_5 D2008_t + \beta_6 D2014_t + \nu_t + \varepsilon_{i,j,t}, \end{aligned} \tag{1}$$

where $Y_{i,j,t}$ is the (average quarterly) credit spread of a syndicated loan received or a bond issued $j$ by firm $i$ at time $t$, $DEBT_{i,j,t}$ is a vector containing loan/bond characteristics, $FIRM_{i,t-1}$ contains firm characteristics at time $t-1$, $OIL_t$ is a vector with oil price, its volatility and the volume of crude oil exports, $MACRO_t$ includes control variables on the macroeconomic situation, $D2008_t$ and $D2014_t$ are dummy variables for both oil price shocks, $\nu_t$ are year fixed effects, and $\varepsilon_{i,j,t}$ is the error term. Following the literature (Chava 2014; Chen et al. 2007;

Table 2: Differences in the Sample

| | Full Sample (mean) (1) | Pre 2008 (mean) (2) | Post 2008 Pre 2014 (mean) (3) | Post 2014 (mean) (4) | Difference (3) - (2) | Difference (4) - (3) |
|---|---|---|---|---|---|---|
| Total Assets | 6747.0027 | 6244.3662 | 7654.3718 | 10326.3615 | 1410.0056*** | 2671.9896*** |
| Total Debt | 1706.4348 | 1674.4294 | 1764.2116 | 2991.0219 | 89.7822* | 1226.8103*** |
| Leverage | 0.8581 | 0.8592 | 0.8561 | 1.4916 | –0.0032 | 0.6355 |
| Profitability | –0.2510 | –0.2030 | –0.3369 | –0.5025 | –0.1340 | –0.1655 |
| Loan Credit Spread | 184.8016 | 163.4243 | 230.8245 | 204.2390 | 67.4002*** | –26.5854*** |
| Loan Amount | 822.5305 | 799.7916 | 871.8274 | 1254.4782 | 72.0359 | 382.6508*** |
| Loan Maturity | 43.2096 | 41.9909 | 45.8515 | 48.6087 | 3.8606*** | 2.7572*** |
| Bond Credit Spread | 2.6297 | 2.3878 | 2.9411 | 3.3838 | 0.5533*** | 0.4427*** |
| Bond Amount | 772.9341 | 753.8738 | 796.2685 | 1071.5869 | 42.4128 | 275.3004*** |
| Bond Maturity | 191.1667 | 203.5953 | 175.1264 | 154.0742 | –28.4688*** | –21.0522*** |

*Note:* *p<0.1; **p<0.05; ***p<0.01

Ghouma et al. 2018; Sengupta et al. 2017; Valta 2012), we use information of the respective previous quarter for all variables other than the characteristics of the respective loan or bond itself. This is the most recent information that is available to the bank or the capital markets to evaluate the firm trying to raise debt and the relevant market conditions. Additionally, this has the advantage that the likely endogeneity between the debt and the financial variables is averted by design. As mentioned above, we can distinguish between the firms' positions in the value chain of the oil industry. To investigate whether there are differences in the effects of oil prices and other controls on the cost of debt along the value chain, we estimate all models for the whole sample as well as separately for firms in the three sub-sectors upstream & support services, midstream, and downstream.

#### 4.2.2. Panel Data Methods for Continuous Credit Spreads

With the TRACE data it is possible to continuously asses these cost of debt based on the development of bond prices and yields on the secondary market. This approach utilises the panel structure of the data by aggregating the high-frequency bond pricing data of TRACE and combining these data together with the quarterly financial data from the Capital IQ database. This results in a larger sample compared to the first approach, since we do not only observe the credit spread at the time of issuance, but also whenever there are trades reported in the TRACE database.

The main advantage of the panel data approach is the possibility to assess the different impacts that oil prices have on a firm depending on its position in the oil industry's supply chain. In particular, we can estimate the model jointly for all oil firms and then test for differences across industry classifications compared to the first approach, where we estimate the models separately for the three sub-sectors. Since the industry classification of the individual firm is time-invariant, we cannot use the classical Fixed Effects (FE) estimation to directly estimate the impact of the position in the oil industry's supply chain. To circumvent this methodological limitation, we utilise an approach, which extends the basic Random Effects (RE) estimation method. Initially proposed by Mundlak (1978) and further developed by Bell and Jones (2015), this within-between approach has the advantage that it allows to decompose the combined effect in the random effect models into between- and within-firm effects. Thus, it is possible to obtain separate estimates for the effect of an explanatory variable on the dependent variable between firms (between-firm estimator) and the effect within a particular higher-level group (within-firm estimator). The model can be expressed in its most general form as:

$$Y_{i,t} = \beta_{0,i} + \beta_1(X_{i,t} - \bar{X}_i) + \beta_2\bar{X}_i + \gamma Z_i + u_{0,i} + \varepsilon_{i,t}, \tag{2}$$

where $Y_{i,t}$ is the dependent variable, $X_{i,t}$ are time variant explanatory variables, and $Z_i$ are time-invariant variables. The interpretation of $\beta_1$ is the same as in the FE model, because it measures the effects of within-firm deviations of $X$ on the within-firm deviations of $Y$. The $\beta_2$ is then indicating how the impact varies with cross-sectional variation in the dependent variable, i.e. across industry classification in our model.[9].

We estimate the following within-between effects model for the determinants of the average quarterly credit spread of a firm:

$$\begin{aligned} Y_{i,t} = \beta_{0,i} &+ \beta_1 DEBT_{i,t} + \beta_2 FIRM_{i,t-1} + \beta_3 OIL_t + \beta_4 MACRO_t \\ &+ \beta_5 D2008_t + \beta_6 D2014_t + \beta_7 INTER_{i,t} + \gamma Z_i + u_{0i} + \varepsilon_{i,t}, \end{aligned} \tag{3}$$

where $Y_{i,t}$ is the (quarterly volume-weighted average) credit spread of the outstanding bonds by firm $i$ at time $t$, $DEBT_{i,t}$ is a vector containing bond characteristics, $FIRM_{i,t-1}$ contains firm characteristics in the previous quarter, $OIL_t$ is a vector with oil price and export information, $MACRO_t$ includes control variables on the macroeconomic situation, $D2008_t$ and $D2014_t$ are dummy variables for both oil price shocks, $INTER_{i,t}$ is an interaction terms between the oil price development and the industry classification of a firm, $u_{0i}$ are random errors of the model predicting $\beta_{0,i}$, and $\varepsilon_{i,t}$ is the error term. Since it is not appropriate to include interaction terms into the model, we are following Giesselmann and Schmidt-Catran (2018) and use their "double-demeaned" estimator for the coefficient of the interaction terms. The estimation of the within-between model is carried out using R Core Team (2018) and the *panelr* package by Long (2019). The results of both empirical approaches are discussed in the following subsection.

## 4.3. Estimation Results

### 4.3.1. Determinants of Credit Spreads at Issuance

In this subsection, we present the estimation results of the model equation (1). Results for syndicated loans are displayed in Table 3, while Table 4 summarises the regression results of the estimation on the individual bond level. In both tables, the respective first columns show

[9] A more detailed description of this model and how to arrive at this notation starting from the general RE model can be found in Bell and Jones (2015)

the results for all firms, while the subsequent columns display the results for each of the firms' sub-samples depending on the position in the oil industry's supply value chain. The dependent variables in both the bond and the loan estimations are the natural logs of the respective credit spreads in basis points.

Overall, we find the expected effects of the various firm characteristics on the credit spreads of loans and bonds at issuance. For the full sample, we find that a positive and significant effect of *Leverage* on the credit spread of loans. This suggests that a higher firm indebtedness increases a firm's cost of debt. As such, an increase in the leverage ratio by 0.1 increases the credit spread by on average 7 percent. When comparing the effects across sub-sectors, the effect is the strongest for firms in the midstream segment with around 14 percent and the weakest and only marginally significant for downstream firms. The effect of indebtedness is quantitatively similar for bonds, when considering the full sample. We find that the credit spreads of bonds at issuance increase by 7.6 percent if leverage increases by 0.1. As for syndicated loans, the effect of leverage is the strongest for midstream firms. This finding is in line with the theory, as higher firm indebtedness increases the risk of default, which translates into a higher risk premium for further debt charged by banks and the capital market.

For both loans and bonds, the impact of firm profitability is somewhat ambiguous. In the case of loans, the coefficient of *Profitability* is negative, but not significant for the full sample. This is caused by the fact that the effect of the profitability differs across the sub-samples and thus the negative impact for the midstream and upstream & support services categories is cancelled out by the strong positive impact for the downstream firms. According to this, it would be especially beneficial for midstream firms to increase their profitability as a 0.1 increase would decrease the bond credit spread by 32 percent. In the bond credit spread model, the effect of *Profitability* is weakly significant for the full sample. Similar to loans, the effect of firm profitability is strongest for midstream firms.

In all models we find highly significant negative effects of firm size, measured by the logarithm of *Total Assets*, on credit spreads at issuance. This means that larger firms on average face lower costs of debt. This effect remains robust across the whole supply chain, while it is quantitatively larger in the case of bonds. For the full sample, a 1 percent increase in total assets leads to a decrease of the bond credit spread by almost 0.3 percent, whereas the loan credit spread increases by only 0.17 percent on average. These results are not surprising, since more assets also translate into more potential collateral which lowers the risk of a complete financial loss for the banks.

In addition to the firm-level characteristics, we included bond/loan-level controls in the models. With respect to the volume of debt raised, we find that the loan amount has a negative effect on the credit spread for the full sample and the sub-samples with the exception of midstream firms. For bonds, however, the results indicate the opposite: the higher the amount raised by issuing a bond, the higher is the bond credit spread. An explanation for these different results might be two opposite effects. On the one hand, a higher volume of debt means higher potential losses for the lender in the case of default, which should translate into higher costs. On the

Table 3: Determinants of the loan credit spread at issuance.

| | *Dependent variable:* log(Loan Credit Spread)$_t$ | | | |
|---|---|---|---|---|
| | Full Sample | Upstream & Support Services | Midstream | Downstream |
| Leverage$_{t-1}$ | 0.7048$^{***}$ | 0.6681$^{***}$ | 1.3887$^{***}$ | 0.4138$^{**}$ |
| | (0.0673) | (0.0699) | (0.1588) | (0.2029) |
| Profitability$_{t-1}$ | −0.0431 | −0.2493 | −3.3084$^{***}$ | 1.4998$^{*}$ |
| | (0.2175) | (0.2082) | (0.7167) | (0.8424) |
| log(Total Assets)$_{t-1}$ | −0.1741$^{***}$ | −0.1653$^{***}$ | −0.1033$^{***}$ | −0.0940$^{***}$ |
| | (0.0087) | (0.0118) | (0.0174) | (0.0206) |
| log(Loan Amount)$_t$ | −0.0383$^{***}$ | −0.0921$^{***}$ | 0.0392$^{*}$ | −0.1910$^{***}$ |
| | (0.0123) | (0.0156) | (0.0206) | (0.0298) |
| Maturity$_t$ | 0.0028$^{***}$ | 0.0021$^{***}$ | 0.0017$^{*}$ | 0.0051$^{***}$ |
| | (0.0006) | (0.0008) | (0.0010) | (0.0011) |
| Credit Spread$_t$ | 0.2658$^{***}$ | 0.1889$^{***}$ | 0.2914$^{***}$ | 0.3083$^{*}$ |
| | (0.0508) | (0.0569) | (0.0881) | (0.1638) |
| Term Spread$_t$ | 0.1381$^{***}$ | 0.0766$^{***}$ | 0.2078$^{***}$ | 0.1809$^{***}$ |
| | (0.0106) | (0.0128) | (0.0170) | (0.0293) |
| Oil volatility$_t$ | −0.0166$^{**}$ | −0.0181$^{**}$ | −0.0169 | −0.0133 |
| | (0.0072) | (0.0082) | (0.0125) | (0.0204) |
| log(Oil Price)$_t$ | 0.1224$^{***}$ | 0.0825 | 0.1719$^{**}$ | 0.1140 |
| | (0.0468) | (0.0577) | (0.0742) | (0.1288) |
| log(Oil Exports)$_t$ | 0.0146 | −0.0099 | 0.0693$^{*}$ | −0.0244 |
| | (0.0264) | (0.0325) | (0.0417) | (0.0743) |
| D2008 | −0.1134 | 0.1267 | −0.2628 | −0.2358 |
| | (0.1395) | (0.1624) | (0.2382) | (0.4118) |
| D2014 | −0.0048 | 0.1336 | 0.0635 | −0.2112 |
| | (0.0939) | (0.1174) | (0.1503) | (0.2436) |
| Constant | −89.9041$^{***}$ | −123.2027$^{***}$ | −29.7483 | −87.8063$^{*}$ |
| | (16.7929) | (20.7913) | (26.2921) | (47.5227) |
| Year fixed effects | Yes | Yes | Yes | Yes |
| Observations | 3047 | 1522 | 1171 | 354 |
| $R^2$ | 0.3568 | 0.4457 | 0.2886 | 0.4477 |
| Adjusted $R^2$ | 0.3540 | 0.4409 | 0.2806 | 0.4266 |
| F-Statistic | 129.3966$^{***}$ | 93.2626$^{***}$ | 36.0977$^{***}$ | 21.2009$^{***}$ |
| | (df = 13; 3033) | (df = 13; 1508) | (df = 13; 1157) | (df = 13; 340) |

*Note:* $^{*}$p<0.1; $^{**}$p<0.05; $^{***}$p<0.01
Standard errors in parentheses.

Table 4: Determinants of the bond credit spread at issuance.

| | *Dependent variable:* log(Bond Credit Spread)$_t$ | | | |
|---|---|---|---|---|
| | Full Sample | Upstream & Support Services | Midstream | Downstream |
| Leverage$_{t-1}$ | 0.7675*** | 0.8615*** | 1.5159*** | 0.9762*** |
| | (0.1076) | (0.1361) | (0.2004) | (0.3427) |
| Profitability$_{t-1}$ | −0.5509 | −0.3099 | −4.2920*** | 1.1002 |
| | (0.4238) | (0.4529) | (1.3431) | (1.6793) |
| log(Total Assets)$_{t-1}$ | −0.2936*** | −0.2851*** | −0.1424*** | −0.3156*** |
| | (0.0118) | (0.0213) | (0.0187) | (0.0388) |
| log(Bond Amount)$_t$ | 0.2231*** | 0.1428*** | 0.1418*** | 0.2636*** |
| | (0.0186) | (0.0382) | (0.0212) | (0.0889) |
| Maturity$_t$ | −0.0002 | 0.0000 | 0.0003** | 0.0007** |
| | (0.0001) | (0.0003) | (0.0002) | (0.0003) |
| Credit Spread$_t$ | 0.5526*** | 0.5252*** | 0.5970*** | 0.5211*** |
| | (0.0478) | (0.0723) | (0.0610) | (0.1404) |
| Term Spread$_t$ | −0.0001 | −0.0225 | −0.0006 | 0.1204** |
| | (0.0171) | (0.0260) | (0.0210) | (0.0573) |
| Oil volatility$_t$ | 0.0079 | 0.0028 | 0.0238** | −0.0024 |
| | (0.0084) | (0.0123) | (0.0108) | (0.0287) |
| log(Oil Price)$_t$ | 0.1086* | 0.0477 | 0.1682** | 0.0176 |
| | (0.0567) | (0.0916) | (0.0712) | (0.1603) |
| log(Oil Exports)$_t$ | 0.0130 | −0.0545 | 0.1162** | −0.1297 |
| | (0.0426) | (0.0680) | (0.0507) | (0.1444) |
| D2008 | −0.1143 | −0.0597 | −0.1843 | −0.6691 |
| | (0.1644) | (0.3451) | (0.1789) | (0.6667) |
| D2014 | −0.1146 | −0.0093 | −0.2623* | −0.2205 |
| | (0.1119) | (0.1734) | (0.1459) | (0.3120) |
| Constant | −47.4611 | −133.8198*** | 76.2765** | −182.7495* |
| | (27.5084) | (43.2281) | (32.9483) | (94.1098) |
| Year fixed effects | Yes | Yes | Yes | Yes |
| Observations | 1511 | 558 | 746 | 207 |
| $R^2$ | 0.4675 | 0.4975 | 0.3906 | 0.5535 |
| Adjusted $R^2$ | 0.4629 | 0.4855 | 0.3798 | 0.5234 |
| F-Statistic | 101.0944*** | 41.4347*** | 36.0937*** | 18.4044*** |
| | (df = 13; 1497) | (df = 13; 544) | (df = 13; 732) | (df = 13; 193) |

*Note:* *p<0.1; **p<0.05; ***p<0.01
Standard errors in parentheses.

other hand, higher volumes of debt are likely to be raised by on average bigger firms, which are likely to be more creditworthy and receive better credit conditions. Thus, the former might be stronger in the case of corporate bonds, while the latter might be more important in the case of syndicated loans.

A higher maturity is expected to lead to higher costs of debt, as it increases the debt exposure of the lender. We find this relationship for syndicated loans: an increase of the loan maturity by one month results in an increase of the loan credit spread by 0.25 percent. When considering the sub-sectors, we find the effect for the upstream & support services and downstream sectors, while it is positive but not significant for midstream firms. In the case of bonds, there is some evidence for a positive effect of maturity on credit spreads for downstream and midstream firms, while the coefficient of *Maturity* is not significant for the full sample.

The variables measuring the overall risk environment of the macroeconomy partly affect the credit spread of newly issued loans. The TED spread, a measure indicating the risk in the banking sector, does not have any effect. However, the term spread seems to raise the spread that firms have to pay for loans. In the corporate bond models, the overall risk environment in the bond market is measured by the *Credit Spread*, which is positively correlated with the bond credit spread the firms in the oil industry are facing. This means that during periods, when the AAA-BAA credit spread widens for the overall economy, it also increases for firms along the oil industry's supply chain.

We now turn to the variables of interest in the models, i.e. various measures of oil prices. The results for the oil price itself differ between the bond and the loan models. In the case of loans, the coefficients of *log(Oil Price)* are negative, but not statistically significant across all specifications. However, the oil price has a positive and significant effect on the credit spreads of corporate bonds issued by oil firms. In fact, we find that a 1 percent increase in the WTI spot price yields a 0.15 percent increase of the bond credit spread. When looking at the different sub-sectors, the effect is only significant for midstream firms. An explanation might be that crude is likely to be an input for midstream firms, such that higher prices mean higher costs for these firms. For loans we find some evidence that the possibility of exporting to other markets, which was banned until December 2015, seems to decrease the perceived riskiness of oil firms and thus their costs of debt. However, there is no such effect for credit spreads of bonds at issuance.

Results regarding the volatility of oil prices, however, are more robust across industry classifications and sources of debt. We find that the credit spreads of loans and bonds at issuance increase with the volatility of the oil price when considering the full samples of firms. When looking at the sub-sectors, the impact of price volatility on credit spreads remains robust. The only exceptions are midstream firms in the case of loans and downstream firms in the case of bonds. Overall, these results indicate that a higher price volatility translates into higher uncertainty about oil prices and thus the economic environment of oil firms. Thus, it is not surprising that banks and capital markets charge higher prices for debt in such periods.

Finally, we do not find convincing evidence for any impact of the oil price shocks in 2008 and 2014. Both shocks are not significant in the models for syndicated loans. For corporate bonds, there is a positive effect of the oil price shock in 2008 on the credit spread. However, this effect seems to be driven by the midstream firms. We do not find any effect on the credit spread for the two other categories however. A reason for this result might be the fact these shocks are captured by other variables in the model. As we explicitly included the oil price and its volatility, it is likely that these variables capture the adverse oil price shocks. Furthermore, the additional effect of the financial crisis parallel to the 2008 oil price shock might be captured by the variables on the macroeconomic conditions. Another possible explanation for the non-significance in the case of loans might be the decrease in issued loans during the financial crisis parallel to the 2008 oil price shock. Thus, there might be a selection effect, i.e. only the few least risky firms received loans, while others were not able to raise debt. Thus, the crisis might have led to a more extent to credit rationing instead of higher spreads.

