What Drives Drilling Up and Prices Down?

A Structural Vector Autoregressive Model of US Natural Gas Markets*

Valentin Winkler†

September 4, 2023

Abstract

This study uses a structural vector autoregressive (SVAR) model to examine the relationships between the intensity of drilling activity (i.e. investment) for natural gas production, natural gas withdrawals, economic activity, and natural gas prices in the United States. The results show that the reaction of drilling to an unexpected change in natural gas prices depends on the source of the price change. Specifically, I find that a price change due to economic activity shocks leads to a stronger response in drilling than a price change due to different demand-side factors such as changing preferences or spillover effects from other energy markets. This can explain why natural gas drilling levels in the US were relatively low after the Russian invasion of Ukraine despite the highest natural gas prices since 2008. In addition, it is shown that demand-side factors were more important than supply-side factors in explaining the 85% drop in natural gas prices from June 2008 to April 2012. This contradicts prevailing explanations focused on shale gas development and should dampen the expectations of policymakers seeking to rapidly expand shale gas production in order to obtain similar cheap energy as the U.S. after 2008.

Keywords: Natural Gas, Drilling, Time Series, SVAR Model, Bayesian Analysis

JEL Classification: C11, C32, Q41, Q43, Q48

*I would like to thank the IMK Düsseldorf for the opportunity to work on this project during a summer internship in 2022 and especially Sven Schreiber for his valuable suggestions and advice during this time. I am also grateful to Emanuel Gasteiger as well as the participants of the brown bag seminar of the TU Vienna Economics Research Group and the course Bayesian Econometrics II in spring 2023 at the Vienna University of Business for their feedback and the opportunity to present my work.

†Mail: valentin.winkler@icloud.com

1 Introduction

The Russian invasion of Ukraine in spring 2022 has led to massive uncertainty about the security of energy supply in Europe and to historic jumps in energy prices, especially for natural gas. As a result, the functioning mechanisms of gas markets and in particular the question of what is responsible for fluctuations in the supply and price of natural gas have become the focus of renewed attention. This paper examines the natural gas market in the US using a structural vector autoregressive model with the aim of gaining additional empirical insights into these issues. Looking at the US experience makes sense for two main reasons: First, since the financial crisis of 2008, many regions around the world, including the European Union, have switched from a pricing policy linked to crude oil to an US-style model of gas on gas competition or consider doing so (IGU, 2021; Mu, 2019). In addition, features that are already largely a reality in the US, such as the physical integration of regional gas markets and the availability of pipeline capacity for natural gas that has been traded on the spot market, are being advanced in other parts of the world as well. One can thus look at the natural gas market in the US as a model towards which other markets are currently evolving. Second, the energy situation in the US has changed over the last 15 years in a direction that governments around the world look upon with envy. After decades of stagnation, natural gas production in the US began to rise rapidly between 2005 and 2007. As Figure 1 shows, this was mainly driven by a massive exploitation of unconventional natural gas reservoirs, commonly refered to as shale gas. After the financial crisis of 2008, this expansion of natural gas production was also accompanied by a sharp decrease in prices: Between June 2008 and April 2012 natural gas prices in the US fell by 85 percent.[1] This drop stood in sharp contrast to rising global prices for liquefied natural gas (LNG) due to increased demand after the Fukushima accident and extremely high crude oil prices (see Figure 2 and Mu, 2019). The United States seemed to have found a recipe to escape the global energy price bonanza. Unsurprisingly, this success story immediately attracted the interest of other governments – for example in Argentina, Poland or Australia – who wanted to work on their own shale gas boom (Wang and Krupnick, 2015). Recently it was the energy crisis following the Russian invasion of Ukraine that has revived debates in many countries on whether shale gas development should be pursued. The US experience after the shale gas boom

[1] In June 2008 one MMBtu of natural gas (about 0.29 MWh) cost 12.7 USD at the Henry Hub, in April 2012 this price was down to 1.95 USD. Source: FRED (2022) 'Henry Hub Natural Gas Spot Price, Not Seasonally Adjusted (MHHNGSP)' (monthly): `https://fred.stlouisfed.org/series/MHHNGSP`.

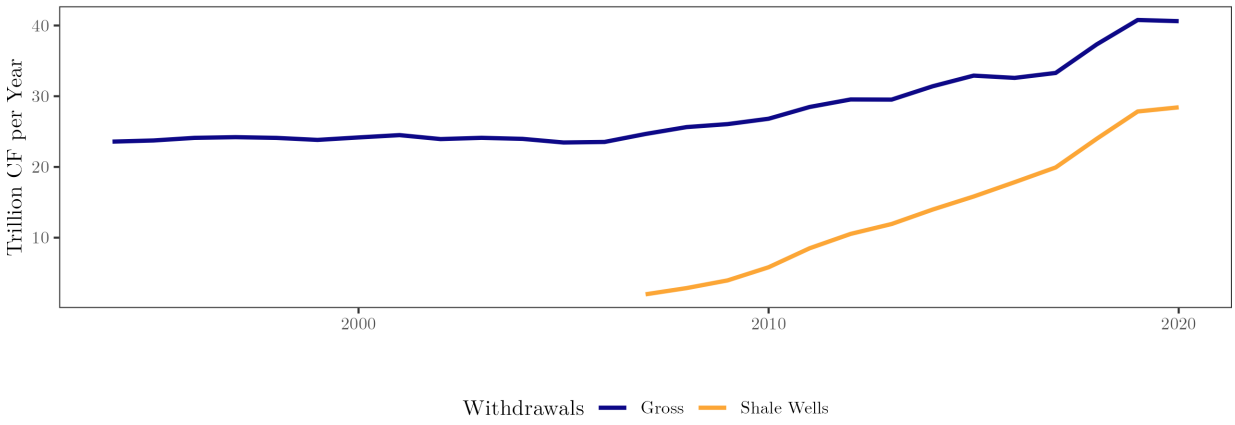


Figure 1: Natural gas withdrawals in the United States and the contribution of shale gas.
Source: Energy Information Administration (2022) 'Natural Gas Gross Withdrawals and Production' (annual): `https://www.eia.gov/dnav/ng/ng_prod_sum_dc_NUS_mmcf_a.htm`.

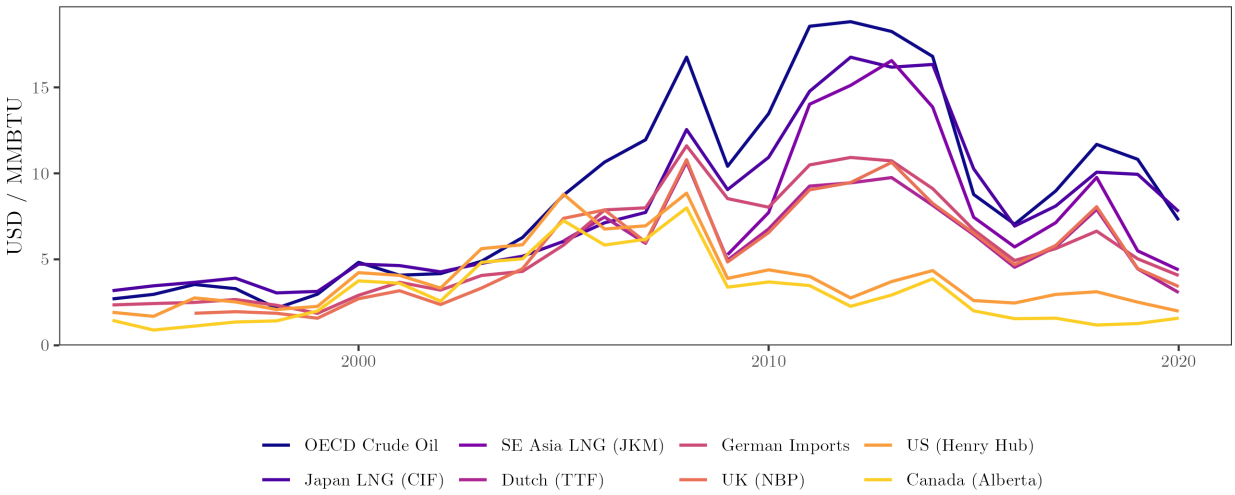


Figure 2: International natural gas prices and price for crude oil.
Source: Statistical Review of World Energy - BP via Our World in Data (2022) 'Natural gas prices' (annual): `https://ourworldindata.org/grapher/natural-gas-prices`.

is already serving as a blueprint for policy makers around the world and thus it is important to understand it. Therefore I not only try to disentangle the general workings of natural gas markets in the US but also study this particular historical episode in some detail.