#### 4.3.2. Determinants of Credit Spreads on the Secondary Market

This subsection discusses the results of the second empirical approach using data on bond credit spread on the secondary market. The results of the within-between effects estimation of the model in equation (3) is presented in Table 5. The results by and large confirm the previous results and provide additional information on the determinants of the cost of debt.

The within-effects can be interpreted as the coefficients of a FE estimation and thus capture the impact of the within-firm variation on the variation of the bond credit spread of the firm. With respect to the firm-specific financial variables, we find that firms with a higher debt-to-asset ratio face higher costs of debt and larger firms, measured by their total assets, have to pay a lower credit spread. The controls for the macroeconomic environment have the expected directions. Both the term and the credit spread have a positive impact on the calculated bond credit spread. Thus, an overall increase of the perceived risk in the macroeconomy also increases the risk premium the market implicitly assigns to firms in the oil industry.

The impact of the oil price and its volatility is also statistically significant. A higher oil price leads to lower financing costs for the firms, whereas rising uncertainty, i.e. higher volatility, increases the cost of debt. The additional possibility that oil firms could diversify into foreign markets, by exporting a certain amount of their production, decreases their financing costs. The two oil price shocks in the fourth quarter of 2008 and 2014, respectively, are decreasing the bond credit spread on the firm-level.

The between-effects need to be interpreted against firms in the midstream sub-sector, which is the reference group. There exist significant differences between the costs of debt for firms in the different industry categories. The results of this estimation show that downstream and upstream & support services do on average face significantly higher financing costs compared to midstream firms by 41 or 46 percent respectively. This meets our expectations given the boxplots in Figure 6 (c). Moreover, the interaction effects between the industry classifications and the oil price show that the impact of the oil price on the financing costs varies with the industry classification of the firms. The results indicate that the negative impact of an oil price increase on the bond credit spread is much stronger for upstream & support services compared to midstream firms. In contrast, downstream firms are relatively weaker affected. An explanation might be the theoretical consideration that oil prices do have a more direct impact on the firms higher up the supply chain.

Table 5: Within-Between effects estimation of the determinants of the bond credit spread on the secondary market for the full sample.

| | *Dependent variable:* | | | | |
|---|---|---|---|---|---|
| | log(Bond Credit Spread)$_t$ | | | | |
| | Est. | Std. Error | t-val. | d.f. | p-Value |
| Within-Effects | | | | | |
| Leverage$_{t-1}$ | 1.16 | 0.06 | 20.66 | 6527 | 0.00 |
| Profitability$_{t-1}$ | −0.80 | 0.11 | −7.11 | 6379 | 0.00 |
| log(Total Assets)$_{t-1}$ | −0.03 | 0.02 | −1.30 | 6410 | 0.19 |
| Avg. Months-to-Maturity | 0.00 | 0.00 | −1.53 | 6358 | 0.13 |
| Credit Spread | 0.52 | 0.02 | 21.64 | 6320 | 0.00 |
| Term Spread | −0.17 | 0.01 | −22.36 | 6343 | 0.00 |
| Oil Volatility | 0.03 | 0.00 | 7.75 | 6314 | 0.00 |
| log(Oil Price) | −0.09 | 0.05 | −1.69 | 6442 | 0.09 |
| log(Oil Exports) | 0.04 | 0.02 | 1.94 | 6321 | 0.05 |
| D2008 | −0.23 | 0.07 | −3.19 | 6301 | 0.00 |
| D2014 | −0.13 | 0.06 | −2.26 | 6310 | 0.02 |
| Between-Effects | | | | | |
| (Intercept) | 9.13 | 4.12 | 2.21 | 244 | 0.03 |
| Leverage$_{t-1}$ | 1.50 | 0.22 | 6.71 | 209 | 0.00 |
| Profitability$_{t-1}$ | −2.72 | 0.79 | −3.45 | 224 | 0.00 |
| log(Total Assets)$_{t-1}$ | −0.13 | 0.03 | −4.72 | 196 | 0.00 |
| Avg. Months-to-Maturity | 0.00 | 0.00 | −6.00 | 196 | 0.00 |
| Credit Spread | −3.03 | 2.02 | −1.50 | 220 | 0.13 |
| Term Spread | 0.63 | 0.42 | 1.49 | 205 | 0.14 |
| Oil Volatility | 0.13 | 0.46 | 0.27 | 244 | 0.79 |
| log(Oil Price) | −0.15 | 1.18 | −0.13 | 245 | 0.90 |
| log(Oil Exports) | −0.07 | 0.16 | −0.44 | 216 | 0.66 |
| D2008 | 18.98 | 12.88 | 1.47 | 224 | 0.14 |
| D2014 | −3.17 | 8.07 | −0.39 | 222 | 0.69 |
| Upstream & Support Services | 0.42 | 0.08 | 5.28 | 192 | 0.00 |
| Downstream | 0.42 | 0.11 | 3.92 | 190 | 0.00 |
| Time Fixed Effects | 0.00 | 0.00 | −1.04 | 6383 | 0.30 |
| Cross-Level Interactions | | | | | |
| log(Oil Price)*Upstream & Support Services | −0.38 | 0.06 | −6.86 | 6476 | 0.00 |
| log(Oil Price)*Downstream | −0.12 | 0.06 | −2.06 | 6446 | 0.04 |
| Random Effects | | | | | |
| Group | Parameter | Std. Dev. | | | |
| Firm ID | (Intercept) | 0.46 | | | |
| Residual | | 0.54 | | | |

p-values calculated using Satterthwaite d.f.

Table 6: Information on the estimated within-between model.

| Model Info & Fit: | | | |
|---|---|---|---|
| Firms | 232 | Quarters | 11-70 |
| Type | Linear mixed effects | Specification | within-between |
| AIC | 11423.73 | BIC | 11627.4 |
| Pseudo-$R^2$ (fixed effects) | 0.51 | Pseudo-$R^2$ (total) | 0.71 |
| Entity ICC | 0.41 | | |

# 5. Conclusion

This paper provides an empirical analysis of the effect of oil prices on the costs of debt in the oil industry. For this analysis, we combine data on syndicated loans and bonds issued with firm-level financial data of firms in the US oil industry. In the case of bonds, we further use data in bond trades on the secondary market. Hence, we capture both the banking sector and the bond market that are both frequently used by oil firms in the US to raise debt. This allows us to compare whether banks and the capital market evaluate the effect of oil prices on the creditworthiness of oil firms differently. Furthermore, we can explicitly check whether certain effects might be driven by specifics of the banking sector or debt markets. Controlling for other factors, as the macroeconomic conditions of the economy, we find that oil prices, in particular their volatility, significantly affect the cost of debt of US oil firms.

For the firm characteristics, we confirm findings of previous studies on various sectors. We find that larger firms have lower costs of debt, while the premium a firm has to pay on its debt increases with its indebtedness. As one would expect, we find that the credit spread for loans increases with the maturity. Furthermore, we find that the cost of debt in the oil industry increases with the perceived credit risk in the general economy.

With respect to oil prices, we find that, even after controlling for loan/bond and firm characteristics, oil prices have an effect on a firm's cost of debt. The within-between effects estimation further reveals that the effect of the oil prices differs across sub-sectors. It is particularly strong for upstream & support services firms. Our results further indicate that particularly the volatility of oil prices seems to be important for the costs of debt. We find that higher oil price volatility seems to lead to higher uncertainty about the market environment of oil firms. As a consequence, banks and capital markets charge higher prices for debt.

Our results on the impact of the oil price shocks in 2008 and 2014 is rather ambiguous. A reason for this might be these shocks are captured by other variables, as the oil price and its volatility, which we explicitly model. Furthermore, the additional effect of the financial crisis parallel to the 2008 oil price shock might be captured by the variables on the macroeconomic conditions. Another possible explanation for the non-significance in the case of loans might be the decrease in issued loans during the financial crisis parallel to the 2008 oil price shock. Thus, there might be a selection effect, i.e. only the few least risky firms received loans, while others were not able to raise debt. Thus, the crisis might have led more to credit rationing instead of higher spreads.

Thus, even after controlling for loan/bond and firm characteristics, oil prices, in particular oil price volatility, have effects on a firm's cost of debt. In addition to directly affecting oil firms' sales revenues, decreasing oil prices increase the price a firm has to pay to raise new debt. Both banks and the bond market seem to consider falling oil prices as well as higher price volatility risks that increase the probability of default and thus reduces the creditworthiness of oil firms. Consequently, banks and the capital market demand higher credit spreads. These results emphasise the link between commodity prices and the costs of debt producers of these commodities are facing. Potential areas for further research could be other industries. In this paper, we only consider the oil industry and it thus might be fruitful to examine, e.g. the coal or natural gas sectors. Furthermore, our results indicate that the effects differ between bank loans and corporate bonds. Thus, a deeper investigation of developments in the banking sector, its regulation and changes of the market for corporate bonds might provide explanations for the differences we find.

## References


Azar, A. (2017): 'Reserve Base Lending And The Outlook For Shale Oil And Gas Finance'. In: *Columbia Center on Global Energy Policy*.

Badoer, D. C. and C. M. James (2016): 'The Determinants of Long-Term Corporate Debt Issuances'. In: *The Journal of Finance* 71 (1), pp. 457–492. DOI: 10.1111/jofi.12264.

Baffes, J., M. A. Kose, F. Ohnsorge and M. Stocker (2015): 'The Great Plunge in Oil Prices: Causes, Consequences, and Policy Responses'. In: *SSRN Electronic Journal*. DOI: 10.2139/ssrn.2624398.

Baumeister, C. and L. Kilian (2016): 'Understanding the Decline in the Price of Oil since June 2014'. In: *Journal of the Association of Environmental and Resource Economists* 3 (1), pp. 131–158. DOI: 10.1086/684160.

Becker, B. and V. Ivashina (2014): 'Cyclicality of credit supply: Firm level evidence'. In: *Journal of Monetary Economics* 62, pp. 76–93. DOI: 10.1016/j.jmoneco.2013.10.002.

Bell, A. and K. Jones (2015): 'Explaining Fixed Effects: Random Effects Modeling of Time-Series Cross-Sectional and Panel Data'. In: *Political Science Research and Methods* 3 (1), pp. 133–153. DOI: 10.1017/psrm.2014.7.

Benmelech, E. and N. K. Bergman (2011): 'Bankruptcy and the Collateral Channel'. In: *The Journal of Finance* 66 (2), pp. 337–378. DOI: 10.1111/j.1540-6261.2010.01636.x.

Benmelech, E., M. J. Garmaise and T. J. Moskowitz (2005): 'Do Liquidation Values Affect Financial Contracts? Evidence from Commercial Loan Contracts and Zoning Regulation'. In: *The Quarterly Journal of Economics* 120 (3), pp. 1121–1154. DOI: 10.1093/qje/120.3.1121.

Berg, T., A. Saunders and S. Steffen (2016): 'The Total Cost of Corporate Borrowing in the Loan Market: Don't Ignore the Fees'. In: *The Journal of Finance* 71 (3), pp. 1357–1392. DOI: 10.1111/jofi.12281.

Bessembinder, H., K. M. Kahle, W. F. Maxwell and D. Xu (2008): 'Measuring abnormal bond performance'. In: *The Review of Financial Studies* 22 (10), pp. 4219–4258.

Board of Governors of the Federal Reserve System (2019): *Treasury Constant Maturity Interest Rates*. Ed. by FRED.

Bolton, P. and D. S. Scharfstein (1996): 'Optimal Debt Structure and the Number of Creditors'. In: *Journal of Political Economy* 104 (1), pp. 1–25. DOI: 10.1086/262015.

Boneva, L. and O. Linton (2017): 'A discrete-choice model for large heterogeneous panels with interactive fixed effects with an application to the determinants of corporate bond issuance'. In: *Journal of Applied Econometrics* 32 (7), pp. 1226–1243. DOI: 10.1002/jae.2568.

Brownlees, C. T. and G. M. Gallo (2006): 'Financial econometric analysis at ultra-high frequency: Data handling concerns'. In: *Computational Statistics & Data Analysis* 51 (4), pp. 2232–2245. DOI: 10.1016/j.csda.2006.09.030.

Cakir Melek, N. (2015): 'What could lower prices mean for US oil production?' In: *Economic Review - Federal Reserve Bank of Kansas City*.

Campbell, J. Y. and G. B. Taksler (2003): ‘Equity Volatility and Corporate Bond Yields’. In: *The Journal of Finance* 58 (6), pp. 2321–2350. DOI: 10.1046/j.1540-6261.2003.00607.x.

Chava, S. and M. R. Roberts (2008): ‘How Does Financing Impact Investment? The Role of Debt Covenants’. In: *The Journal of Finance* 63 (5), pp. 2085–2121. DOI: 10.1111/j.1540-6261.2008.01391.x.

Chava, S. (2014): ‘Environmental Externalities and Cost of Capital’. In: *Management Science* 60 (9), pp. 2223–2247. DOI: 10.1287/mnsc.2013.1863.

Chava, S., D. Livdan and A. Purnanandam (2009): ‘Do Shareholder Rights Affect the Cost of Bank Loans?’ In: *Review of Financial Studies* 22 (8), pp. 2973–3004. DOI: 10.1093/rfs/hhn111.

Chen, L., D. A. Lesmond and J. Wei (2007): ‘Corporate Yield Spreads and Bond Liquidity’. In: *The Journal of Finance* 62 (1), pp. 119–149. DOI: 10.1111/j.1540-6261.2007.01203.x.

Dennis, S., D. Nandy and L. G. Sharpe (2000): ‘The Determinants of Contract Terms in Bank Revolving Credit Agreements’. In: *Journal of Financial and Quantitative Analysis* 35 (01), pp. 87–110.

Dick-Nielsen, J. (2009): ‘Liquidity Biases in TRACE’. In: *The Journal of Fixed Income* 19, pp. 43–55. DOI: 10.3905/jfi.2009.19.2.043.

– (2014): ‘How to Clean Enhanced TRACE Data’. In: *SSRN*. DOI: 10.2139/ssrn.2337908.

Dick-Nielsen, J., P. Feldhütter and D. Lando (2012): ‘Corporate bond liquidity before and after the onset of the subprime crisis’. In: *Journal of Financial Economics* 103 (3), pp. 471–492. DOI: 10.1016/j.jfineco.2011.10.009.

Domanski, D., J. Kearns, M. J. Lombardi and H. S. Shin (2015): ‘Oil and debt’. In: *BIS Quarterly Review*.

EIA (2017): *Monthly Energy Review – July 2017*. Tech. rep. https://www.eia.gov/totalenergy/data/monthly/archive/00351707.pdf. U.S. Energy Information Administration.

Farrant, K., M. Inkinen, M. Rutkowska and K. Theodoridis (2013): ‘What can company data tell us about financing and investment decisions?’ In: *Bank of England Quarterly Bulletin* 53 (4), pp. 361–370.

Faulkender, M. and M. A. Petersen (2006): ‘Does the Source of Capital Affect Capital Structure?’ In: *Review of Financial Studies* 19 (1), pp. 45–79. DOI: 10.1093/rfs/hhj003.

Federal Reserve Bank of St. Louis (2018a): *TED Spread [TEDRATE] retrieved from FRED, Federal Reserve Bank of St. Louis;* https://fred.stlouisfed.org/series/TEDRATE.

– (2018b): *Term Spread – 10-Year Treasury Constant Maturity Minus 3-Month Treasury Constant Maturity*. https://fred.stlouisfed.org/series/T10Y3M.

Ghouma, H., H. Ben-Nasr and R. Yan (2018): ‘Corporate governance and cost of debt financing: Empirical evidence from Canada’. In: *The Quarterly Review of Economics and Finance* 67, pp. 138–148. DOI: 10.1016/j.qref.2017.06.004.

Giesselmann, M. and A. Schmidt-Catran (2018): *Interactions in Fixed Effects Regression Models*. Discussion Papers of DIW Berlin 1748. DIW Berlin, German Institute for Economic Research.

Gilchrist, S., J. W. Sim and E. Zakrajšek (2014): *Uncertainty, Financial Frictions, and Investment Dynamics*. Working Paper 20038. National Bureau of Economic Research. DOI: 10.3386/w20038.

Gilchrist, S. and E. Zakrajšek (2012): 'Credit Spreads and Business Cycle Fluctuations'. In: *American Economic Review* 102 (4), pp. 1692–1720. DOI: 10.1257/aer.102.4.1692.

Goss, A. and G. S. Roberts (2011): 'The impact of corporate social responsibility on the cost of bank loans'. In: *Journal of Banking & Finance* 35 (7), pp. 1794–1810. DOI: 10.1016/j.jbankfin.2010.12.002.

Graham, J. R., M. T. Leary and M. R. Roberts (2014): *How does government borrowing affect corporate financing and investment?* Working Paper 20581. National Bureau of Economic Research. DOI: 10.3386/w20581.

Greenwood, R., S. Hanson and J. C. Stein (2010): 'A Gap-Filling Theory of Corporate Debt Maturity Choice'. In: *The Journal of Finance* 65 (3), pp. 993–1028. DOI: 10.1111/j.1540-6261.2010.01559.x.

Greenwood, R. and D. Vayanos (2014): 'Bond Supply and Excess Bond Returns'. In: *Review of Financial Studies* 27 (3), pp. 663–713. DOI: 10.1093/rfs/hht133.

Hamilton, J. D. (2009): *Causes and Consequences of the Oil Shock of 2007-08*. Working Paper 15002. National Bureau of Economic Research. DOI: 10.3386/w15002.

Hart, O. and J. Moore (1994): 'A Theory of Debt Based on the Inalienability of Human Capital'. In: *The Quarterly Journal of Economics* 109 (4), pp. 841–879. DOI: 10.2307/2118350.

Ivashina, V. (2007): 'The effects of syndicate structure on loan spreads'. In: *Working Paper*.

– (2009): 'Asymmetric information effects on loan spreads'. In: *Journal of Financial Economics* 92 (2), pp. 300–319. DOI: 10.1016/j.jfineco.2008.06.003.

Ji, Q. and Y. Fan (2011): 'A dynamic hedging approach for refineries in multiproduct oil markets'. In: *Energy* 36 (2), pp. 881–887. DOI: 10.1016/j.energy.2010.12.025.

Kilian, L. (2009): 'Not All Oil Price Shocks Are Alike: Disentangling Demand and Supply Shocks in the Crude Oil Market'. In: *American Economic Review* 99 (3), pp. 1053–1069. DOI: 10.1257/aer.99.3.1053.

Kilian, L. and R. J. Vigfusson (2011): 'Are the responses of the U.S. economy asymmetric in energy price increases and decreases?' In: *Quantitative Economics* 2 (3), pp. 419–453. DOI: 10.3982/qe99.

Kim, H. and H. Kung (2016): 'The Asset Redeployability Channel: How Uncertainty Affects Corporate Investment'. In: *Review of Financial Studies* 30 (1), pp. 245–280. DOI: 10.1093/rfs/hhv076.

Kinda, T., M. Mlachila and R. Ouedraogo (2016): 'Commodity Price Shocks and Financial Sector Fragility'. In: *IMF Working Papers* 16 (12), pp. 1–48. DOI: 10.5089/9781498328722.001.

Krishnamurthy, A. and A. Vissing-Jorgensen (2011): *The Effects of Quantitative Easing on Interest Rates: Channels and Implications for Policy*. Working Paper 17555. National Bureau of Economic Research. DOI: 10.3386/w17555.