This paper contains two main contributions to the existing natural gas literature. First, in addition to economic activity, gas production and the gas price I include natural gas drilling in my econometric model. Recent studies of the oil and gas market have shown that production from existing wells depends primarily on geological characteristics rather than the price of the fuel, and producers respond to price fluctuations through the rate of drilling (Anderson et al., 2018; Mason and Roberts, 2018). This makes the rate of drilling the most important decision variable of oil and gas companies and a key determinant of domestic fossil fuel production

in the medium and long run. Nevertheless, it has so far been unusual to include the rate of drilling in macroeconomic models of the natural gas market. By doing this, I find that the reaction of drilling to a change in the natural gas price strongly depends on the source of the price change. In particular, price changes due to strong economic activity produce a stronger reaction of drilling than price changes due to other demand-side factors such as spillovers from other energy markets or changed preferences. This new result could serve as an explanation for why natural gas drilling in the US was relatively low after the Russian invasion of Ukraine, even though prices were the highest they had been since the great recession. Second, I show that demand-side factors contributed more to the 85 percent natural gas price drop between June 2008 and April 2012 than the shale gas boom. Since this fall in prices impressed analysts and governments around the world, studying its causes is of great relevance, as an incorrect assessment of the role that the shale gas boom played for this price drop in the US could lead to exaggerated expectations regarding the benefits of shale gas development elsewhere. This, in turn, carries the risk of accepting too great a financial and environmental cost for this undertaking. My model suggests that policy makers, even if they were as successful as the US in expanding shale gas production, should not expect as sharp a price drop as the US has experienced after 2008.

The remainder of the paper is structured as follows: In Section 2 I will provide an overview of the existing literature, Section 3 sets up the structural vector autoregressive model and explains how identification and inference are conducted. In Section 4 the main results are presented and discussed, Section 5 contains several robustness checks of my key results and Section 6 concludes.

2 Literature Review

This paper builds primarily on research into three aspects of the fossil fuel market: The methodological literature on macroeconometric analysis of oil markets, the application of SVAR models to the natural gas market, and the primarily microeconomic work on drilling. In what follows, I will outline this paper's connection points to these three strands of literature and explain the gaps I am addressing with my work.

Econometrics of oil markets. Much of the research on natural gas markets draws heavily on methods and discussions previously developed with respect to oil markets and my model is no exception. Since Kilian (2009) structural vector autoregressive (SVAR) models are the pre-

dominant tool for analyzing the oil market. Using recursive exclusion restrictions and modelling oil production, economic activity and oil prices, this influential study comes to the conclusion that demand for oil stemming from changes in global economic activity was mainly responsible for rising oil prices in the 2000s, which could explain why this price increase seemed to not curb economic growth. Before, studies like Hamilton (2003) focused mainly on oil price changes due to supply disruptions. Kilian and Murphy (2012) and Kilian and Murphy (2014) develop the basic model of Kilian (2009) further, using sign and elasticity restrictions instead of more restrictive exclusion restrictions. This article's model follows the approach of these two papers to a large extent. Recently Baumeister and Hamilton (2019) used an innovative approach to challenge the predominant role of demand-side shocks for oil price fluctuations provided by the modelling approach of Kilian and Murphy (2014). In addition to allowing a larger price elasticity of oil supply than Kilian and Murphy (2014) they specify a prior distribution for A_0 – the contemporaneous elasticities between endogenous variables – instead of imposing prior restrictions on A_0^{-1} – which are the instantaneous effects of the structural shocks. This is relevant because a supposedly uninformative prior for one of the matrices may imply an unintentionally informative prior for its inverse (Baumeister and Hamilton, 2015; Kilian and Lütkepohl, 2018, p. 453-454). Even though the methodological debate about how to correctly identify structural shocks in the oil market is of interest for analyzing natural gas markets, it is to vast to be summarized here in full. Kilian and Zhou (2020) give an overview of current discussions.

Gas market models. In the last decade there has also evolved a literature on natural gas market SVAR models. One of the first articles is Nick and Thoenes (2014), who study the German natural gas market. Since their specification uses weekly data they are not able to include any variables concerning real economic activity. However, they highlight the importance of oil and coal prices for explaining natural gas prices in the long term. Later papers built on the three variables and the recursive identification strategy of Kilian (2009) to study the US natural gas market and extended the baseline model in some dimension: Arora and Lieskovsky (2014) include residential natural gas consumption as a fourth variable, Hou and Nguyen (2018) set up a Markov switching SVAR model and Rubaszek and Uddin (2020) use a threshold SVAR model to estimate effects of structural shocks depending on natural gas inventory levels. All three studies highlight that supply side factors play a limited role for explaining natural gas price fluctuations.

The two natural gas market studies most closely related to this paper are Wiggins and

Etienne (2017) and Rubaszek et al. (2021). The former use the sign restrictions of Kilian and Murphy (2014) combined with a time varying parameter (TVP) and stochastic volatility specification while the latter largely apply the oil market model of Baumeister and Hamilton (2019) to the natural gas market of the United States. Like this paper, both studies find that demand-side shocks played a major role in the decline of natural gas prices after 2008. Since both studies use competing modelling approaches this speaks for the robustness of the result – especially since the Baumeister and Hamilton (2019) approach allows for a larger contribution of supply side factors to oil price fluctuations than the model of Kilian and Murphy (2014). However, neither paper explicitly calculates or performs inference on a measure of the relative importance of demand-side shocks to the gas price decline in a given time window. To the best of my knowledge, this paper is the first to do so.

Drilling literature. Apart from the macroeconomic literature on fossil fuel markets discussed so far, an independent literature on drilling has developed in recent years. These studies have been developed against the background of major improvements in (horizontal) drilling and are characterised by the use of microdata, deep understanding of the drilling industry and reference to engineering knowledge. There are three major contributions of Ryan Kellogg, one of which is with coauthors: Kellogg (2011) shows how inter-firm learning increases the productivity of certain matches between fossil fuel companies and drilling contractors, which can explain the stability of contract relationships and the consistent improvement of productivity in oil and natural gas drilling. Kellogg (2014) uses data on oil drilling in Texas to show how oil price volatility influences drilling negatively. Finally, Anderson et al. (2018) show that oil production from existing wells in Texas is unresponsive to oil prices, which can be explained by laws of geology and petro-engineering. Drilling, on the other hand responds strongly to oil price changes. A similar result for natural gas is provided by Mason and Roberts (2018), who study natural gas wells in Wyoming and report price elasticities of supply for existing wells that are small in magnitude. The last two papers are crucial for my work, since they motivate including drilling activity into the gas market model and justify some of my identifying assumptions. Further evidence on the relationship between prices and drilling is provided by Khalifa et al. (2017) as well as Shakya et al. (2022) who conduct time series analysis of aggregate data. Using quantile regression techniques, the former show that there is a strong relationship between the rig count and oil prices in the US. The authors conclude, however, that this relationship is not contemporaneous but that the reaction of drilling to price changes takes time. The latter article

conducts Granger causality analysis to study the relationship between oil drilling, gas drilling, oil prices and gas prices. The authors use a rolling window approach to account for structural change and find that information transmitted between drilling and prices has increased since the development of shale oil and gas.

All in all, there seems to be a lack of structural macro-econometric models trying to explain the dynamics governing and effects of oil or gas drilling. Existing papers on drilling mainly focus on the micro-level or conduct traditional time series analysis using reduced form equations and ignoring general macroeconomic conditions or treating them as exogenous. Structural gas or oil market models, on the other hand, usually do not consider drilling. This paper tries to close some part of this gap.

3 The Econometric Model

3.1 Structural Vector Autoregression

To model the US natural gas market I propose a structural vector autoregressive (SVAR) model of the form

$$A_0 y_t = c + \sum_{j=1}^{p} A_j y_{t-j} + A_x x_t + \epsilon_t, \quad (\epsilon_t) \overset{iid}{\sim} N(0, I_4), \tag{1}$$

where y_t is a $n \times 1$ vector of endogenous variables and x_t is a $q \times 1$ vector of exogenous variables. The $n \times n$ matrix A_0 captures contemporaneous structural relationships between the endogenous variables, while the $n \times n$ matrices $A_1, ..., A_p$ and the $n \times q$ matrix A_x model the influence of past endogenous and current exogenous variables. The structural shocks ϵ_t are assumed to be Gaussian, uncorrelated and have unit variance. Left-multiplying (1) by A_0^{-1} yields the reduced form

$$y_t = d + \sum_{j=1}^{p} B_j y_{t-j} + B_x x_t + e_t, \quad (e_t) \overset{iid}{\sim} N(0, \Sigma), \tag{2}$$

where $B := [d, B_1, ..., B_p, B_x] = A_0^{-1}[c, A_1, ..., A_p, A_x]$ and $\Sigma = A_0^{-1}(A_0^{-1})'$. To estimate the reduced form, I choose a Bayesian approach. As is common (see Inoue and Kilian, 2013; Arias et al., 2018), I propose an agnostic Gaussian-Inverse Wishart prior and assume that

$$\Sigma \sim \mathcal{IW}_n(S_*, n_*), \quad \text{vec}(B)|\Sigma \sim N(\beta_*, V_* \otimes \Sigma), \tag{3}$$

with the hyper-parameters $n_*, S_*, \beta_*, V_*^{-1}$ being all set to zero, which yields an improper prior distribution. The Gaussian-Inverse Wishard prior is a natural conjugate prior, and therefore the posterior distribution has the same form as the prior-distribution (see Kilian and Lütkepohl, 2018, p. 162-165 for details). This has the advantage that after computing the posterior hyper-parameters one can sample from the posterior distribution of $[B, \Sigma]$ using an off-the-shelve random number generator.