Leary, M. (2009): ‘Bank Loan Supply, Lender Choice, and Corporate Capital Structure’. In: *The Journal of Finance* 64 (3), pp. 1143–1185. DOI: 10.1111/j.1540-6261.2009.01461.x.

Lehn, K. and P. Zhu (2016): *Debt, Investment and Production in the U.S. Oil Industry: An Analysis of the 2014 Oil Price Shock*. SSRN Scholarly Paper ID 2817123. Rochester, NY: Social Science Research Network.

Lemmon, M. and M. R. Roberts (2010): ‘The Response of Corporate Financing and Investment to Changes in the Supply of Credit’. In: *The Journal of Financial and Quantitative Analysis* 45 (3), pp. 555–587.

Li, S. and N. Richie (2016): ‘Income smoothing and the cost of debt’. In: *China Journal of Accounting Research* 9 (3), pp. 175–190. DOI: 10.1016/j.cjar.2016.03.001.

Lim, J., B. A. Minton and M. S. Weisbach (2014): ‘Syndicated loan spreads and the composition of the syndicate’. In: *Journal of Financial Economics* 111 (1), pp. 45–69. DOI: 10.1016/j.jfineco.2013.08.001.

Lips, J. (2019): ‘Debt and the Oil Industry – Analysis on the Firm and Production Level’. In: *Journal of Energy Markets* 12 (3). forthcoming.

Long, J. A. (2019): *panelr: Regression Models and Utilities for Repeated Measures and Panel Data*. R package version 0.7.0.

Merton, R. C. (1974): ‘On the Pricing of Corporate Debt: The Risk Structure of Interest Rates’. In: *The Journal of Finance* 29 (2), pp. 449–470. DOI: 10.1111/j.1540-6261.1974.tb03058.x.

Moody’s (2018): *Credit Spread – Moody’s Seasoned Baa Corporate Bond Yield-Moody’s Seasoned Aaa Corporate Bond Yield*. https://fred.stlouisfed.org/graph/?g=D9J.

Mundlak, Y. (1978): ‘On the Pooling of Time Series and Cross Section Data’. In: *Econometrica* 46 (1), pp. 69–85. DOI: 10.2307/1913646.

Ortiz-Molina, H. and G. M. Phillips (2014): ‘Real Asset Illiquidity and the Cost of Capital’. In: *Journal of Financial and Quantitative Analysis* 49 (1), pp. 1–32. DOI: 10.1017/S0022109014000210.

Popov, A. and G. F. Udell (2012): ‘Cross-border banking, credit access, and the financial crisis’. In: *Journal of International Economics* 87 (1), pp. 147–161. DOI: 10.1016/j.jinteco.2012.01.008.

R Core Team (2018): *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria.

Ravazzolo, F. and P. Rothman (2013): ‘Oil and U.S. GDP: A Real-Time Out-of-Sample Examination’. In: *Journal of Money, Credit and Banking* 45 (2-3), pp. 449–463. DOI: 10.1111/jmcb.12009.

Rodziewicz, D. (2018): 'Energy Investment Variability within the Macroeconomy'. In: *Economic Review - Federal Reserve Bank of Kansas City* 103 (3), pp. 53–75. DOI: 10.18651/ER/3q18Rodziewicz.

Rossi, M. (2014): 'Realized Volatility, Liquidity, and Corporate Yield Spreads'. In: *Quarterly Journal of Finance* 04 (01), pp. 1–42. DOI: 10.1142/S2010139214500049.

Sengupta, R., B. W. Marsh and D. Rodziewicz (2017): 'Do Adverse Oil-Price Shocks Change Loan Contract Terms for Energy Firms?' In: *Economic Review - Federal Reserve Bank of Kansas City* 102 (4). DOI: 10.18651/ER/4q17senguptamarshrodziewicz.

Stiglitz, J. E. and A. Weiss (1981): 'Credit Rationing in Markets with Imperfect Information'. In: *The American Economic Review* 71 (3), pp. 393–410.

Sufi, A. (2007): 'Information Asymmetry and Financing Arrangements: Evidence from Syndicated Loans'. In: *The Journal of Finance* 62 (2), pp. 629–668. DOI: 10.1111/j.1540-6261.2007.01219.x.

– (2009): 'The Real Effects of Debt Certification: Evidence from the Introduction of Bank Loan Ratings'. In: *Review of Financial Studies* 22 (4), pp. 1659–1691. DOI: 10.1093/rfs/hhm061.

Valta, P. (2012): 'Competition and the cost of debt'. In: *Journal of Financial Economics* 105 (3), pp. 661–682. DOI: 10.1016/j.jfineco.2012.04.004.

# Appendices

## A. Classification of SIC and NAICS codes

Table 7: SIC and NAICS codes used to query firms' information from the Compustat Capital IQ database and their categorisation along the oil industries' value chain.

| SIC | NAICS | Industry Classification |
|---|---|---|
| 1311 | 211111 | Upstream |
| 1321 | 211112 | Downstream |
| 1381 | 213111 | Upstream |
| 1382 | 213112 | Support Services |
| 1382 | 541360 | Support Services |
| 1389 | 213112 | Support Services |
| 1389 | 237120 | Support Services |
| 1389 | 238910 | Support Services |
| 1623 | 237120 | Support Services |
| 1629 | 237120 | Support Services |
| 2819 | 211112 | Upstream |
| 2865 | 325110 | Downstream |
| 2869 | 325110 | Downstream |
| 2911 | 324110 | Downstream |
| 2990 | | Downstream |
| 2992 | 324191 | Downstream |
| 2999 | 324199 | Downstream |
| 3533 | 333132 | Support Services |
| 4612 | 486110 | Midstream |
| 4613 | 486910 | Midstream |
| 4619 | 486990 | Midstream |
| 4922 | 486210 | Midstream |
| 4923 | 221210 | Midstream |
| 4923 | 486210 | Midstream |
| 4924 | 221210 | Midstream |
| 4925 | 221210 | Midstream |
| 4931 | 221210 | Midstream |
| 4932 | 221210 | Midstream |
| 4939 | 221210 | Midstream |
| 5171 | 424710 | Downstream |
| 5171 | 454310 | Downstream |
| 5172 | 424720 | Downstream |
| 5900 | | Downstream |
| 5983 | 454310 | Downstream |
| 5984 | 454310 | Downstream |
| 5989 | 454310 | Downstream |
| 6792 | 523910 | Downstream |
| 6792 | 533110 | Downstream |
| 7373 | | Support Services |
| 8741 | 237120 | Support Services |

# B. Definition of variables

Based on the data from the Capital IQ database additional variables can be calculated, which are commonly used in the empirical corporate finance literature.

Table 8: Definition of variables used in the empirical analysis.

| Variable | Description |
|---|---|
| Leverage | Debt to Asset ratio of firms |
| Profitability | Ratio of EBITDA to Total Assets |
| Total Assets | Borrowers total amount of assets measured in million USD |
| Loan Credit Spread | Spread in base points over benchmark interest rate, LIBOR or EURIBOR |
| Loan Amount Issued | Loan facility amount measured in million USD |
| Loan Maturity | Loan term measured in months |
| Bond Credit Spread at Issuance | Spread between the coupon rate of the issued bond and the interest rate of a US treasury bond with matching maturities |
| Bond Amount Issued | Issued bond amount measured in million USD |
| Bond Maturity at Issuance | Bond term measured in months at issuance |
| Bond Credit Spread (continuous) | Spread between calculated bond yield, based on secondary market transactions, and the interest rate of a US treasury bond with matching maturities |
| Bond Maturity (continuous) | Remaining months to maturity measured continuously |
| Oil Price | Average quarterly WTI spot price |
| Oil Volatility | Quarterly volatility of the WTI spot price |
| Oil Exports | Logarithm of quarterly US crude oil exports measured in thousand barrels |
| D2008 | Dummy variable that equals one in 2008:Q4, zero otherwise |
| D2014 | Dummy variable that equals one in 2014:Q4, zero otherwise |
| Credit spread | Difference between Aaa and Baa corporate bond yield by Moody's obtained from https://fred.stlouisfed.org/graph/?g=D9J. |
| TED spread | Series is calculated as the spread between 3-Month LIBOR based on US dollars and 3-Month Treasury Bill and is directly obtained from https://fred.stlouisfed.org/series/TEDRATE |
| Term spread | Difference between the 10-year Treasury yield and the 3-month T-bill yield. |

# C. Exploratory Data Analysis

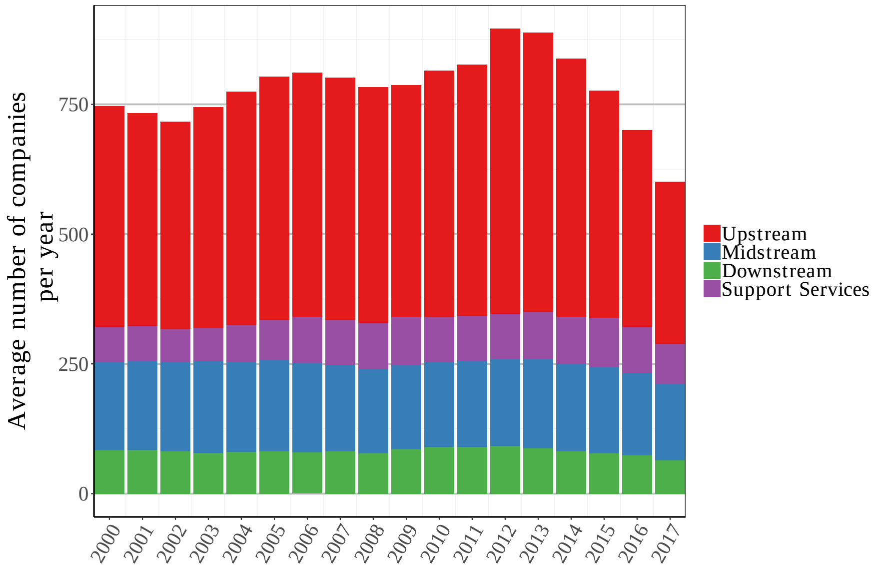



Figure 8: Average number of firms per industry classification and per year.

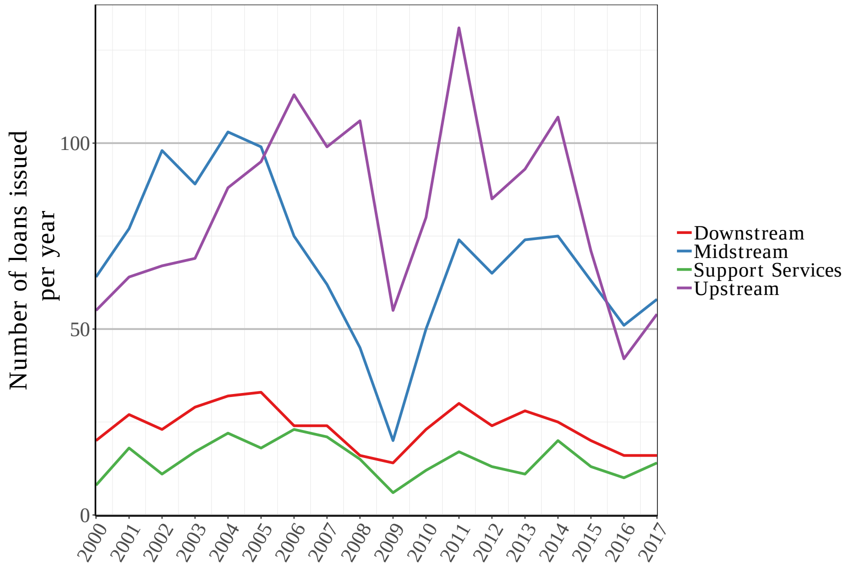



Figure 9: Number of loans issued per industry classification

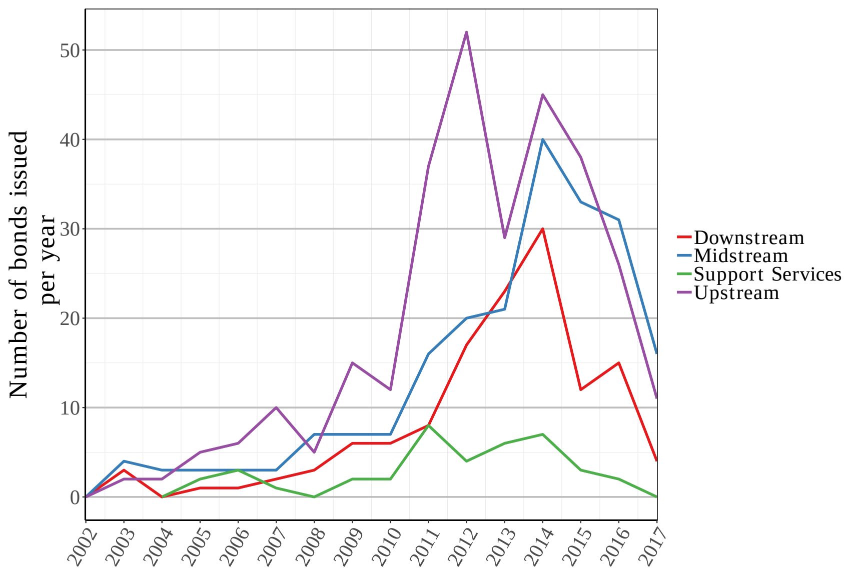



Figure 10: Number of bonds issued per industry classification


Figure 11: Median bond credit spread at issuance bonds per industry.

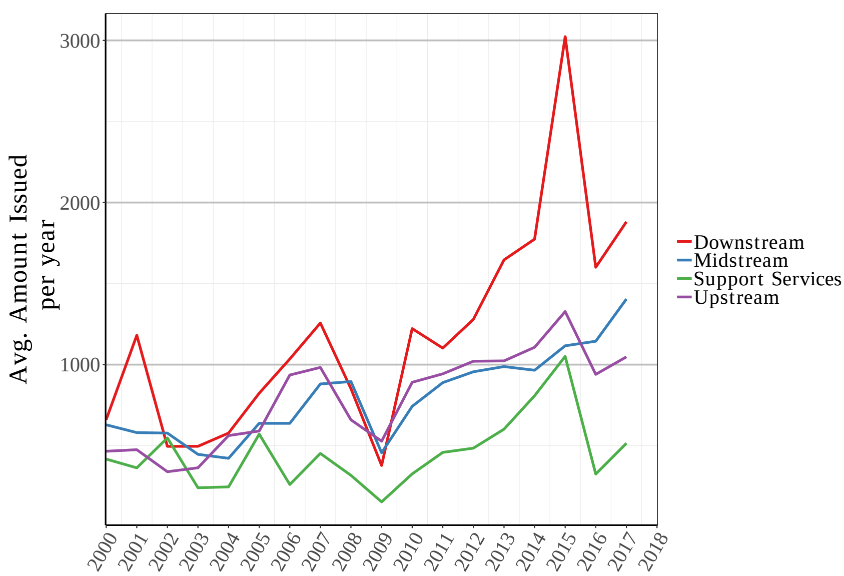



Figure 12: Volume of loans issued per industry classification.

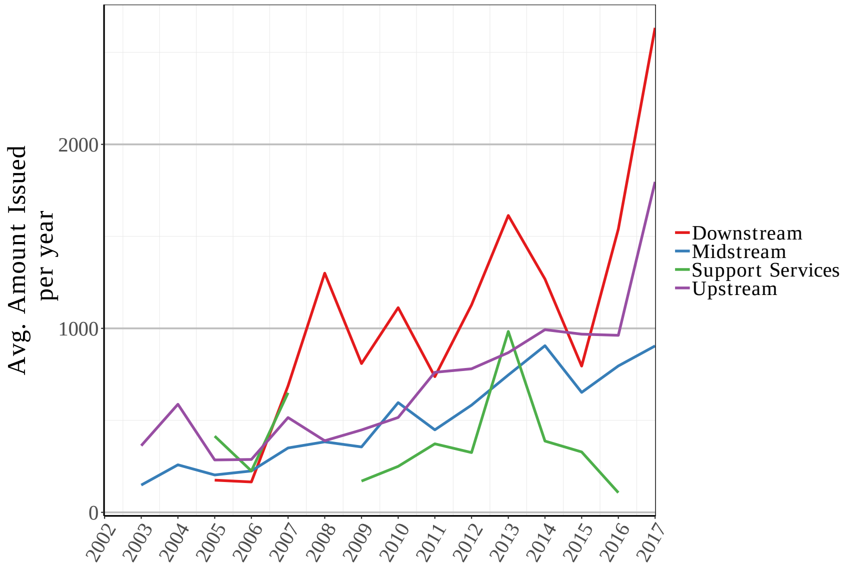



Figure 13: Volume of bonds issued per industry classification.

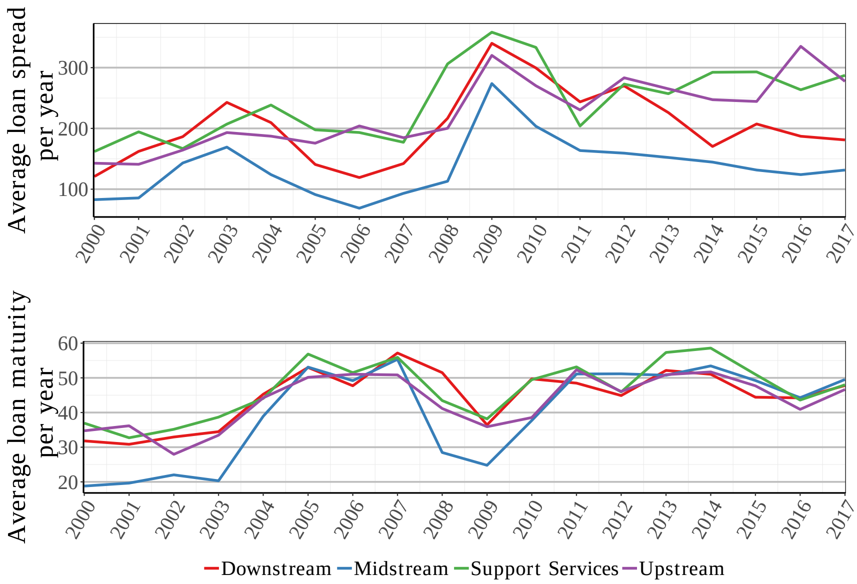



Figure 14: Average loan credit spread at issuance and average maturity of loan facilities.