3.2 Data

To model the joint dynamics of general macro-economic and gas-market specific conditions, I include four endogenous variables into the SVAR model: A measure of active drilling rigs (r_t), gas production (q_t), industrial production (i_t) and the natural gas price (p_t). Therefore, the vector of endogenous variables looks as follows:

$$y_t' = \begin{bmatrix} y_{t1} & y_{t2} & y_{t3} & y_{t4} \end{bmatrix} = \begin{bmatrix} r_t & q_t & i_t & p_t \end{bmatrix}. \tag{4}$$

As a measure of drilling, r_t, I include the number of active rigs currently used for gas drilling in the United States as provided by the Baker Hughes Rig Count.[2] Since drilling rigs are used to drill new wells and to deepen or sidetrack existing ones, their use serves as a broad measure of investment in natural gas development. Furthermore, since drilling is expensive and the economic viability of new production capacity is highly dependent on prices in the distant future (Mu, 2019), the number of active rigs can be interpreted as a barometer of expectations in the oil and gas industry. However, the most important reason to model drilling is recent research, which shown that the price elasticity of oil and natural gas production from existing wells is small, since production there is mainly determined by geological characteristics. The seminal paper here is (Anderson et al., 2018), who study oil markets, and Mason and Roberts (2018) provide a similar result for natural gas markets. As a consequence, oil and gas producers mainly react to price changes via the drilling rate, which is their most important decision variable. Therefore, drilling should be considered when modelling the natural gas market.

The choice of the remaining endogenous variables needs less explanation, since at least since Kilian (2009), oil production, the oil price and economic activity lie at the center of many oil market SVAR models and this framework has been adopted to study natural gas markets.

[2] Baker Hughes Rig Count via Energy Information Administration (2022) 'Crude Oil and Natural Gas Drilling Activity' (monthly): `https://www.eia.gov/dnav/ng/ng_enr_drill_s1_m.htm`.

Following Arora and Lieskovsky (2014) and Wiggins and Etienne (2017), I use marketed US natural gas withdrawals[3] as a measure of gas production, q_t, and real US industrial production[4] as i_t. While in oil market models it is common to include a *global* measure of economic activity such as the freight rate index of Kilian (2009), the US natural gas market can be considered as local. Li et al. (2014) show that between 1997 and 2011 there is no evidence for convergence between North American and European or Asian natural gas prices. Especially after 2008 the lack of or fully utilised capacity for liquefied natural gas exports served as a physical barrier, that prevented North American gas prices from converging to the much higher world market level. Finally, as the gas price, p_t, the Henry Hub price deflated by the US consumer price index[5] is used. The monthly Henry Hub price is provided by the Wall Street Journal and available for 1993M11-2014M02.[6] From 1997M01 on a monthly Henry Hub series is made available by the EIA.[7] Like Arora and Lieskovsky (2014) I first use the Wall Street Journal series, but I switch to the EIA series when the former is not available anymore.

As a next step, I present the exogenous variables x_t. Since variation of natural gas usage in the residential sector should to a large extent come from changes in the weather, I follow Brown and Yücel (2008) as well as Nick and Thoenes (2014) and include measures of temperature into the model. Heating or cooling degree days, as provided by the EIA[8], measure how cold or warm a certain location is. For an aggregate measure that can be used to assess national energy demand, local measures are population-weighted. Additionally to this I add dummy variables for each month to capture seasonal effects and two dummies to exclude two extraordinary events, the hurricanes Katrina in 2005M09 and Gustav in 2008M09. Katrina greatly reduced natural gas output, especially in the Gulf of Mexico, in an already tight market and led to a spike in natural gas prices (see DOE, 2006). Three years later, Gustav also reduced natural gas production but was not followed by a price hike since the economic crisis of 2008 already reduced gas demand (see CRS, 2009). The two hurricanes led to short-term reductions of gas production that are clearly extraordinary and violate the assumption of symmetric shocks. It

[3]Energy Information Administration (2022): 'US Natural Gas Marketed Production' (monthly): `https://www.eia.gov/dnav/ng/hist/n9050us2M.htm`

[4]FRED (2022) 'IPB50001N' (monthly): `https://fred.stlouisfed.org/series/IPB50001N`.

[5]FRED (2022) 'USACPIALLMINMEI' (monthly): `https://fred.stlouisfed.org/series/USACPIALLMINMEI`.

[6]Wall Street Journal via FRED (2016) 'GASPRICE' (monthly): `https://fred.stlouisfed.org/series/GASPRICE`.

[7]Energy Information Administration via FRED (2022) 'MHHNGSP' (monthly): `https://fred.stlouisfed.org/series/MHHNGSP`.