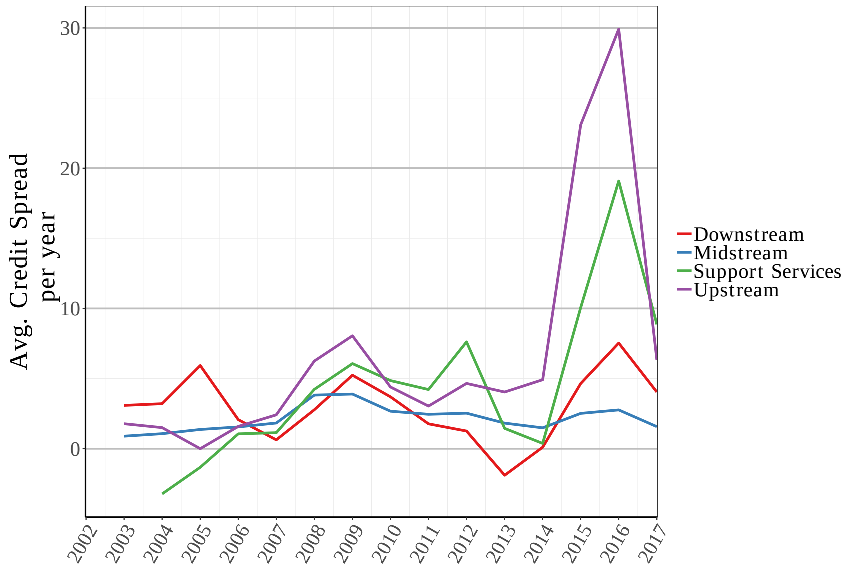
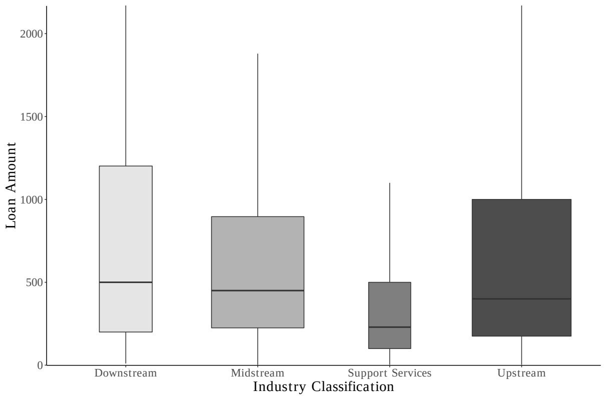
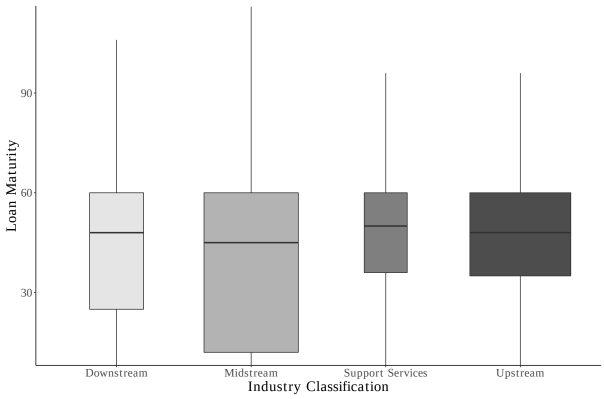


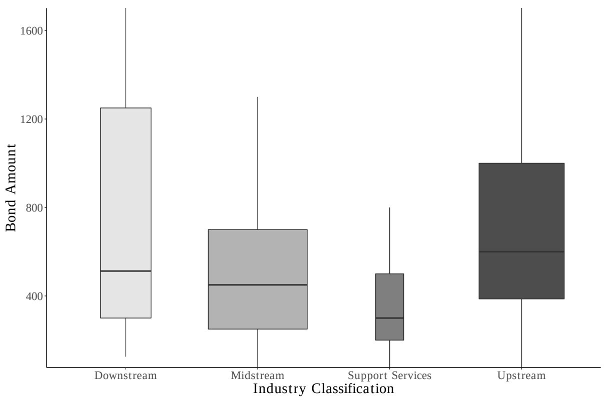

Figure 15: Average credit spread of bonds traded on the secondary market per industry classification.



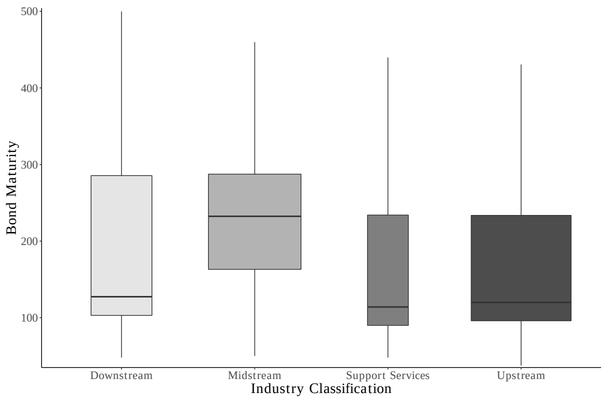

(a) Loan Amount




(b) Loan Maturity




(c) Bond Amount




(d) Bond Maturity

Figure 16: Issued loan and bond amounts and maturities.

### C.0.1. Only Dealscan and Trace Data

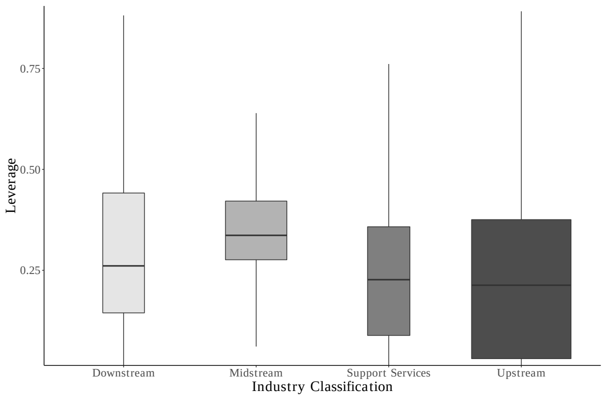
Table 9: Summary Statistics - TRACE and Dealscan for all Firms. All monetary variables are in million US dollars.

| Full Sample | n | $\varnothing T$ | mean | sd | min | Q0.25 | Q0.5 | Q0.75 | max |
|---|---|---|---|---|---|---|---|---|---|
| Total Assets | 615 | 16.91 | 17497.48 | 40607.76 | 0.12 | 1671.86 | 4762.22 | 15653.33 | 410074.00 |
| Total Debt | 615 | 16.91 | 4673.80 | 9123.97 | 0.00 | 573.75 | 1586.00 | 4980.65 | 138237.00 |
| Leverage | 615 | 16.91 | 0.36 | 0.20 | 0.00 | 0.25 | 0.34 | 0.44 | 4.91 |
| Profitability | 601 | 16.77 | 0.03 | 0.08 | -3.62 | 0.02 | 0.03 | 0.05 | 3.79 |
| Loan Credit Spread | 591 | 5.55 | 184.80 | 139.80 | 12.50 | 100.00 | 150.00 | 239.21 | 1325.00 |
| Loan Amount | 592 | 5.58 | 822.53 | 1375.37 | 2.00 | 195.00 | 400.00 | 909.41 | 29762.75 |
| Loan Maturity | 592 | 5.58 | 43.21 | 21.48 | 1.00 | 26.00 | 48.00 | 60.00 | 324.00 |
| Bond Credit Spread | 234 | 30.49 | 2.63 | 2.26 | -10.66 | 1.30 | 2.10 | 3.70 | 11.55 |
| Bond Amount | 219 | 5.11 | 791.32 | 1003.09 | 0.00 | 300.00 | 500.00 | 850.00 | 11000.00 |
| Bond Maturity | 235 | 30.56 | 191.17 | 99.82 | 37.50 | 112.56 | 177.00 | 252.00 | 779.00 |

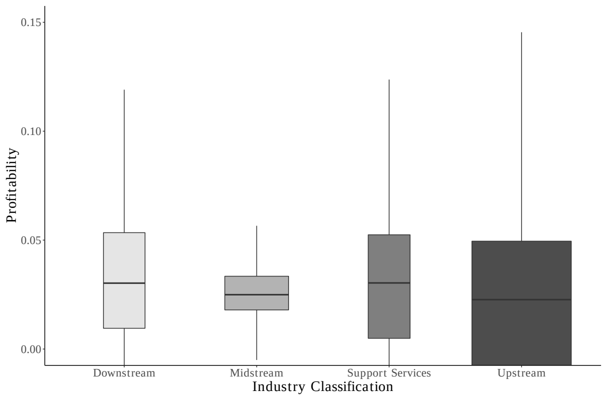
Table 10: Differences in the Sample - TRACE and Dealscan only using all Firms

| | Full Sample (mean) (1) | Pre 2008 (mean) (2) | Post 2008 Pre 2014 (mean) (3) | Post 2014 (mean) (4) | Difference (2) - (3) | Difference (3) - (4) |
|---|---|---|---|---|---|---|
| Total Assets | 17497.4847 | 15635.5969 | 20569.8836 | 23972.8701 | 4934.2867*** | 3402.9866*** |
| Total Debt | 4673.7949 | 4488.9078 | 4978.8867 | 7305.3390 | 489.9789*** | 2326.4523*** |
| Leverage | 0.3625 | 0.3759 | 0.3405 | 0.4332 | −0.0354*** | 0.0928*** |
| Profitability | 0.0284 | 0.0275 | 0.0300 | 0.0092 | 0.0025* | −0.0208*** |
| Loan Credit Spread | 184.8016 | 163.4243 | 230.8245 | 204.2390 | 67.4002*** | −26.5854*** |
| Loan Amount | 822.5305 | 799.7916 | 871.8274 | 1254.4782 | 72.0359 | 382.6508*** |
| Loan Maturity | 43.2096 | 41.9909 | 45.8515 | 48.6087 | 3.8606*** | 2.7572*** |
| Bond Credit Spread | 2.6297 | 2.3878 | 2.9411 | 3.3838 | 0.5533*** | 0.4427*** |
| Bond Amount | 791.3209 | 776.7861 | 808.5527 | 1071.5869 | 31.7666 | 263.0342*** |
| Bond Maturity | 191.1677 | 203.5953 | 175.1264 | 154.0742 | −28.4688*** | −21.0522*** |

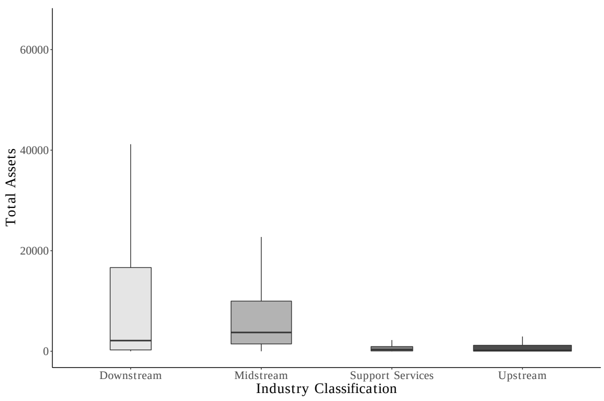
*Note:* *p<0.1; **p<0.05; ***p<0.01




(a) Leverage




(b) Profitability




(c) Total Assets

Figure 17: Boxplots for the exogenous variables of companies with a syndicated loan or a bond.

# Chapter 5 Benford's law and its application to detecting financial fraud and manipulation.

**Reference for this Handbook Chapter:**

Bannier, C., C. Ewelt-Knauer, P. Winker and J. Lips (In press): "Benford's law and its application to detecting financial fraud and manipulation." In: *Corruption and Fraud in Financial Markets: Malpractice, Misconduct and Manipulation.* Ed. by C. Alexander and D. Cumming. John Wiley & Sons, Ltd.

# Benford's law and its application to detecting financial fraud and manipulation.

Christina Bannier Corinna Ewelt-Knauer Johannes Lips*
Peter Winker

This version: September 2019

This article is going to be published as a chapter in the handbook "Corruption and Fraud in Financial Markets: Malpractice, Misconduct and Manipulation." edited by Carol Alexander and Douglas Cumming.

*Department of Economics, Justus Liebig University of Giessen, Licher Str. 64, 35394 Giessen, Germany. E-mail: Johannes.Lips@wirtschaft.uni-giessen.de

# 1. Introduction

Large data sets may contain relevant information about substantive relationships, but also provide indications of at potential fraud, malpractice and manipulation. In the digital age, the amount of large data sets available for analysis is growing exponentially. Consequently, statistical methods to analyze such data sets in order to detect potential inconsistencies regain interest.

We will focus, exemplary, on a classical method used in this context, namely digital analysis based on Benford's law. When studying this approach more thoroughly, it becomes obvious, that statistical methods cannot be expected to provide unequivocal evidence. This conclusion also applies to more refined methods, e.g., based on machine learning including deep learning. All these data driven methods can only provide signals. These signals bear the risk of marking an incident as suspicious, although it is not, or of missing a real case of fraud. This also cuts down the usefulness of these methods in legal proceedings.[1] Thus, statistical methods such as Benford's law are valuable tools for providing first indications of potential misconduct. However, without further evidence, they do not provide sufficient proof for a conviction. Therefore, it might be advisable to focus digital analysis on incidents where a priori a higher likelihood of fraud could be expected so that the risk of false alarms will be reduced. Alternatively, such methods might be used as pre-screening procedures.

Even more important than taking the risk of misleading signals into account is the insight into detection methods fraudsters may obtain. Thus, if people are aware that some financial data will be inspected with regard to its fit to some distributional assumption such as Benford's law, one might even expect that manipulated data fit Benford's law better than real data, at least after some adjustment period. Therefore, it appears important to consider the predator-and-prey perspective when analyzing the performance of data-based procedures for fraud detection. In fact, a reliable procedure should still work even if the potential fraudster is aware of its use. Otherwise, procedures must be elaborated in a way that they keep pace with the fraudsters, a scenario which is more likely.

Inspired by the observation that the first pages of logarithmic tables wear out quicker than the last ones, the discovery of Benford's law dates back to the 19th century, when Newcomb (1881), in a short note, provided a mathematical model for the distribution of the first significant digit of numbers. In non-technical terms, it states that smaller values of the first digit occur more often than larger values. The frequency decreases monotonically from about 30% for a 1 to less than 5% for a 9 as leading digit, while a uniform distribution would predict each digit to occur with the same frequency of about 11.1%. Not only did Benford (1938) provide a formal representation, but also empirical evidence for the law, based on 20,229 observed numbers from many different data-generating processes.[2] Hill (1995) provides a sound statistical base for Benford's law and, consequently, together with technological progress, contributed to an increasing number of applications. A comprehensive literature review of applications and extensions of Benford's law up to the early 2000s is provided by Hürlimann (2006).

According to a non-representative survey among practitioners Bierstaker et al. (2006) using digital analysis software which addresses conformity with Benford's law was rated as quite

[1] See, for example, a decision by the Lower Saxony Finance Court rejecting the use of digital analysis as proof of manipulations (Niedersächsisches Finanzgericht 15. Senat, 17. November 2009, Az: 15 K 12031/08). Other German court decisions also refer to the normal or uniform distribution as potential benchmarks (FG Münster, 05.12.2002 - 8 V 5774/02 E,G,U, FG Münster, 14.08.2003 - 8 V 2651/03 E,U, FG Düsseldorf, 13.04.2004 - 11 V 632/04 A(U), FG Düsseldorf, 31.03.2008 - 14 V 4646/07 A(E,G,U,H(L))).

[2] Diaconis and Freedman (1979) provide convincing evidence that Benford himself manipulated part of his data to obtain a better fit to the theoretically assumed distribution.

effective for fraud detection in an accounting framework. In the same survey, it was listed number 33 out of 34 fraud prevention methods regarding its actual implementation. Bierstaker et al. (2006) argue that this might be due to the resources required for implementing digital analysis software. Thus, there might be room for a broader implementation.

Considering applications in financial markets, for example, Durtschi et al. (2004) discuss when Benford's law might be an effective tool to detect fraud and when it might not be expected to perform well. In particular, they stress that real data conform to Benford's law if they are combinations of numbers from different sources as is often the case with aggregate accounting information. By contrast, when only few observations are available, when the share of manipulated observations is small, or when the source of real data does not satisfy Benford's law, the tool loses efficiency. In our case study, we consider time series with only a small number of observations for each period. Thus, we analyze to what extent Benford's law is informative in this situation and suggest alternative methods.

This chapter contributes to the literature on digital analysis in two ways. First, it provides an overview of applications of Benford's law in the fields of accounting, controlling, taxation, finance, and related areas with a focus on opportunities and limitations. Second, the application to LIBOR (London Interbank Offered Rate) data will demonstrate whether the manipulation of LIBOR data might have been detected earlier under the use of Benford's law or some generalization and which alternative approaches could be used in this and similar settings.

The remainder of this chapter is organized as follows. Section 2 describes Benford's law and some basic limitations and straightforward generalizations to circumvent the limitations. The following Section 3 summarizes applications described in the literature in the domains of accounting, controlling, taxation, and finance. The application to LIBOR data is presented in Section 4, while Section 5 draws some policy conclusions on the use of digital analysis for fraud detection. Concluding remarks and an outlook for future research in the field are provided in Section 6 .

## 2. Benford's law and Generalizations

### The Basic Principles of Benford's Law

Newcomb (1881) had already discovered that the leading significant digits (meaning the first non-zero digit, regardless of the decimal separator) of naturally occurring numbers are not uniformly distributed, but rather follow a logarithmic distribution. The rediscovery of this property and the empirical evidence provided by Benford (1938) triggered research in this area and led to the attribution of this law to Benford. It is, thus, another example for Stigler's law of eponomy, which posits that many scientific discoveries are attributed and named after people other than their respective originators (Goodman 2016).

We start with the most elementary version of Benford's law, which considers only the first significant, i.e., non-zero digits. For example, the numbers 0.123, 123 and 123,000 all share the first significant digit of 1. Consequently, the set of possible outcomes for the first significant digit $d_1$ is given by the set $\{1,2,...,9\}$. Benford's law assumes that the probability distribution of a randomly selected first digit $D_1$ is given by

$$Prob(D_1 = d_1) = log_{10}(d_1 + 1) - log_{10}(d_1) = log_{10}\frac{d_1 + 1}{d_1}.$$

To provide an example for the calculation, according to this version of Benford's law for the first digit, the probability of observing a 1 as the first significant digit in a number is given by

$Prob(D_1 = 1) = log_{10}(\frac{1+1}{1}) = 0.3010$. Thus, it is straightforward to calculate the theoretically expected frequency distribution of the first significant digit according to Benford's law (see Table1) and use it as a benchmark against which the observed empirical frequency of the digit can be tested.

### Illustration of Benford's Law

To illustrate the application of Benford's law we apply it to a data set similar to one of those which were part of the original illustration in Benford (1938). Namely, we analyze the distribution of the first digits for house numbers in Great Britain which are part of the OpenStreetMap (2018) database. This results in a total of N=93,087 observations. [3] To compare the distribution of first digits with Benford's law, the first digit is extracted from each house number and the relative frequency is calculated.

Table 1 provides the expected frequencies for each possible first digit in the second column and the empirical frequencies found for the house numbers in the third column. The final column shows the absolute difference between the two frequencies. The similarity of both sets of frequencies is striking, which is also supported by the graphical representation in Figure 1. A possible explanation for the close conformity, implying high frequencies for low digits, is that in Great Britain the scheme of numbering houses in most cases starts with 1 for houses closest to the city center. Thus, lower digits naturally have a higher prevalence than all of the larger digits, since the length of streets is finite and, whenever you have a 4, you had a 1 before, whenever you have a 22, you had 10, 11,...,19 before etc. However, only 11.6% of the numbers exceed 100 and only 20 or 0.02% of those are larger than 1000.

| Digit | Benford Distribution | Empirical Distribution | Absolute Difference |
|---|---|---|---|
| 1 | 0.3010 | 0.3066 | 0.0056 |
| 2 | 0.1761 | 0.1828 | 0.0067 |
| 3 | 0.1249 | 0.1270 | 0.0021 |
| 4 | 0.0969 | 0.0953 | 0.0016 |
| 5 | 0.0792 | 0.0782 | 0.0010 |
| 6 | 0.0669 | 0.0654 | 0.0015 |
| 7 | 0.0580 | 0.0543 | 0.0037 |
| 8 | 0.0512 | 0.0480 | 0.0032 |
| 9 | 0.0458 | 0.0425 | 0.0033 |

Table 1: Empirical distribution of first digits of house numbers in Great Britain from the OpenStreetMap (2018) database (third column) and the theoretical distribution according to Benford's law (second column). The observed absolute differences between the two distributions (right column) are rather small. Thus, the first digits for house numbers appear to be a prime example for naturally occurring numbers closely following Benford's law.

### Testing for Conformity with Benford's Law

The comparison of the empirical distribution of digits with the distribution implied by Benford's law is either carried out by means of graphical presentations of both distributions as shown in Figure 1 or based on statistical procedures. In the latter case, Pearson's chi-squared

[3] It has to be noted that the OpenStreetMap (2018) data set does not include all house numbers in Great Britain, but only covers a subset.