[8]Energy Information Administration (2022) 'Heating degree-days by census division' (monthly); 'Cooling degree-days by census division' (monthly): `https://www.eia.gov/totalenergy/data/monthly/index.php`

is hard to imagine a productivity shock rising gas production multiple percentage points above the level expected in the previous month before coming down rapidly. I also use a broken trend dummy of the form $\begin{bmatrix}0 & ... & 0 & 1 & 2 & 3 & ...\end{bmatrix}$ to account for a constant rise in gas production from the mid 2000s onwards. As is described by Kellogg (2011), Anderson et al. (2018), Wang and Krupnick (2015) and Mu (2019), among many others, from the early 2000s on innovations in hydraulic fracturing ('fracking') and horizontal drilling led to higher cost effectiveness and production levels in the oil and gas industry. The start of this broken trend is determined by applying the Bai and Perron (1998) procedure to the output series q_t, which gives 2005M05 as start point of the broken trend, which is also historically plausible as start of the shale gas boom (Mu, 2019). One could object to modelling the shale gas boom as an exogenous break in the trend. High natural gas prices after 2000 helped making already existing fracking techniques economically viable and set an incentive for developing new techniques. In this sense, the shale gas boom must be partly seen as an endogenous outcome. However, due to the singularity of the event and the non-linearity as well as long delay times and horizons involved – as Wang and Krupnick (2015) highlight, efforts to exploit shale gas started multiple decades before 2000 – the SVAR model is clearly not able to disentangle the effects that led to the shale boom. The exogenous dummies have to be interpreted as an auxiliary construction to deal with this shortcoming. As is shown in Section 5, however, the main results do not change when the broken trend is not included.

I use six lags of endogenous variables in my model, which is way above the number suggested by the AIC (one), since according to Kilian and Zhou (2020) up to two years of lags are needed to capture the relationship between economic activity and oil market variables. Additionally, Kellogg (2014) and Khalifa et al. (2017) report that drilling takes multiple months to react to price changes. The endogenous variables are included in log-levels, despite some of them being integrated of order 1, since this makes the historical decomposition of endogenous variables easier to interpret. The exogenous temperature series have the mean of the last five observations that were in the same month subtracted from itself to remove any possible trend as well as their seasonal pattern. Lastly, the SVAR model is estimated using monthly data starting in 1993M11 through 2020M01. I do not use data after the latter date because the Covid pandemic and the Russian invasion of Ukraine, which led to historic dislocations in global energy markets, were such exceptional events that it is questionable whether the linear SVAR model is appropriate for this period. This may raise legitimate doubts about the external validity of the results of this

study for the future. In the end, it boils down to the question of whether these extraordinary events will fundamentally change the functioning of energy markets in the long run. It is probably too early to tell, but caution is definitely warranted.

3.3 Identifying Assumptions

After estimating the reduced form (2), the structural form (1) is still to be identified. For a given reduced form model $[B, \Sigma]$ any set of structural parameters that satisfies

$$A_0^{-1} = CQ, \tag{5}$$

$$\begin{bmatrix} c & A_1 & ... & A_p & A_x \end{bmatrix} = A_0 \begin{bmatrix} d & B_1 & ... & B_p & B_x \end{bmatrix}, \tag{6}$$

is a corresponding structural model, where $C = \text{chol}(\Sigma)$ and $Q \in \mathcal{O}(n) := \{Q \in \mathbb{R}^{n \times n} | QQ' = Q'Q = I_n\}$ is an arbitrary orthonormal matrix. Therefore, without restricting the set of admissible structural models, the set of possible structural models is of the same magnitude as the $(n-1)n/2$-dimensional real numbers – even when ignoring the uncertainty involved in estimating the reduced form. As is common in the SVAR literature, I restrict the values of A_0^{-1} using a-priori knowledge. Table 1 summarizes the sign/exclusion restrictions that I place. First, I assume that the first of the four structural shocks that span the innovations in the reduced form model is a drilling shock. This shock leads to a rise in gas drilling and is the only shock that can affect the number of active rigs contemporaneously, since for contractual, legal and technical reasons the decision to drill takes about three months to result in actual drilling activities (Kellogg, 2014). With this assumption, the first structural shock as well as the first column of A_0^{-1} are point identified. The second structural shock, a gas supply shock, decreases gas production unexpectedly, which should increase the gas price and therefore decrease industrial production. The third shock, an economic activity shock, raises industrial production, which should push the gas price up, which should trigger a rise in natural gas production. Lastly, a demand shock specific to the gas market raises gas prices and thus gas production while decreasing industrial production. A gas demand shock can be thought of as a shift in preferences, representing changed heating and cooling behavior or a transformation of energy production due to environmental concerns. Another possible cause of gas demand shocks are substitution or spillover effects: Brown and Yücel (2008) report crude oil prices to be one of the main drivers