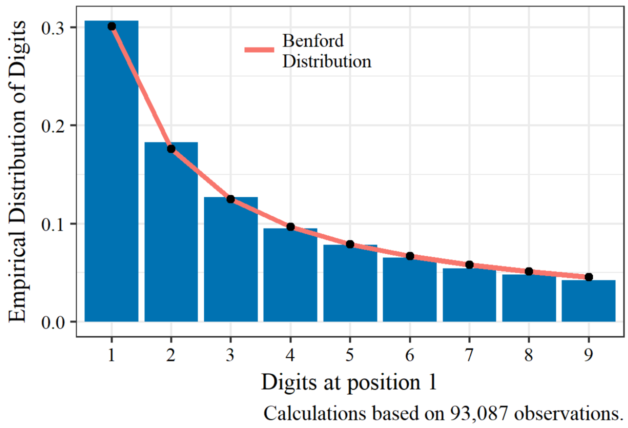



Figure 1: The graphical display of the relative frequencies of first digits of house numbers in the UK (blue) and the theoretical expected distribution according to Benford's law (red) indicated the close conformity to Benford's law.

($\chi^2$) test is an appropriate procedure. In the present setting, it tests the null hypothesis that the observed digits are generated in accordance with Benford's law. To this end, the empirical frequencies of all digits are compared with their theoretical counterparts, the probabilities from Benford's law. The test statistics is calculated by the below formula:

$$\chi^2 = \sum_{i=1}^{9} \frac{(ac_i - ec_i)^2}{ec_i}$$

$ac_i$ corresponds to the actual count of digit $i$, $ec_i$ represents the expected count which is given by the probability of observing an $i$ under Benford's law multiplied with the number of observations $N$, i.e., $ec_i = Prob(D_1 = i) \times N$. The $\chi^2$ test statistic takes on large values if these differences are large, while it should be small when the data are generated in accordance with Benford's law. Asymptotical critical values for the $\chi^2$ test statistics are obtained from the $\chi^2$-distribution with 9 degrees of freedom.[4] For the house number example, the $\chi^2$ test statistics amount to 105.64. Provided that this value is larger than the critical value at the 5% level (and also above the 1% and 0.1% level), the null hypothesis of conformity of the distribution of the first digits with Benford's law has to be rejected despite the seemingly good fit shown in Figure 1.

In the context of searching fraud, malpractice, or misconduct, the fact that failing to reject the null hypothesis, i.e., apparent conformity of the data with Benford's law does not automatically imply that numbers occurred naturally and, thus, are not manipulated has to be considered. Alternatively, the number of observations might be too small to allow for a significant result, or the data follow Benford's law despite manipulations as the fraudsters might be aware of tests conducted.

On the other hand, large deviations resulting in a rejection of the null hypothesis do not prove manipulations as well. For example, when testing at a significance level of 5%, such rejections occur at a rate of 5% even if the data follow Benford's law. Of course, a significant test result

[4] These critical values are 16.919, 19.023, and 21.666 for a level of significance of 0.1, 0.05 and 0.01, respectively. For the analysis of second or further significant digits, asymptotical critical values are obtained from the $\chi^2$-distribution with 10 degrees of freedom: 18.307, 20.483, and 23.209 for a level of significance of 0.1, 0.05, and 0.01, respectively.

might also be owed to the fact that the true distribution does not comply with Benford's law for the specific data considered. Thus, finding significant deviations should only be a starting point for further investigations in order to find out if they can be attributed to fraudulent behavior or if they arose otherwise.

Another important caveat in the application of the $\chi^2$- test statistic is the problem of excess power, as described by Nigrini (2012). This means that in cases, where the number of observations becomes large, even small and practically irrelevant deviations from Benford's law will result in test statistics which exceed the critical values at the usual significance levels. This leads to the rejection of the null hypothesis and thus to the false conclusion that there is a relevant – and not just statistically significant – deviation from Benford's law. When applying the $\chi^2$- test statistic in an attempt to detect financial fraud, it is important to have a large sample available for analysis. This unwelcome aspect might be addressed by using stricter levels of significance, e.g., 1%, 0.1% or even below if the number of observations becomes very large.

An alternative indicator, which is used by Shi et al. (2018), Judge and Schechter (2009) and Schündeln (2018) amongst others, is based on the Euclidian distance between the distributions and calculated as

$$d^* = \frac{1}{M}\sqrt{\sum_{i=1}^{9}\left(\frac{ac_i}{N} - \frac{ec_i}{N}\right)}$$

where $M$ is the maximum possible distance, which is obtained if all observed digits are equal to 9.The normalization with $M$ guarantees that the values of $d^*$ fall in the interval [0,1]. Values close to zero indicate conformity with Benford's law. While the measure has the advantage of being not sensitive to sample size and appears useful for comparing changes over time, it appears difficult to derive thresholds for deciding if a distribution follows Benford's law or not.

A further alternative has been proposed by Drake and Nigrini (2000). It might be used to determine the degree of conformity that an empirical distribution exhibits, namely the mean absolute deviation (MAD). This measure is the average of the absolute differences between the empirically observed relative frequencies and the ones determined in accordance with Benford's law. According to Drake and Nigrini (2000), it has the advantage of being less affected by the number of observations used in the analysis. Unfortunately, however, the MAD indicator does not follow a well-defined probability distribution. Therefore, Drake and Nigrini (2000) provide some simulated critical values and suggest the following conclusions. When considering only the first significant digit, they label values in the range up to 0.006 as "close conformity", values above 0.006 and below 0.012 as "acceptable conformity", values above 0.012 and below 0.015 as "marginally acceptable conformity" and only values above 0.015 as "nonconformity".[5] For the house number example, the value of the MAD is obtained by adding up the absolute differences listed in the last column of Table 1 and dividing this sum by the number of digits considered (9). The resulting value is 0.0032 and thus, based on the categories Drake and Nigrini (2000), we might assume "close conformity" to Benford's law.

### Considering further Digits with Benford's Law

Similar arguments that apply to the first digit also hold true for further significant digits. In particular, for the second significant digit $d_2$, the probability distribution over the set of possible values $\{0,1,...,9\}$ is calculated as follows:

[5]The relevant intervals for second digits and the combination of first-two or first-three digits are provided in Table 3 in the Appendix.

$$Prob(D_2 = d_2) = \sum_{i=1}^{9} log_{10}\left[1 + \frac{1}{(10i + d_2)}\right].$$

Accordingly, the probabilities decrease from 11.97% for a zero to 8.50% for a 9 as the second digit. However, the differences are much smaller compared to those of the first digit.

To illustrate this, we again use the house number data. The comparison between theoretical probabilities and empirical frequencies is shown in Figure 2. A first point to note is that there are a fewer observations, only 70,446, for the second digit, which is due to the fact that some (small) house numbers only have one digit. The results for testing the conformity with Benford's law for the second digit are similar to the results obtained for the first digit, since with a $\chi^2$- test statistic of 155.54, the null hypothesis of conformity between the empirically observed and theoretically expected frequencies has to be rejected. In contrast to this, the value of the MAD indicator of 0.00383 again indicates "close conformity" according to the intervals provided in Table 3.

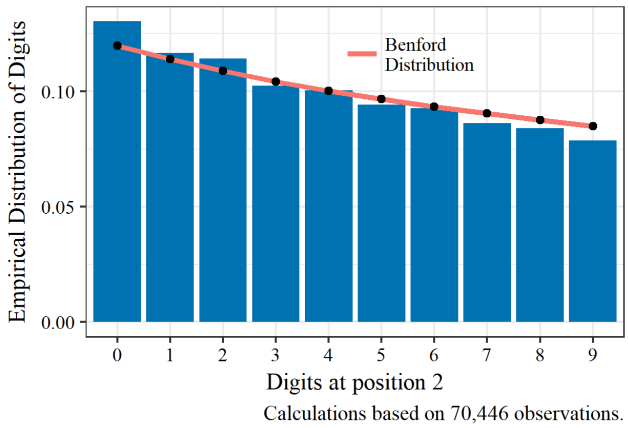



Figure 2: The graphical display of the relative frequencies of second digits of house numbers in the UK (blue) and the theoretical expected distribution according to Benford's law (red) confirm the observation and still indicates the close conformity to Benford's law.

The following formula generalizes the probability distribution for all subsequent digits. For the n-th position in a number, when $n > 1$, the probability to observe a digit $d \in \{0, 1, ..., 9\}$ is calculated as follows:

$$Prob(D_n = d) = \sum_{k=10^{n-2}}^{10^{n-1}-1} log_{10}\left[1 + \frac{1}{(10k + d)}\right].$$

It is important to note that, as might be expected, the theoretical distribution for subsequent digits gets closer to uniform when moving further to the right within the number. Already for the fourth significant digit, the probabilities for all digits range between 10.02% for a 0 and 9.98% for a 9, while they would all be 10% in case of a uniform distribution.

Since the first and second (and further) digits are not distributed independently, considering several significant digits in combination makes perfect sense in order to increase the discriminatory power of the analysis. Hill (1995) derives the formal generalization of the significant-digit law for the first $k$ digits ($k > 1$). It states that for positive integers with k digits, all first digits $d_1 \in \{1, 2, ...9\}$ in combination with subsequent digits $d_j \in \{0, 1, ..., 9\}, j = 2, ..., k$, should follow the joint probability distribution

$$Prob(D_1 = d_1, ..., D_k = d_k) = log_{10}\left[1 + \left(\sum_{i=1}^{k} d_i \times 10^{k-i}\right)\right]$$

To provide an example for the calculation according to this general version of Benford's law, the probability to observe a number with the first two significant digits being 1 and 2, e.g. 12, 125 or 1209, is given by $Prob(D_1 = 1, D_2 = 2) = log_{10}(1 + (12)^{-1}) = 0.03476$.

Again, we use the example of house numbers in the UK to illustrate how this analysis is conducted. Figure 3 exhibits the empirical and the theoretically expected frequencies of a combination of the first two digits. It is apparent that the first two digits do not correspond to the generalized Benford's law. This evidence is underpinned by the calculated test statistics, which both result in a rejection of the null hypothesis of conformity with Benford's law. The $\chi^2$- test statistic takes on a value of 59,456, which has to be compared to the critical values of a $\chi^2$- distribution with 89 degrees of freedom (106.469, 112.022, and 122.942 for the 10%, 5% and 1%-level, respectively, and only 147.35 for the 0.01% level). The MAD indicator adds up to 0.004939326. This value clearly falls into the range of nonconformity according to the values provided in Table 3 in the appendix (above 0.0022).

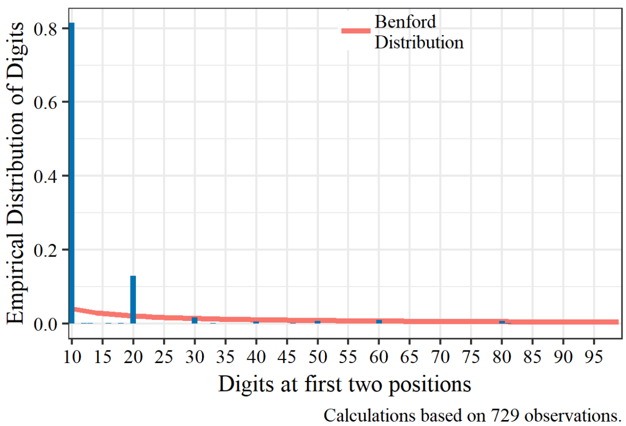



Figure 3: The graphical display of the relative frequencies of the first two digits of house numbers in the UK (blue) and the theoretical expected distribution according to Benford's law (red) indicate that the first two digits do not match the generalized Benford distribution. It is apparent that numbers with the leading digits 10 and 20 are much more frequently observed than any other digit combination.

Recently, Barabesi et al. (2017) propose a hierarchical sequence of tests starting with a joint test on the first k digits and, conditional on rejection, going down to smaller sets. They provide simulated critical values and demonstrate that the procedure helps reducing the rate of false alarms.

### When Do Data Conform To Benford's law?

Benford's law does not apply to all sets of numerical data. Thus, it is important to verify whether all necessary conditions are met. Nigrini and Mittermaier (1997) and Fewster (2009) describe the most essential of these conditions.

A first requirement refers to the range of values the numbers might take on. One of the most basic and most important findings is the fact that Benford's law provides a better fit if the distribution of the numbers covers several orders of magnitude (Raimi 1976; Smith 1997). The order of magnitude of a number can be described using the number of powers of 10 contained therein. The house numbers in our previous application ranging from 1 to 2,473 only cover four orders of magnitude since all numbers lie within the range from $1 \times 10^0$ to $1 \times 10^4$. According to Fewster (2009) a range of six might be sufficient to obtain a good approximation to Benford's law. He also provides an intuitive explanation and several examples for this requirement. Second, the numbers should not be assigned but occur naturally. This implies that Benford's law might not work well considering personal income figures of a certain income tax bracket due to the arbitrary cut-off points. Consequently, the probabilities for certain digits might deviate considerably from Benford's law. Another example for assigned numbers in the same context are social security numbers, which follow a certain pattern as they are the result of human thought. By contrast, the total sales volume of firms should meet this requirement.

Third, as proved by Hill (1995), many naturally occurring numbers, which follow Benford's law, are mathematical combinations of numbers which are unbiasedly sampled from various distributions. This is also the explanation why Benford's law is often used to analyze accounting data at the firm level, since these accounting data are the result of the multiplication of randomly sampled numbers of, for example, quantities and prices and their summation.

In a prior contribution, Durtschi et al. (2004) provide an overview of applications of Benford's law in the context of financial accounting. They also give examples as to when the conditions for the application of Benford's law can be expected to hold. In particular, they stress the relevance of combinations of numbers, e.g., accounts receivable as product of quantity sold times prices, and the advantage of large data sets as even small deviations might become significant. However, as mentioned above, recently, this feature has been judged more ambivalently. Furthermore, a mean larger than the median, i.e., a positive skewness of the distribution, supports conformity to Benford's law (see also Wallace (2002)). Nigrini and Mittermaier (1997) suggest that auditors might assume that the data follow Benford's law when considering "items, such as accounts receivable or payable, inventory counts, fixed asset acquisitions, daily sales, and disbursements". Durtschi et al. (2004) state that Benford's law should not be applied when numbers are assigned, influenced by human thought, containing a large number of firm-specific numbers or built-in minimum or maximum.[6] In addition, they point out that Benford's law and similar methods cannot be applied if fraud is conducted without any record, e.g., theft or kickbacks.

A more rigorous mathematical analysis of the conditions to be met for a set of numbers to follow Benford's law can be found in Boyle (1994), Chenavier et al. (2018), Dümbgen and Leuenberger (2008), Gauvrit and Delahaye (2009), Pinkham (1961), and Wallace (2002).

### Limitations of Using Benford's Law for Identification of Manipulations

The requirements stated above limit the range of possible applications of Benford's law for detecting fraud, malpractice, and manipulation in financial markets. If the distribution of non-manipulated data does not follow Benford's law, the test statistics introduced above will reject the null hypothesis too often and result in false alarms. Consequently, generalizations of the approach, which do not require conformity with Benford's law in the strict sense, might be more appropriate and are discussed below.

Furthermore, even if real data follow Benford's law closely enough, a substantial number of observations for potentially manipulated data is required for a powerful statistical test. This

[6] Lu and Boritz (2005) propose an adaptive version of Benford's law allowing to take artificial cut-off values into account.

limits the applicability of the method as an early warning or real-time check, when only few observations become available in each time period, e.g., aggregate earnings or profits at the firm level. Only after a sufficiently long period of malpractice, a retrospective analysis of the data might become feasible.

When considering fraud and manipulation of data, a further shortcoming of methods relying on Benford's law in the strict sense is their vulnerability under the predator-and-prey perspective, i.e., the observation that impostors will adjust their behavior to the means used in surveillance. If potential manipulators become aware that such methods are commonly used for checking the integrity of data, they can decide either to stop manipulations or to adjust the way they produce manipulated data. Unfortunately, it does not present a challenge to produce data in compliance with Benford's law when using computer tools. There are even websites offering the generation of such data or providing a code for generating the data. For obvious reasons, we do not want to advertise these offers, but it is a rather safe bet to assume that potential manipulators will find and use them.[7] Therefore, relying on Benford's law as a stand-alone instrument for detecting fraud, malpractice or manipulations is not recommendable.

### Generalizations of Benford's Law for Identification of Manipulations

In order to deal with some of the limitations mentioned above, the literature proposes some generalizations and extensions. In the context of accounting, Winter et al. (2012) propose using a modified Benford distribution which makes allowance for accounting limits. The resulting distribution should be closer to observed real numbers and, consequently, reduce the rate of false positives. A different and more general approach followed, amongst others, by Rodriguez (2004) and Hürlimann (2009) is to consider classes of distributions and to decide on the specific benchmark based on the fit to the available data. Again, having a closer approximation to the actual distribution should reduce the risk of false alarms.

Given that the digital analysis and the tests for homogeneity, i.e., equal distributions, are mainly non-parametric, a further straightforward extension consists in considering the empirical distribution of digits as a benchmark. If it can be assumed that the share of manipulated data is small, the distribution of digits over all available observations should represent a good approximation to the true distribution (Schräpler 2011). If the distribution for a subsample differs significantly, this generates a signal to look at the subsample more thoroughly.

Developing this idea further, it could be assumed that part of the available data is correct and could provide the benchmark distribution of digits, while the other part exhibits a different distribution of significant digits. A generalized procedure could try to identify both parts by an optimization procedure over different sample splits attempting to maximize the distance between the distributions of the two parts. However, such a procedure would be computationally complex and could not indicate a priori which of the two parts contains the manipulated data.

All three approaches, generalized parametric distributions, non-parametric approaches based on the empirical distribution of all available observations and non-parametric approaches derived from an optimization procedure share the advantage that the benchmark used for the analysis is not known a priori. Thus, potential manipulators do not know which benchmark to meet unless they have access to all data. Even if this were the case, the sketched optimization approach might still impose a hurdle for generating manipulated data, which might remain undetected. Finally, the idea of combining several indicators as proposed by Bredl, Winker, et al.

[7] In a laboratory experiment, Watrin et al. (2008) provide evidence that the distribution of digits of manipulated numbers becomes more similar to Benford's law when the subjects are informed about its use. However, the differences were not found to be statistically significant.

(2012) in the context of survey data, which focus on different aspects of data quality and not just on the distribution of digits, might help to improve discriminatory power.

# 3. Usage of Benford's Law for Detecting Fraud and Deviant Behavior

Since the initial publication in 1881, the number of published research papers using or analyzing Benford's law has increased considerably, in particular after 2000 with a peak during the financial crisis. Admittedly, following the financial crisis the number of publications decreased. Figure 4 provides a graphical representation of this development and is based on a comprehensive listing by Hürlimann (2006) and Berger et al. (2017).[8]

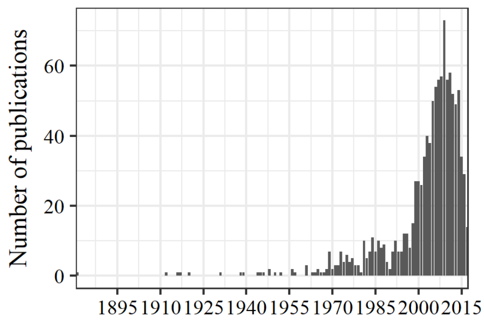



Figure 4: Development of the yearly number of academic publications using or analyzing Benford's law. Initially, the first publication by Newcomb (1881) did not entail a lot of follow-up publications, and also the publication by Benford (1938) increased the interest in this topic only slightly. Only the availability of increasing computational resources and large digital data sets lead to an increasing usage of Benford's law in order to detect fraud and deviant behavior resulting in a growing number of publications during the 1990s and the early 2000s. The figure is based on data provided by Berger et al. (2017).

Even though the applications of Benford's law are not restricted to the analysis of fraud and manipulations in financial markets, the number of publications in this particular area is quite substantial. Therefore, the literature review presented in this section has to be selective. It is organized along the topics (1) forensic accounting, auditing and internal control systems, (2) finance, and (3) surveys and research.

### Forensic accounting in the context of auditing, internal control systems, and taxation

Detecting fraud, malpractice, or manipulations is one of the main tasks of audits and internal control systems. Carslaw (1988) was the first who applied Benford's law on accounting data by hypothesizing that when important key indicators, such as net income, are slightly below specific 'psychological boundaries', managers typically round these numbers up to evoke the impression that they are larger (this rounding up phenomenon is also known as '$1.99 pricing phenomenon'). For instance, a net income of $187,000 or $9.54 million is rounded up

[8] Hürlimann (2006) compiled an impressive list, which is continuously updated and provided at http://www.benfordonline.net/list/chronological by Berger et al. (2017) and as of (09/25/2018) lists more than 1,000 publications.See alsoMir and Ausloos (2018) for an in-depth bibliometric analysis of citations to Newcomb (1881) and Benford (1938).

to \$200,000 or \$10 million, respectively. This technically implies an increase of second digit 0s and a shortage of second digit 9s. Indeed, Carslaw (1988) found in a data sets from New Zealand more second digit 0s and fewer second digit 9s than expected. This result is also in line with the observations of Thomas (1989), who found an excess of second digit 0s based on a rounding-up phenomenon in quarterly U.S. net income data, when companies report a positive net income. In contrast, when companies have negative net incomes, he observed fewer second digit 0s and more second digit 9s, meaning that firms avoid rounding losses. This effect also holds when analyzing net income on a per share level (earnings per share – EPS). More precisely, multiples of 5 cents and 10 cents are more often used than expected when reporting EPS, while EPS ended less often than expected on the ending digit 9. Recently, Henselmann et al. (2015) add to this stream of literature by showing a high degree of deviation from Benford's law in suspect firm years (defined as firm-years just meeting or beating zero (last-year) earnings), compared to other firm-years that clearly miss or beat the thresholds. In the same vein,Amiram et al. (2015) inter alia demonstrate that restated financial statements conform more closely to Benford's law compared to the respective misstated financial statements. In addition, they show that earnings persistence decreases, when divergence from Benford's law increases.

The detection of fraudulent behavior conducted by individuals is mainly based on Hill (1988), who ran an experiment with 742 students asking them to randomly guess a six-digit number. The results show that the first digit 1 occurs more often than expected, whereas the numbers 8 and 9 occur less often than expected as first digit based on Benford's law. Moreover, the second digits are distributed more uniformly than the first digits. This idea is picked up by Nigrini (1994) – as cited in Nigrini and Mittermaier (1997) – by hypothesizing that numbers made up by people will not conform to the expected digital frequency. In detail, he used cases of payroll fraud and found that fraudulent numbers deviated significantly from Benford's law. This effect is even more pronounced when individuals have more routine in faking numbers, because they get used to certain numbers. Nigrini and Mittermaier (1997) introduced Benford's law to a specific audit context. In detail, they show how digital and number tests can help to assess the authenticity of lists of numbers in the planning stages of audits. Durtschi et al. (2004) provide a practitioner's guide for auditors on when and how Benford's law can help to detect suspect accounts based on all available data. Nigrini and Miller (2009) introduce a 'second-order Benford test' to the auditing literature to find errors in transactional data. They show that digits of the differences between amounts approximate the frequencies of Benford's law for of the most data sets when the amounts are sorted from the smallest to the largest amounts. Moreover, Nigrini (2012) demonstrates how Benford's law can help to detect inconsistencies within the accounts receivables based on firm-wide invoice-level data. More recently, Benford's law has been used to detect target-driven earnings management especially by Ullmann and Watrin (2017).

Focusing on taxation,Christian et al. (1993) investigate tax returns to detect whether tax payers reduced taxable income from above a tax table bracket to a taxable income below a tax table bracket (so called "secondary evasion") to get a lower tax rate. They assume that that the ending digits of taxable incomes should be uniformly distributed over the 00 to 99 range and that the expected frequency of the third, fourth and fifth digits represent a near-uniform distribution. The results indicate a clear bias toward taxpayers having taxable incomes slightly below a specific tax table bracket. Regarding tax compliance, Nigrini (1996) shows that there was a bias towards lowering taxable income by using (1) low digits for interest received and high digits for interest paid. Watrin et al. (2008) show that Benford's law is a valuable tool to select firms for an on-site tax audit if the non-manipulated data conforms to the Benford distribution resulting in more efficient and effective on-site tax audits.

Barabesi et al. (2017) illustrate their hierarchical testing procedure to trade data reported by Italian traders. Since customs and value added tax are calculated based on these declared values, there is a considerable incentive to underreport. On the other hand, money laundering schemes might provide incentives to increase reported valuations. For two selected traders, the authors demonstrate that their hierarchical testing procedure allows for a better differentiation between false alarms and indeed suspicious cases.

**Finance**

In the finance context, Ley (1996) is one of the first to employ Benford's law. He finds that the series of daily returns on two of the most important U.S. stock indexes, the Dow-Jones Industrial Average Index (DJIA) over the period 1900 to 1993 and the Standard and Poor's Index (S&P) over the period 1926 to 1993, follow Benford's law. His work followed earlier research studying the importance of certain numerical values of major stock indexes (Koedijk and Stork 1994; Ley and Varian 1994). While earlier papers mainly focused on the psychological impact of specific index levels (such as, e.g., multiples of 1,000) in order to predict future stock market movements, the application of Benford's law was subsequently employed to examine manipulative behavior on financial markets. In this respect, Corazza et al. (2010) discuss the S&P 500 stock market index and find that sequences of trading days not confirming with the Benford distribution are rather short. Moreover, they observe that days on which the stock index distribution does not follow Benford's law are related to extreme market events such as the attack on the Twin Towers on September 11, 2001.

Relatedly, Rauch, Göttsche, Brähler, et al. (2011) use Benford's law to check the accuracy of governmental macroeconomic statistics relevant for compliance with the Stability and Growth Pact criteria of the European Union. They consider all relevant data from the 27 EU member states from 1999 to 2009. As macroeconomic data comes from different sources with different distributions, a Benford distribution is to be expected. Moreover, since the Benford distribution is the only distribution of first significant digits that is scale invariant, this property makes Benford's law particularly helpful in a macroeconomic context where data need to be converted from one currency to another. As might have been expected, Rauch, Göttsche, Brähler, et al. (2011) find the data reported by Greece to show the greatest deviation from Benford's law among all countries in the Euro area. Deleanu (2017) comes to a similar conclusion with respect to a data sets of self-reported indicators of compliance and efficiency in the fight against money laundering among European Union member states. Her results, based on Benford's law, hint at potential manipulations of these indicators for countries that faced sufficient incentives and opportunities to misinform the community about their efforts to fight money laundering.

Nye and Moul (2007) apply Benford's law as a tool for assessing the quality of macroeconomic indicators on a broader scale. They focused on the GDP series of OECD countries and of certain African nations and find that only a subset of the data - particularly from the developing countries - shows non-conformity consistent with deliberate manipulation of the underlying series. In a follow-up study, Gonzalez-Garcia and Pastor (2009) enlarge the data set to 80 countries and report only little indication of a rejection of the first-digit law for most series. Even more importantly, they show that the observed deviations from Benford's law may be a result of structural breaks captured in the data series and caution against interpreting them as a signal of poor quality in macroeconomic data.

In contrast to this, current studies have found evidence in favor of strategic manipulation of macroeconomic data by governments. Michalski and Stoltz (2013), for instance, consider the balance of payments data from the IMF between 1989 and 2007. Since some countries have already been caught misreporting their information (e.g., Ukraine) to this data set, there is some valid reason to re-consider the full data set. The authors find that there is evidence from Benford's law that countries which are more vulnerable to capital flow reversals (e.g., those with fixed exchange rates or countries with current and fiscal deficits) show irregular, non-Benford behavior of the first digits in their data series. Interestingly, China, which has often come under scrutiny for its official statistics on GDP development, has been shown to largely conform with Benford's law (Holz 2014).

Further applications of Benford's law have been inspired by observations of peculiar price movements in the financial crisis starting in 2007. Irregularities on reference rates such as the LIBOR have questioned the integrity of the interbank market and led to investigations using Benford's law (Abrantes-Metz, Villas-Boas, and Judge 2011; Mollenkamp 2008; Mollenkamp and Whitehouse 2008). Our case study in Section 4 demonstrates how Benford's law can be employed in this context. Further analyses indicate that other reference rates, such as EURIBOR or TIBOR, may have been affected by manipulation as well (Rauch, Göttsche, El Mouaaouy, et al. 2013).

Hofmarcher and Hornik (2013) employ Benford-like distributions to CDS market data and find consistency for the US-CDS market but huge fluctuations in the distributions of first significant digits for the European market during the financial crisis. They attribute these differences to the reorganization procedures in bankruptcy that are much more lenient on borrowers in the US as compared to Europe. Ausloos et al. (2016) stress the importance of employing Benford's law for checking the credibility of CDS data whose pricing processes are often opaque and trading volumes highly variable. They examine the daily sovereign CDS spreads of thirteen European countries between 2008 and 2015. Their results show that Benford's law tends to be violated more often in the more liquid CDS instruments and in the core European countries' (France, Germany, United Kingdom) CDS. The authors nevertheless point out that the development of CDS spreads is strongly affected by liquidity constraints, which might have a stronger impact than the underlying sovereign risk perception.

However, manipulations of interest rates appear to reflect only the tip of the iceberg as other benchmarks, e.g., deriving from commodity, currency, or other financial markets, seem to be affected as well (Wheatley 2012) . In this respect, El Mouaaouy (2018) examines FX benchmarks via the use of Benford's law. He finds anomalies in several foreign exchange rates that were also exposed to the LIBOR rigging and blames coordinated interventions of market participants for influencing the benchmark rates to their advantage. In a similar analysis,Stenfors (2018) also utilizes Benford's law as a screening device in order to detect artificial patterns in FX data of USD/JPY and USD/NOK swap markets. The empirical results reveal patterns that suggest that some form of coordination between market participants has taken place.

El Mouaaouy and Riepe (2018) employ Benford's law to analyze capital allocation processes within firms. They consider publicly available segment-level data from the German banking industry from 2004 to 2011 and show that managerial interventions lead to stronger deviations from Benford's law with increasing complexity of the underlying business model. Similar results are obtained by Lin et al. (2018) with regard to earnings management in Taiwanese firms. They find stronger deviations from Benford's law for firms in which board members have the power to increase their own pay substantially.

**Surveys and Research**

Finally, two further fields of applications, which are at least indirectly related to financial markets, warrant a brief mentioning. The first are surveys, which often focus on topics related to financial markets, for instance investor expectations or household finance. Bredl, Storfinger, et al. (2013) provide a recent review on the usage of Benford's law in this context. Second, there is a line of research testing whether academic research itself might be affected by possible manipulations or fraud.

With regard to surveys, in particular those on investor expectations, a first observation is that they do not lend themselves to a digital analysis as they include only a small number of questions, out of which a large part is rather qualitative ("stock prices will increase", "stay constant", "fall"). Hence, there are not enough metric variables with several digits in the data set required for a good estimate of the empirical distribution of digits for a single interview or at least for all interviews conducted by one interviewer. Furthermore, the metric variables included in the data set should follow Benford's law in real data, which might be the case for many variables included in financial markets or household finance surveys. However, it is not sufficient that the real data conform to Benford's law. The actual numbers reported by respondents have to comply with Benford's law, too. At this level, it is well known that respondents tend to round numbers in a way that at least for non-leading digits a clustering at 0 and 5 can often be observed (see, e.g., Schräpler (2011)). Obviously, this should not constitute a signal for possible fraud by interviewers. Therefore, care should be taken when deciding which digits from which variables to use for comparison with Benford's distribution in survey data.

Judge and Schechter (2009) and Schräpler (2011) also use Benford's law for identifying suspect observations or interviewers. Both employ household surveys including a large number of metric variables related to agricultural production and monetary income components. They succeed in identifying suspicious cases. According to Schräpler (2011), a follow up done for one interviewer actually resulted in discovering falsifications, which remained unnoticed as more traditional methods were used. In line with one of the generalizations of Benford's law discussed in Section 2, Schräpler (2011) suggests using the distribution of digits over all available observations as benchmark instead of Benford's distribution. Assuming that most observations are real, this might provide a better approximation to the underlying distribution, in particular when rounding is relevant. However, this method has to be refined in case of a large share of falsifiers as, e.g., found by Bredl, Winker, et al. (2012). In a related application, Schündeln (2018) compares the conformity to Benford's law for repeated interviews on consumption expenditures in the Ghana Living Standards Survey. He finds smaller deviations for early responses and takes this finding for an indication of higher data quality resulting from early measurements.

In particular, if the analysis is based on only a limited number of metric variables as might be the typical setting in financial markets surveys, the standard approach will not withstand the predator-and-prey perspective. If falsifiers are aware of controls based on Benford's law, they have to exert just a minor additional effort to generate falsified data in line with Benford's law. Then, the test will lose its power for identifying suspicious cases. In order to overcome this limitation, Bredl, Winker, et al. (2012) combine several indicators and used the multivariate distribution for discriminating groups of interviewers. In this way, the benchmark becomes higher dimensional and depends on the other interviewers, making it much more difficult for a cheater to replicate it Winker (2016).

Finally, given that the peer-review process typically does not include a check of research data, there are incentives for researchers to manipulate their results in order to increase the possibility of publication of their research. Diekmann (2007) analyzes empirical distributions of digits in statistical estimates, as they can often be found in empirical research. According to his analysis based on more than 1,000 regression coefficients, the first digit of these numbers closely follows Benford's law. The author uses students to fabricate regression coefficients, supporting a certain hypothesis. The results indicate that fabricated data most significantly differ not at the first digit, but rather at the subsequent ones. This is also in line with research by Mosimann, Dahlberg, et al. (2002) and Mosimann, Wiseman, et al. (1995).

Another strand of literature uses Benford's law to check the integrity of reported academic results and for example Tödter (2009) reports a substantial number of articles in his sample, which do not conform to Benford's law and are thus probably manipulated. He concludes that the empirical results presented in these articles might have been manipulated. In a direct response, Diekmann and Jann (2010) challenge the conclusions drawn by Tödter (2009) and doubt that Benford's law is an appropriate tool to distinguish between manipulated and untampered estimates. They especially highlight the relatively high probability of Type I error, i.e., "false positives", when using Benford's law. This property was also analyzed in more detail by Bauer and Groß (2011). They conclude that a comparison of research data with Benford's law can only provide indications for possible fraud, and these indications are relevant if certain conditions are met.

# 4. A Case Study: Benford's Law and the LIBOR

This section provides a case study demonstrating possible ways to use Benford's law for identifying suspected fraud. During the climax of the financial crisis, in April and May 2008, first reports surfaced in the Wall Street Journal (WSJ), indicating possible problems with the reliability of the London Interbank Offered Rate (LIBOR) (Mollenkamp 2008; Mollenkamp and Whitehouse 2008).

Since the LIBOR is supposed to be a measure of the interest rate at which banks lend to each other, it can be used as an indicator of individual bank's risk, which is expected to rise when banks face higher risks. Thus, during the peak of the financial crisis individual banks had the incentive to report lower interest rates to mask their real borrowing costs and, consequently, to obscure the extent of their risk exposure towards the market. A further incentive for fraud, which a priori does not predetermine the direction of possible misreported interest rates, is the possibility that market participants with a large exposure to derivatives rated based on the LIBOR can profit by strategically moving the LIBOR into the desired direction.

Gyntelberg and Wooldridge (2008) raise the issue of reliability of interest rate fixings in connection with the drying up of liquidity in major interbank markets in the second half of 2007. By comparing spreads and correlations between interbank fixings, they are able to show that in the period between August 2007 and January 2008 the spreads widened considerably while LIBOR rose substantially less than similar interest rates fixes.

Monticini and Thornton (2013) use this as a starting point to check the effect of misreporting on the LIBOR rate and find significant breaks in the spread for the period from late 2007 to late 2008 and early 2009. Our case study focuses on the one month LIBOR for the US-$. For this series, we find a break as indication for suspected under-reporting as shown in Figure 5, following the initial publication of the report in the WSJ on April 16, 2008 (date marked by the vertical line in Figure 5) that immediately increased the reported rate.

Abrantes-Metz, Villas-Boas, and Judge (2011)[9] use digital analysis based on Benford's law to analyze the submitted interest rates and try to discern possible patterns indicating the manipulation of the reported interest rates. Given the overall level of interest rates in the period under consideration, a manipulation of the first digit would be too obvious to remain undetected.

[9] The working paper version Abrantes-Metz and Villas-Boas (2010)provides additional technical details on the analysis.

Therefore, the authors focus their analysis on the second digit. They show that, in comparison with the years 1987 to 2005, in the run up and during the financial crisis from 2007 to 2008, the deviations between the empirical distribution of the second digits and the expected distribution according to Benford's law increased.

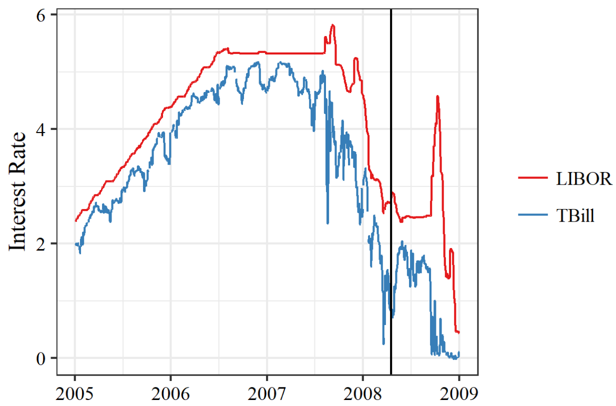



Figure 5: Historical development of the one-month LIBOR and Treasury bill interest rate from 1/4/2005 to 12/31/2008 in percentage terms. The black vertical line indicates the initial release of the Wall Street Journal report on possible manipulations of the LIBOR submissions. It is found that the LIBOR interest rate follows the Treasury bill interest rate closely only until the beginning of the financial crisis. In September 2008, the spread becomes substantially larger than usual.

Obviously, as a test case to check whether Benford's law might be useful in detecting possible manipulations or fraud during this episode provides an interesting example. Ex post, it became apparent that the interest rates submitted by individual banks were often rigged in order to profit in the derivative markets or in order to understate their liquidity risk during the financial crisis (Ashton and Christophers 2015).

However, it has to be taken into account that interest rates cannot be expected to follow Benford's law closely as they do not fulfill most of the conditions discussed in Section 2. Thus, the following analysis does not mainly focus on the question if the LIBOR submissions follow Benford's law, but rather on the question whether Benford's law could act as a sensible benchmark against which the distribution of the digits can be checked. Therefore, the first step of our analysis shown in Figure 6 is just an illustration that the second significant digits of the interest rates under consideration are not distributed according to Benford's law. Nevertheless, the distribution of the second digits for the four weeks Treasury Bills (TBills) interest rates is much more similar to Benford's distribution than the LIBOR submissions during the period from 1/4/2005 to 12/31/2008. The analysis for the LIBOR is based on the submissions of interest rates for US Dollar lending with a maturity of one month from individual banks, which were published by Rogers in The Guardian. In order to provide a benchmark interest rate, the TBills secondary market interest rate[10] is used, which is supposedly less prone to human intervention in the determination of single digits. In fact, the high proportion of the digit 3 for the LIBOR submissions is striking, while the TBills interest rate rather exhibits too high frequencies of larger values (5 to 9) compared to the distribution of the second digit according to

[10]Data on the TBills interest rate provided by the Board of Governors of the Federal Reserve System (2017) through their FRED economic data tool.