	Drilling Shock	Gas Supply Shock	Economic Activity Shock	Gas Demand Shock
Rig Count	*	0	0	0
Gas Production	*	−	+	+
Ind. Production	*	−	+	−
Gas Price	*	+	+	+

Table 1: Impact sign and exclusion restrictions for A_0^{-1}.
Notes: Additional elasticity restrictions are described in the text.

of natural gas prices in the US and for Germany Nick and Thoenes (2014) also conclude that oil prices are an important determinant of natural gas prices. Of course, the decoupling of oil and natural gas prices in the United States after the shale gas boom might have mitigated the importance of oil price shocks in recent years.

Not all structural models that satisfy the imposed impact sign restrictions are also economically plausible. Some may for example imply elasticities which are known to be false. As an answer, Kilian and Murphy (2012) suggest using prior knowledge about such elasticities to further restrict the space of admissible models. Therefore I use microeconometric evidence provided by Mason and Roberts (2018), who study price elasticities of already existing natural gas wells in Wyoming. They report a price elasticity of gas production of 0.0262 for the whole sample, which is almost exactly the value that Kilian and Murphy (2014) use as an upper bound for the impact price elasticity of oil supply. Splitting samples in four quartiles according to peak production rates gives Mason and Roberts (2018) no price elasticity that is significantly different from zero for the first two quartiles. The third quartile has an elasticity of 0.0311 and the fourth quartile of 0.0646. To be prudent I use 0.065 as an upper bound for the price elasticity of natural gas supply and discard all model draws, for which a_{23}/a_{43} or a_{24}/a_{44} are above this threshold, where a_{ij} denote the elements of A_0^{-1}. Kilian and Murphy (2012) also restrict the impact response of economic activity to a gas demand shock to a small interval. The numerical value for this restriction is the hardest to motivate on empirical grounds so Kilian and Murphy choose an arbitrary value close to zero and conduct sensitivity tests in which they vary that threshold. I restrict the immediate impact of a gas demand shock on industrial production to lie in the interval $(-0.004, 0)$, where 0.004 is half the standard deviation of the reduced form error of industrial production. As some kind of guideline, one can use the theoretical result of Hamilton (2013) that the energy price elasticity of national output should equal the energy bill as a fraction of GDP times the price-elasticity of energy demand. This result only refers to direct, supply-side effects and leaves aside changes in employment or capital formation which

one can assume to operate with some lag. In 2008, the US natural natural gas bill was the highest in the sample with about 200 billion 2019 Dollars[9], amounting to below 1.2 percent of GDP. Rubaszek et al. (2021) estimate a short term price elasticity of gas demand of -0.42. Even using the extremely cautious value of -1 for this elasticity would give a lower bound of roughly -0.0014 for the impact response of economic output to an unexpected price increase of 0.114 – one standard deviation of the reduced form error of the real gas price. This is about a third of the lower bound I use for a_{44}. Further justification for a bound close to zero is given by Baumeister and Hamilton (2019) who use a prior mean of -0.05 for the oil price elasticity of economic activity in the first month and obtain a posterior of almost exactly zero.

3.4 Identification Procedure

When using (dynamic) sign or elasticity restrictions to identify a model it is common to proceed as follows (Kilian and Lütkepohl, 2018, 421-490):

1. Additional to priors for Σ and B propose also a prior for Q on $\mathcal{O}(n)$. Usually a uniform prior is used.

2. Sample from the prior distribution of Σ, B and (most likely multiple times per reduced form draw) Q. Drawing from a uniform distribution on $\mathcal{O}(n)$ can be done by drawing $n \times n$ matrices with independent standard normal entries, computing the QR decomposition and retaining the matrix Q.

3. Check if $\text{chol}(\Sigma)Q$ satisfies all the imposed restrictions. If not, do not save the draw and continue sampling. If yes, save the structural model with $A_0^{-1} = \text{chol}(\Sigma)Q$ and $\begin{bmatrix} c & A_1 & ... & A_p & A_x \end{bmatrix} = A_0 B$.

This approach works, because when using sign restrictions the set of all structural parameters satisfying the sign restrictions has always positive Lebesgue measure in the set of all structural parameters (Arias et al., 2018).[10] Thus, one will always have a positive probability of drawing an admissible model when proceeding as above. In the case of zero restrictions, however, the probability of drawing a model that matches the restrictions will always be zero. This is why Arias et al. (2018) developed an algorithm to draw structural models that satisfy just the *zero*

[9] Source: EIA, Today in Energy, September 9, 2021.