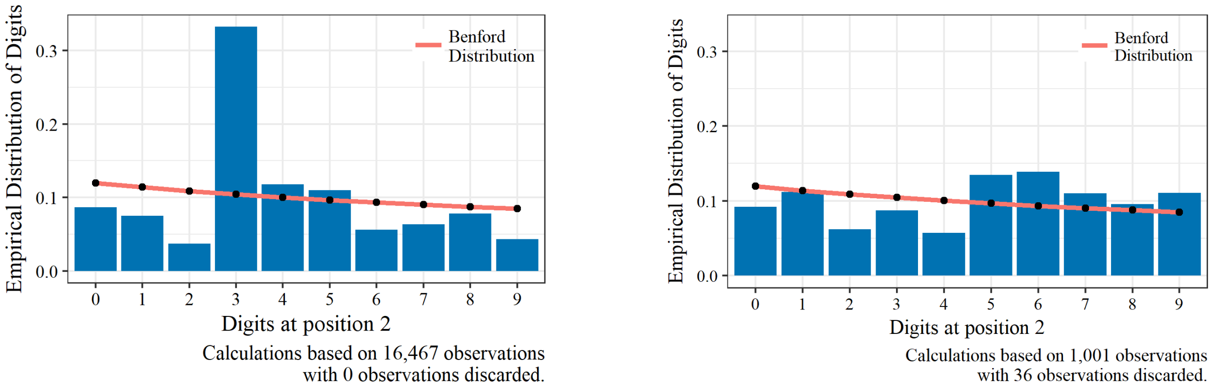



Figure 6: Comparison of the distribution of second digits in the LIBOR submissions (left) and TBills interest rate (right) from 1/4/2005 to 12/31/2008 with the expected theoretical distribution based on Benford's law (red line). The high share of digit 3 in the LIBOR submissions signals a clear departure from Benford's law, while the distribution of second digits in the TBills interest rates is much closer to the theoretical distribution according to Benford's law.

Benford's law. The high frequency of digit 3 is caused by a long period from mid-2006 to mid-2007, when the LIBOR was at about 5.3% and – as will be shown below – all submissions from individual banks were almost identical. Per se, this does not provide any proof for manipulation or fraud, but could have provided a first warning signal even before the first articles in the WSJ.

Figure 7 exhibits the distribution of the second significant digits for a more recent period starting with the newly created LIBOR submission regime and covering the period 2/3/2014 – 2/8/2017.[11] One of the major changes to increase the transparency is that all individual submissions are made available after a three-month delay by the Intercontinental Exchange (ICE) London.[12] Still, the distribution does not conform to Benford's law, but deviations are much smaller than for the earlier period, and the overall shape becomes more similar to the one for the TBills interest rate.

The comparison of LIBOR submissions in each of the two sub periods as depicted in Figure 6, and Figure 7 shows that the transparency measures introduced after the LIBOR scandal appear to have led to a digit distribution more similar to Benford's distribution.

The formal tests based on the $\chi^2$ statistics confirm the descriptive evidence that the empirical distributions would deviate significantly from the distribution if Benford's law applied. Table 2 reports the test statistics for both sub periods and the three pairs of distributions (TBills, LIBOR and theoretical distribution according to Benford's law), which all exceed the critical value at the 1%-level of 23.209. However, the size of the test statistic differs substantially between LIBOR and TBills when it comes to the comparison with Benford's distribution. For the first sub period, the value for LIBOR is larger than 10,000, providing a very strong signal for deviation from Benford's law. This value shrinks to 867 for the second sub period reflecting the descriptive evidence that the distribution became much closer to Benford's distribution. However, for both sub periods, the distance between the empirical distribution of second significant digits for the TBills and Benford's distribution is much smaller.

[11] During this period, the Treasury Bills interest rate became smaller than 0.1% for a substantial number of days including a few days when it actually reached zero. As only two decimal places are reported for the Treasury Bills rate, no second significant digit is available for these days. Therefore, the sample had to be restricted to those days with rates equal to or above 0.1% (303 out of 756 days).

[12] For more information on the newly created regime, please see `https://www.theice.com/iba/libor` and for detailed data please see `https://www.theice.com/marketdata/reports/186`.

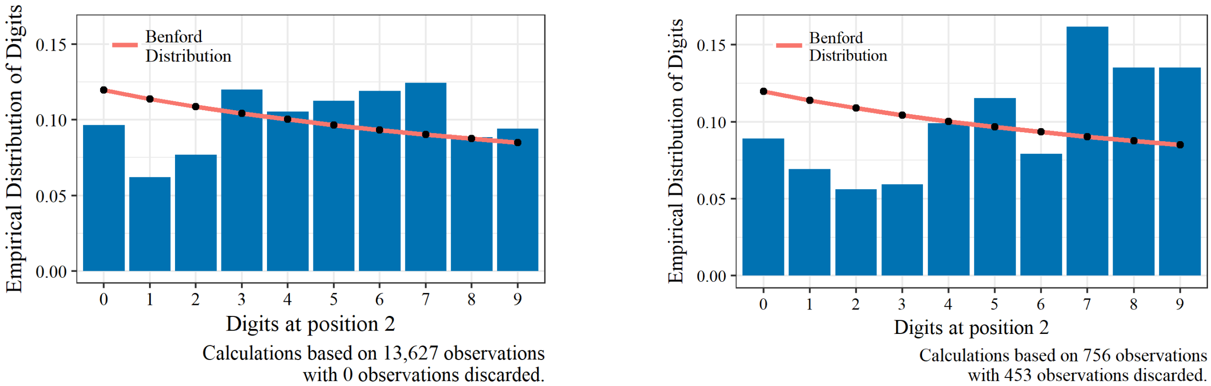



Figure 7: Comparison of the distribution of second digits in the LIBOR submissions (left) and Treasury Bills interest rate (right) from 2/3/2014 to 2/8/2017 with the expected theoretical distribution based on Benford's law (red line). The differences for the LIBOR rate become much smaller as compared to the earlier period shown in Figure 6.

| Distributions compared | $\chi^2$-Test Statistic | MAD – Test Statistic |
|---|---|---|
| 2nd Digits LIBOR vs. Benford 2005-2008 | 10,185 | 0.05184 |
| 2nd Digits TBills vs. Benford 2005-2008 | 94 | 0.02733 |
| 2nd Digits TBills vs. LIBOR 2005-2008 | 463 | 0.06134 |
| 2nd Digits LIBOR vs. Benford 2014-2017 | 867 | 0.02135 |
| 2nd Digits TBills vs. Benford 2014-2017 | 57 | 0.03765 |
| 2nd Digits TBills vs. LIBOR 2014-2017 | 31 | 0.02707 |

Table 2: $\chi^2$ test statistic for different periods and different underlying interest rate time series testing the goodness of fit between the distributions of second digits. Based on the $\chi^2$ test statistic, the null hypothesis of equal distributions has to be rejected for all cases. Moreover, the MAD test statistic in all cases is above the threshold of 0.012 indication non conformity between the pairs of distributions. The size of the $\chi^2$ test statistic shrinks substantially for the second sample period, i.e., the distributions of second digits become more similar.

It is interesting to note that the $\chi^2$-test cannot only be used to compare any empirical distribution with the one according to Benford's law, but also two empirical distributions like the TBills with the LIBOR. The large test statistics for this comparison, in particular for the first sub period, provide further signals regarding a potential problem with the LIBOR data, as the null hypothesis that both distributions of second significant digits follow the same probability distribution for the period 2005 to 2008 has to be rejected at all conventional levels of significance. Therefore, either one has a good rationale why second significant digits should be distributed differently for two interest rates with otherwise similar properties or it has to be interpreted (with caution) as an indication of potential manipulations in at least one of the series. Apparently, the changes to the LIBOR submission regime had the desired impact and made the frequency distributions of the two interest rates much more similar, visible in the decrease of the $\chi^2$-test statistic from 463 to 31.

The values of the MAD test statistic in Table 2 confirm the test results obtained by the $\chi^2$ test, since for the second digit they are all well above the upper threshold provided in Table 3 in the appendix. They thus indicate that the digits of both the LIBOR and the TBills do not conform to Benford's law. Nevertheless, they also indicate that the conformity of the LIBOR time series increased, since the value for the second period is considerably smaller.

Additionally, the deviations between the LIBOR and TBills decreased from 0.06134 to 0.02707 which also indicates that the newly created regulatory regime might have improved the quality of the LIBOR as a proxy for short term interest rates.

In the first part of the analysis, it became apparent that the interest rates under consideration do not conform to Benford's law no matter whether they were based on real transactions such as the TBills or on submissions by individual banks. Nevertheless, since Benford's law offers an easily applicable benchmark it is possible to monitor the extent of these deviations from the benchmark over time, which might be used to detect suspicious changes.

In Figure 8, the change of these deviations is plotted against time. Each observation corresponds to one quarter of data. It becomes obvious that the deviations are largest over the period when the LIBOR rate was almost constant. In order to check whether this substantial increase of the deviations is triggered by particular banks, we repeat the analysis for the individual submissions of each bank separately. The results are provided in Figure 9. Apparently, for the last quarter of 2006 and the first two quarters of 2007 all banks exhibit the same level of deviation, as they reported the same interest rate of 5.32% for a substantial part of this period.

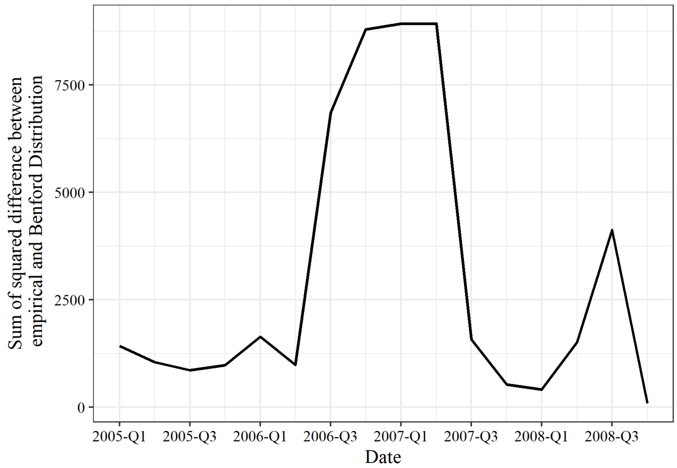



Figure 8: Quarterly development of the sum of squared differences between the empirical and Benford's distribution of the second digit of the LIBOR interest rate submissions from January 2005 to December 2008. Interestingly, there is an unusually large deviation beginning in 2006 until the end of 2007.

Comparing the quarterly development of the squared deviations between the empirical digit distribution and Benford's distribution in Figure 8 and Figure 10, it is important to note the overall decrease of deviations. The mean and median for these two periods decreased from 3,042 to 386 and 1,468 to 256, respectively. This can be regarded as an indication that the newly created regime to increase transparency – and perhaps also the information that academics used Benford's law to analyze historical submissions – led to a more similar distribution of digits and fewer deviations from Benford's law.

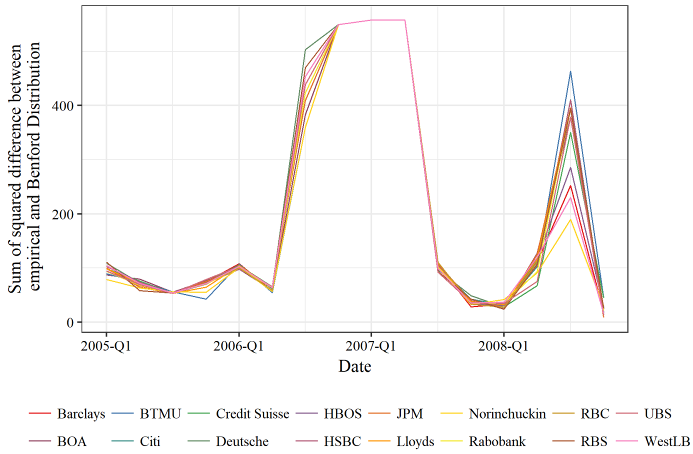



Figure 9: Quarterly development of the sum of squared differences between the empirical and Benford's distribution of the second digit of the LIBOR interest rate submissions per individual bank from January 2005 to December 2008. It becomes apparent that the large deviations from 2006 to the end of 2007 are not caused by one single bank, but are rather the result of a period when all banks submitted the exact same interest rate.

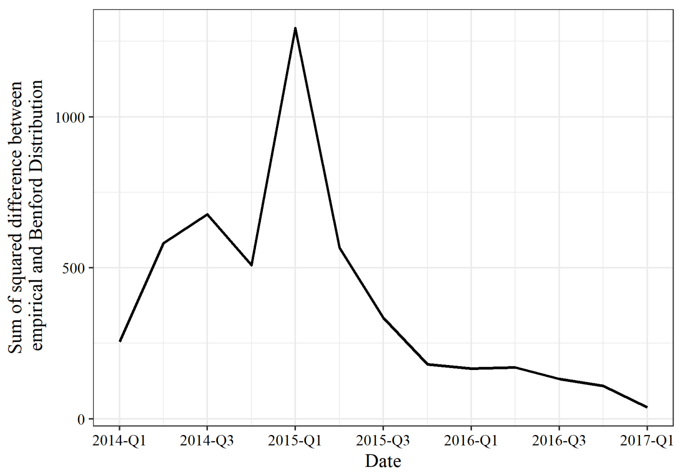



Figure 10: Quarterly development of the sum of squared differences between the empirical and Benford's distribution calculated for the second digits of the LIBOR interest rate submissions from February 2014 to February 2017. Although there are still substantial deviations it has to be noted that the absolute level of the sum of squared differences is much lower than during the period from 2005 to 2008.

# 5. Policy Implications

In order to make productive use of this type of digital analysis in identifying fraud and manipulations, they should be accompanied by further considerations. First, any misconduct – whether within companies or on financial markets – is typically triggered by an incentive to misrepresent the data at hand. It is therefore important to keep the incentive structure of the decision environment in mind when conducting an analysis based on Benford's law or similar statistical techniques. Particularly for decisions taken on upper management levels, a consideration of management compensation schemes will be important. In this respect, Morrison and Thanassoulis (2017) show that the fraud-related impact of compensation systems in firms is strongly dependent on the ethical standards of their employees so that strong incentives via bonuses are not necessarily harmful.[13] Second, and equally important, the increasing complexity of the decision environment will raise the ability of market participants to coordinate their behavior in such a way that the manipulation of data can be concealed efficiently. Examining the conformity with a given data set using Benford's law in combination with an assessment of the incentive and complexity structure of the corresponding decision should help to enhance the validity and informative power of the analysis' output.

Furthermore, the higher the interests at stake, the more likely it is that manipulators are aware of simple methods such as digital analysis based on Benford's law. From a predator-and-prey perspective, it would be naïve to rely on a routine application of such methods as manipulators could avoid detection at low cost. Instead, continuous effort is required to develop new statistical tools for detecting fraud and manipulation. While any new method, unknown to the agents in the field of application, might work well for some time, more sustainable effects can be achieved if the methods do not allow for an easy adjustment of manipulated data even if they were known. The generalizations of digital analysis based on Benford's law discussed in Section 2, which use empirical distributions as a benchmark, represent such an approach.

Finally, from a regulation point of view, it appears rather obvious that benchmarks with a financial impact such as the LIBOR (similar examples include inflation rates used in indexed contracts, GDP determining contributions to international organizations, credit ratings) should not be determined by agents immediately affected.

# 6. Outlook

Benford's law can be a valuable method to detect fraud, for instance with respect to the LIBOR manipulations. Therefore, regulators such as the United States Securities and Exchange Commission (SEC) or the United States Federal Reserve System (Fed) as well as other organizations with similar responsibilities should make use of tests relying on methods like Benford's law to detect fraudulent behavior proactively. A stronger focus on statistical methods for fraud detection will also increase the number of lawsuits against companies and employees, which may have a deterrent effect on other potential offenders. Moreover, courts should encourage companies to use statistical methods for identifying first suspicions of a fraud case. However, there is always a risk of errors either due to the properties of a statistical test or due to the fact that not all data considered necessarily follow Benford's law (see also Cleary and Thibodeau

[13] Irrespective of compensation structures,there is a positive relation of certain individual personality traits, such as narcissism, with fraud (Rijsenbilt and Commandeur 2013). However, effects of personality traits of individual decision makers will be much more difficult to assess as compared to a general incentive structure invoked by a compensation scheme. For a discussion of incentives and personal characteristics in the context of interviewer falsifications see, e.g., Winker et al. (2015) and Winker (2016).

(2005)). Therefore, not all significant deviations from the Benford distribution should be taken as conclusive evidence of manipulation. While fabrication of numbers often does result in a deviation (unless the fabricators are aware of checks based on Benford's law – in this case, one would not expect any deviation), even simple rounding can have the same effect. Consequently, non-conformity with Benford's law can be seen as a signal of potential misconduct that might warrant further investigation. As such, Benford's law has become a well-known pre-test in many different fields of manipulation detection. However, for sentencing companies or individuals, courts need a complete chain of evidence based on indisputable facts.

Focusing on fraud prevention in the future, digitalization of the accounting system and new technologies like blockchain technology might make data manipulations more difficult. In detail, blockchain technology can help to implement an automated accounting process especially across companies based on a joint register. In this joint register, all transactions are recorded and cryptographically sealed. Thus, the life cycle of each accounting incident is fully reflected in the blockchain and all relevant documents are stored in the blockchain making data manipulation practically impossible. However, even though the blockchain technology has a lot of potential, to date it is still in the experimental phase. Furthermore, it is recommendable to monitor the development of this technology with a focus on gateways for misconduct. In any case, until blockchain (or an alternative) technology is widely used, detecting fraud remains a crucial issue, and thus, statistical methods are of utmost importance.

Given the limitations of rather simple methods of digital analysis, such as Benford's law, further developments in this field are required. More advanced methods, including those mentioned as generalizations of Benford's law, require more data, e.g., several variables from financial statements. Then, both classical and more recent multivariate classification tools might be used, which exhibit the advantage of making it more difficult for manipulators to adjust their data in a way to remain undetected. Some of these tools require a supervised learning step, and it might be difficult to obtain appropriate data covering already identified manipulated data as well as correct data (Ravisankar et al. 2011). Consequently, the use of other – non-supervised – machine learning tools might be an interesting approach for future research.

# References

Abrantes-Metz, R. M. and S. B. Villas-Boas (2010): "Tracking the Libor Rate." In: *Available at SSRN: https://ssrn.com/abstract=1646600*. DOI: 10.2139/ssrn.1646600.

Abrantes-Metz, R. M., S. B. Villas-Boas, and G. Judge (2011): "Tracking the Libor rate." In: *Applied Economics Letters* 18 (10), pp. 893–899. DOI: 10.1080/13504851.2010.515197.

Amiram, D., Z. Bozanic, and E. Rouen (2015): "Financial statement errors: Evidence from the distributional properties of financial statement numbers." In: *Review of Accounting Studies* 20 (4), pp. 1540–1593. DOI: 10.1007/s11142-015-9333-z.

Ashton, P. and B. Christophers (2015): "On arbitration, arbitrage and arbitrariness in financial markets and their governance: Unpacking LIBOR and the LIBOR scandal." In: *Economy and Society* 44 (2), pp. 188–217. DOI: 10.1080/03085147.2015.1013352.

Ausloos, M., R. Castellano, and R. Cerqueti (2016): "Regularities and discrepancies of credit default swaps: A data science approach through Benford's law." In: *Chaos, Solitons & Fractals* 90, pp. 8–17. DOI: 10.1016/j.chaos.2016.03.002.

Barabesi, L., A. Cerasa, A. Cerioli, and D. Perrotta (2017): "Goodness-of-Fit Testing for the Newcomb-Benford Law With Application to the Detection of Customs Fraud." In: *Journal of Business & Economic Statistics* 78, pp. 1–13. DOI: 10.1080/07350015.2016.1172014.

Bauer, J. and J. Groß (2011): "Difficulties Detecting Fraud? The Use of Benford's Law on Regression Tables." In: *Journal of Economics and Statistics* 231 (5-6), pp. 734–748. DOI: 10.1515/jbnst-2011-5-611.

Benford, F. (1938): "The Law of Anomalous Numbers." In: *Proceedings of the American Philosophical Society* 78 (4), pp. 551–572.

Berger, A., T. P. Hill, and E. Rogers (2017): *Benford Online Bibliography*.

Bierstaker, J. L., R. G. Brody, and C. Pacini (2006): "Accountants' perceptions regarding fraud detection and prevention methods." In: *Managerial Auditing Journal* 21 (5), pp. 520–535. DOI: 10.1108/02686900610667283.

Board of Governors of the Federal Reserve System (2017): *4-Week Treasury Bill: Secondary Market Rate [DTB4WK]*. Ed. by FRED.

Boyle, J. (1994): "An application of Fourier series to the most significant digit problem." In: *The American Mathematical Monthly* 101 (9), pp. 879–886.

Bredl, S., N. Storfinger, and N. Menold (2013): "A Literature Review of Methods to Detect Fabricated Survey Data." In: *Interviewers' deviations in surveys*. Ed. by N. Menold, P. Winker, and R. Porst. Schriften zur empirischen Wirtschaftsforschung. Frankfurt am Main: Peter Lang, pp. 3–24.

Bredl, S., P. Winker, and K. Kötschau (2012): "A statistical approach to detect interviewer falsification of survey data." In: *Survey methodology* 38 (1), pp. 1–10.

Carslaw, C. A. P. N. (1988): "Anomalies in Income Numbers: Evidence of Goal Oriented Behavior." In: *The Accounting Review* 63 (2), pp. 321–327.

Chenavier, N., B. Massé, and D. Schneider (2018): "Products of random variables and the first digit phenomenon." In: *Stochastic Processes and their Applications* 128 (5), pp. 1615–1634. DOI: 10.1016/j.spa.2017.08.003.

Christian, C. W., S. Gupta, and S. Lin (1993): "Determinants of tax preparer usage: Evidence from panel data." In: *National Tax Journal* 46 (4), pp. 487–503.

Cleary, R. and J. C. Thibodeau (2005): "Applying Digital Analysis Using Benford's Law to Detect Fraud: The Dangers of Type I Errors." In: *Auditing* 24 (1), pp. 77–81. DOI: 10.2308/aud.2005.24.1.77.

Corazza, M., A. Ellero, and A. Zorzi (2010): "Checking financial markets via Benford's law: The S&P 500 case." In: *Mathematical and Statistical Methods for Actuarial Sciences and Finance*. Ed. by M. Corazza and C. Pizzi. Milano: Springer Milan, pp. 93–102. DOI: 10.1007/978-88-470-1481-7{\textunderscore}10.

Deleanu, I. S. (2017): "Do Countries Consistently Engage in Misinforming the International Community about Their Efforts to Combat Money Laundering? Evidence Using Benford's Law." In: *PLOS ONE* 12 (1), pp. 1–19. DOI: 10.1371/journal.pone.0169632.

Diaconis, P. and D. Freedman (1979): "On Rounding Percentages." In: *Journal of the American Statistical Association* 74 (366), pp. 359–364. DOI: 10.2307/2286335.

Diekmann, A. (2007): "Not the First Digit! Using Benford's Law to Detect Fraudulent Scientific Data." In: *Journal of Applied Statistics* 34 (3), pp. 321–329. DOI: 10.1080/02664760601004940.

Diekmann, A. and B. Jann (2010): "Benford's Law and Fraud Detection: Facts and Legends." In: *German Economic Review* 11 (3), pp. 397–401. DOI: 10.1111/j.1468-0475.2010.00510.x.

Drake, P. D. and M. J. Nigrini (2000): "Computer assisted analytical procedures using Benford's Law." In: *Journal of Accounting Education* 18 (2), pp. 127–146. DOI: 10.1016/S0748-5751(00)00008-7.

Dümbgen, L. and C. Leuenberger (2008): "Explicit bounds for the approximation error in Benford's law." In: *Electronic Communications in Probability* 13, pp. 99–112.

Durtschi, C., W. Hillison, and C. Pacini (2004): "The effective use of Benford's law to assist in detecting fraud in accounting data." In: *Journal of forensic accounting* 5 (1), pp. 17–34.

El Mouaaouy, F. (2018): "Financial crime 'hot spots' – empirical evidence from the foreign exchange market." In: *The European Journal of Finance* 24 (7-8), pp. 565–583. DOI: 10.1080/1351847X.2017.1303528.

El Mouaaouy, F. and J. Riepe (2018): "Benford and the Internal Capital Market: A Useful Indicator of Managerial Engagement." In: *German Economic Review* 19 (3), pp. 309–329. DOI: 10.1111/geer.12129.

Fewster, R. M. (2009): "A Simple Explanation of Benford's Law." In: *The American Statistician* 63 (1), pp. 26–32.

Gauvrit, N. and J.-P. Delahaye (2009): "Scatter and regularity imply Benford's law... and more." In: *arXiv preprint arXiv:0910.1359*.

Gonzalez-Garcia, J. and G. C. Pastor (2009): "Benford's Law and Macroeconomic Data Quality." In: *IMF Working Paper* WP/09/10.

Goodman, W. (2016): "The promises and pitfalls of Benford's law." In: *Significance* 13 (3), pp. 38–41. DOI: 10.1111/j.1740-9713.2016.00919.x.

Gyntelberg, J. and P. Wooldridge (2008): "Interbank rate fixings during the recent turmoil." In: *BIS Quarterly Review, March 2008*.

Henselmann, K., D. Ditter, and E. Scherr (2015): "Irregularities in Accounting Numbers and Earnings Management—A Novel Approach Based on SEC XBRL Filings." In: *Journal of Emerging Technologies in Accounting* 12 (1), pp. 117–151. DOI: 10.2308/jeta-51247.

Hill, T. P. (1988): "Random-Number Guessing and the First Digit Phenomenon." In: *Psychological Reports* 62 (3), pp. 967–971. DOI: 10.2466/pr0.1988.62.3.967.

– (1995): "A Statistical Derivation of the Significant-Digit Law." In: *Statistical science* 10 (4), pp. 354–363.

Hofmarcher, P. and K. Hornik (2013): "First Significant Digits and the Credit Derivative Market During the Financial Crisis." In: *Contemporary Economics* 7 (2), pp. 21–29.

Holz, C. A. (2014): "The quality of China's GDP statistics." In: *China Economic Review* 30, pp. 309–338. DOI: 10.1016/j.chieco.2014.06.009.
Hürlimann, W. (2006): "Benford's law from 1881 to 2006." In: *ArXiv Mathematics e-prints* arXiv:math/0607168.
– (2009): "Generalizing Benford's Law Using Power Laws: Application to Integer Sequences." In: *International Journal of Mathematics and Mathematical Sciences* 2009 (4), pp. 1–10. DOI: 10.1155/2009/970284.
Judge, G. and L. Schechter (2009): "Detecting Problems in Survey Data Using Benford's Law." In: *Journal of Human Resources* 44 (1), pp. 1–24. DOI: 10.3368/jhr.44.1.1.
Koedijk, K. G. and P. A. Stork (1994): "Should we care? psychological barriers in stock markets." In: *Economics Letters* 44 (4), pp. 427–432. DOI: 10.1016/0165-1765(94)90116-3.
Ley, E. (1996): "On the Peculiar Distribution of the US Stock Indexes' Digits." In: *The American Statistician* 50 (4), pp. 311–313. DOI: 10.2307/2684926.
Ley, E. and H. R. Varian (1994): "Are there psychological barriers in the Dow-Jones index?" In: *Applied Financial Economics* 4 (3), pp. 217–224. DOI: 10.1080/758526902.
Lin, F., L.-J. Lin, C.-C. Yeh, and T.-S. Wang (2018): "Does the board of directors as Fat Cats exert more earnings management? Evidence from Benford's law." In: *The Quarterly Review of Economics and Finance* 68, pp. 158–170. DOI: 10.1016/j.qref.2017.09.005.
Lu, F. and J. E. Boritz (2005): "Detecting Fraud in Health Insurance Data: Learning to Model Incomplete Benford's Law Distributions." In: *Machine Learning: ECML 2005: 16th European Conference on Machine Learning, Porto, Portugal, October 3-7, 2005. Proceedings*. Ed. by J. Gama, R. Camacho, P. B. Brazdil, J. Alipio, and L. Torgo. Springer Berlin Heidelberg.
Michalski, T. and G. Stoltz (2013): "Do Countries Falsify Economic Data Strategically? Some Evidence That They Might." In: *Review of Economics and Statistics* 95 (2), pp. 591–616. DOI: 10.1162/REST{\textunderscore}a{\textunderscore}00274.
Mir, T. A. and M. Ausloos (2018): "Benford's law: A "sleeping beauty'' sleeping in the dirty pages of logarithmic tables." In: *Journal of the Association for Information Science and Technology* 69 (3), pp. 349–358. DOI: 10.1002/asi.23845.
Mollenkamp, C. (2008): "Bankers cast doubt on key rate amid crisis." In: *Wall Street Journal, 04/16/2008* 16.
Mollenkamp, C. and M. Whitehouse (2008): "Study Casts Doubt on Key Rate: Wall Street Journal (WSJ) Analysis Suggests Banks May Have Reported Flawed Interest Data for Libor." In: *Wall Street Journal, 05/29/2008*.
Monticini, A. and D. L. Thornton (2013): "The effect of underreporting on LIBOR rates." In: *Journal of Macroeconomics* 37, pp. 345–348. DOI: 10.1016/j.jmacro.2013.02.002.
Morrison, A. D. and J. E. Thanassoulis (2017): "Ethical standards and cultural assimilation in financial services." In: *Available at SSRN: https://ssrn.com/abstract=2766996*.
Mosimann, J. E., J. Dahlberg, N. Davidian, and J. Krueger (2002): "Terminal Digits and the Examination of Questioned Data." In: *Accountability in Research* 9 (2), pp. 75–92. DOI: 10.1080/08989620212969.
Mosimann, J. E., C. V. Wiseman, and R. E. Edelman (1995): "Data fabrication: Can people generate random digits?" In: *Accountability in Research* 4 (1), pp. 31–55. DOI: 10.1080/08989629508573866.
Newcomb, S. (1881): "Note on the Frequency of Use of the Different Digits in Natural Numbers." In: *American Journal of Mathematics* 4 (1), pp. 39–40. DOI: 10.2307/2369148.
Nigrini, M. (1994): "Using digital frequencies to detect fraud." In: *The white paper* 8 (2), pp. 3–6.

Nigrini, M. J., ed. (2012): *Benford's Law*. Hoboken, NJ, USA: John Wiley & Sons, Inc. DOI: 10.1002/9781119203094.

Nigrini, M. J. (1996): "A Taxpayer Compliance Application of Benford's Law." In: *Journal of the American Taxation Association* 18 (1), pp. 72–91.

Nigrini, M. J. and S. J. Miller (2009): "Data Diagnostics Using Second-Order Tests of Benford's Law." In: *Auditing* 28 (2), pp. 305–324.

Nigrini, M. J. and L. J. Mittermaier (1997): "The use of Benford's law as an aid in analytical procedures." In: *Auditing* 16 (2), pp. 52–67.

Nye, J. and C. Moul (2007): "The Political Economy of Numbers: On the Application of Benford's Law to International Macroeconomic Statistics." In: *The B.E. Journal of Macroeconomics* 7 (1), pp. 1–14. DOI: 10.2202/1935-1690.1449.

OpenStreetMap (2018): *Planet dump retrieved from https://planet.osm.org*.

Pinkham, R. S. (1961): "On the distribution of first significant digits." In: *The Annals of Mathematical Statistics* 32 (4), pp. 1223–1230.

Raimi, R. A. (1976): "The First Digit Problem." In: *The American Mathematical Monthly* 83 (7), pp. 521–538. DOI: 10.2307/2319349.

Rauch, B., M. Göttsche, G. Brähler, and S. Engel (2011): "Fact and Fiction in EU-Governmental Economic Data." In: *German Economic Review* 12 (3), pp. 243–255. DOI: 10.1111/j.1468-0475.2011.00542.x.

Rauch, B., M. Göttsche, F. El Mouaaouy, and F. Geidel (2013): "Empirical Methods in Competition Analysis Applying Benford's Law to the Western Australian Petroleum Market." In: *Available at SSRN: https://ssrn.com/abstract=2364384*. DOI: 10.2139/ssrn.2364384.

Ravisankar, P., V. Ravi, G. Raghava Rao, and I. Bose (2011): "Detection of financial statement fraud and feature selection using data mining techniques." In: *Decision Support Systems* 50 (2), pp. 491–500. DOI: 10.1016/j.dss.2010.11.006.

Rijsenbilt, A. and H. Commandeur (2013): "Narcissus Enters the Courtroom: Chief Executive Officer (CEO) Narcissism and Fraud." In: *Journal of Business Ethics* 117 (2), pp. 413–429. DOI: 10.1007/s10551-012-1528-7.

Rodriguez, R. J. (2004): "First Significant Digit Patterns from Mixtures of Uniform Distributions." In: *The American Statistician* 58 (1), pp. 64–71.

Schräpler, J. (2011): "Benford's Law as an Instrument for Fraud Detection in Surveys Using the Data of the Socio-Economic Panel (SOEP)." In: *Journal of Economics and Statistics* 231 (5-6), pp. 685–718. DOI: 10.1515/jbnst-2011-5-609.

Schündeln, M. (2018): "Multiple Visits and Data Quality in Household Surveys." In: *Oxford Bulletin of Economics and Statistics* 80 (2), pp. 380–405. DOI: 10.1111/obes.12196.

Shi, J., M. Ausloos, and T. Zhu (2018): "Benford's law first significant digit and distribution distances for testing the reliability of financial reports in developing countries." In: *Physica A: Statistical Mechanics and its Applications* 492, pp. 878–888. DOI: 10.1016/j.physa.2017.11.017.

Smith, S. W. (1997): *The scientist and engineer's guide to digital signal processing*. 1. ed. San Diego, Calif.: California Technical Publ.

Stenfors, A. (2018): "Bid-ask spread determination in the FX swap market: Competition, collusion or a convention?" In: *Journal of International Financial Markets, Institutions and Money* 54, pp. 78–97. DOI: 10.1016/j.intfin.2017.08.001.

Thomas, J. K. (1989): "Unusual Patterns in Reported Earnings." In: *The Accounting Review* 64 (4), pp. 773–787.

Tödter, K.-H. (2009): "Benford's Law as an Indicator of Fraud in Economics." In: *German Economic Review* 10 (3), pp. 339–351. DOI: 10.1111/j.1468-0475.2009.00475.x.

Ullmann, R. and C. Watrin (2017): "Detecting Target-Driven Earnings Management Based on the Distribution of Digits." In: *Journal of Business Finance & Accounting* 44 (1-2), pp. 63–93. DOI: 10.1111/jbfa.12223.

Wallace, W. A. (2002): "Assessing the quality of data used for benchmarking and decision-making." In: *The Journal of Government Financial Management* 51 (3), pp. 18–22.

Watrin, C., R. Struffert, and R. Ullmann (2008): "Benford's Law: An instrument for selecting tax audit targets?" In: *Review of Managerial Science* 2 (3), pp. 219–237. DOI: 10.1007/s11846-008-0019-9.

Wheatley, M. (2012): *The Wheatley Review of LIBOR: Final report*. London: HM Treasury.

Winker, P. (2016): "Assuring the quality of survey data: Incentives, detection and documentation of deviant behavior." In: *Statistical Journal of the IAOS* 32 (3), pp. 295–303. DOI: 10.3233/SJI-161012.

Winker, P., K.-W. Kruse, N. Menold, and U. Landrock (2015): "Interviewer effects in real and falsified interviews: Results from a large scale experiment." In: *Statistical Journal of the IAOS* 31 (3), pp. 423–434. DOI: 10.3233/SJI-150908.

Winter, C., M. Schneider, and Y. Yannikos (2012): "Model-Based Digit Analysis for Fraud Detection Overcomes Limitations of Benford Analysis." In: *2012 Seventh International Conference on Availability, Reliability and Security*, pp. 255–261. DOI: 10.1109/ARES.2012.37.

# Appendices

## A. Appendix

| Digits | Range | Conclusion |
|---|---|---|
| First Digits | 0.000 to 0.006 | Close conformity |
| | 0.006 to 0.012 | Acceptable conformity |
| | 0.012 to 0.015 | Marginally acceptable conformity |
| | Above 0.015 | Nonconformity |
| Second Digits | 0.000 to 0.008 | Close conformity |
| | 0.008 to 0.010 | Acceptable conformity |
| | 0.010 to 0.012 | Marginally acceptable conformity |
| | Above 0.0022 | Nonconformity |
| First-Two Digits | 0.0000 to 0.0012 | Close conformity |
| | 0.0012 to 0.0018 | Acceptable conformity |
| | 0.0018 to 0.0022 | Marginally acceptable conformity |
| | Above 0.0022 | Nonconformity |
| First-Three Digits | 0.00000 to 0.00036 | Close conformity |
| | 0.00036 to 0.00044 | Acceptable conformity |
| | 0.00044 to 0.00050 | Marginally acceptable conformity |
| | Above 0.00050 | Nonconformity |

Table 3: Drake and Nigrini (2000) propose the MAD of digit frequencies as an alternative to the $\chi^2$-statistic. Given that no theoretical results are available on the probability distribution of their statistics, categories corresponding to differing degrees of conformity to Benford's law are provided following their proposal.

# Affidavit

Ich erkläre hiermit, dass ich die vorgelegten und nachfolgend aufgelisteten Aufsätze selbstständig und nur mit den Hilfen angefertigt habe, die im jeweiligen Aufsatz angegeben oder zusätzlich in der nachfolgenden Liste aufgeführt sind. In der Zusammenarbeit mit den angeführten Koautoren war ich mindestens anteilig beteiligt. Bei den von mir durchgeführten und in den Aufsätzen erwähnten Untersuchungen habe ich die Grundsätze guter wissenschaftlicher Praxis, wie sie in der Satzung der Justus-Liebig-Universität Gießen zur Sicherung guter wissenschaftlicher Praxis niedergelegt sind, eingehalten.

Johannes Lips
Gießen, 8th November 2019

**Submitted Papers:**


1. Lips, J. (2017): “Do They Still Matter? – Impact of Fossil Fuels on Electricity Prices in the Light of Increased Renewable Generation”. In: *Journal of Time Series Econometrics* 9 (2), pp. 1–30. DOI: 10.1515/jtse-2016-0018
2. Lips, J. (2019): “Debt and the oil industry: analysis on the firm and production level”. In: *Journal of Energy Markets* 12 (4), pp. 1–29. DOI: 10.21314/JEM.2019.189
3. Funk, C., K. Kempa and J. Lips (2019): “Oil Price Shocks and Cost of Debt – Evidence from Oil Firms”. In: *Working Paper*
4. Bannier, C., C. Ewelt-Knauer, P. Winker and J. Lips (In press): “Benford’s law and its application to detecting financial fraud and manipulation.” In: *Corruption and Fraud in Financial Markets: Malpractice, Misconduct and Manipulation.* Ed. by C. Alexander and D. Cumming. John Wiley & Sons, Ltd.