[10] Of course, the only exception is the degenerate case where there is not a single structural model that satisfies the sign restrictions.

restrictions after which all the models that do not satisfy the *sign* restrictions can be deleted. In my case there is no need to apply this algorithm since – due to the fact that A_0^{-1} is assumed to be block lower triangular – one can use sub-rotations (Kilian and Lütkepohl, 2018, p. 473-475). It can be shown that the structural model satisfies the zero restrictions of Table 1 if and only if

$$Q = \begin{bmatrix} 1 & 0_{1\times 3} \\ 0_{3\times 1} & Q^* \end{bmatrix}, \tag{7}$$

where $Q^* \in \mathcal{O}(3)$ is an arbitrary orthonormal matrix. Therefore one can propose an uniform prior for the smaller-dimensional rotation Q^*. Afterwards, instead of sampling Q directly, one can sample $Q^* \in \mathcal{O}(3)$, build Q as in (7) and retain Q if the structural model satisfies the sign restrictions – the zero restrictions will be satisfied automatically due to the sampling procedure. This approach is not only faster and easier to implement than the more general procedure of Arias et al. (2018), it also has a nice geometric interpretation, which I discuss in more detail in the Appendix.

3.5 Inference on Impulse Response Functions

Usually, when using SVAR models, one wants to conduct inference, especially with regards to the structural impulse response functions (IRFs). The IRF $\theta_{ij,h}$ is the derivative of the endogenous variable i with respect to the structural shock j lagged by h periods. Even when working with point identified models, inference is not straightforward here, since usually one is interested in the joint behavior of the impulse response $\theta_{ij,h}$ for multiple values of i, j and h and therefore point-wise confidence intervals will be too narrow. This is a typical multiple inference problem, the solutions of which Lütkepohl et al. (2020) give an overview of. Additionaly to this problem, in the presence of sign restrictions there is not a single but many structural models that correspond to any point estimate of the reduced form parameters. Therefore, it is not clear which point estimate to report for the impulse response functions. It is common to just report the point-wise median as point estimate and point-wise quantiles as error bands. As Inoue and Kilian (2013) argue, however, the resulting confidence regions will not only be too small, i.e. they usually contain a way smaller fraction of all structural models than intended, but also the resulting point estimate might not have the shape of a single admissible IRF. As a solution, they suggest to report the modal IRF as the point estimate and a corresponding highest density region as confidence band, which is both optimal given a Dirac Delta loss function (Inoue and

Kilian, 2013, 2018).[11] In more detail, they do the following:

1. Draw M structural models $\{\begin{bmatrix} A_0 & A_1 & ... & A_p & A_x \end{bmatrix}_i\}_{i=1}^M$ and compute the IRFs $\{\Theta_i\}_{i=1}^M$.

2. Compute the posterior density $f_i = f(\Theta_i)$ for every IRF.

3. Compute the $(1-\alpha)$-quantile $f^{1-\alpha}$ of all IRF posterior values f_i.

4. Report the modal model as point estimate and the set $\mathcal{S}_\alpha$ as confidence region:

$$\hat{\Theta} := \underset{\Theta_i : i=1,..,M}{\arg\max} f(\Theta_i), \tag{8}$$

$$\mathcal{S}_\alpha := \{\Theta_i : f_i \geq f^{1-\alpha}\}. \tag{9}$$

In contrast to reporting point-wise quantiles, this approach yields a confidence region that is big enough to contain a fraction α of all structural IRFs and a point estimate that has the shape of an actual IRF in the posterior distribution. Since I am especially interested in the shape of the impulse response functions I choose this approach. The posterior density of the IRFs in my model is different from the one reported in Inoue and Kilian (2018), however, since I do not sample Q directly but draw the smaller-dimensional Q^* and construct Q as in (7). The IRF posterior density for this case is derived in the Appendix.

4 Results and Discussion

4.1 Natural Gas Market Dynamics

At first, I present the impulse response functions of the SVAR model. Figure 3 reports the modal model and a 68% highest density region that is constructed as described above. Since I am primarily interested in the determinants of natural gas drilling and prices, I only discuss the responses of the rig count and the real gas price to the four structural shocks, which are plotted in the first and last row. A drilling shock has no significant impact on natural gas prices in the first year and the same is true for the gas supply shock. For the latter, however, the point estimate is clearly positive, implying about 5 percent higher prices in the first six months following an initial supply shortfall of 1 percent. The economic activity shock and the gas specific demand shock both are followed by a significant positive price response. However, the

[11] Recently, Inoue and Kilian (2022) have generalized this approach using other loss functions than the Dirac Delta.

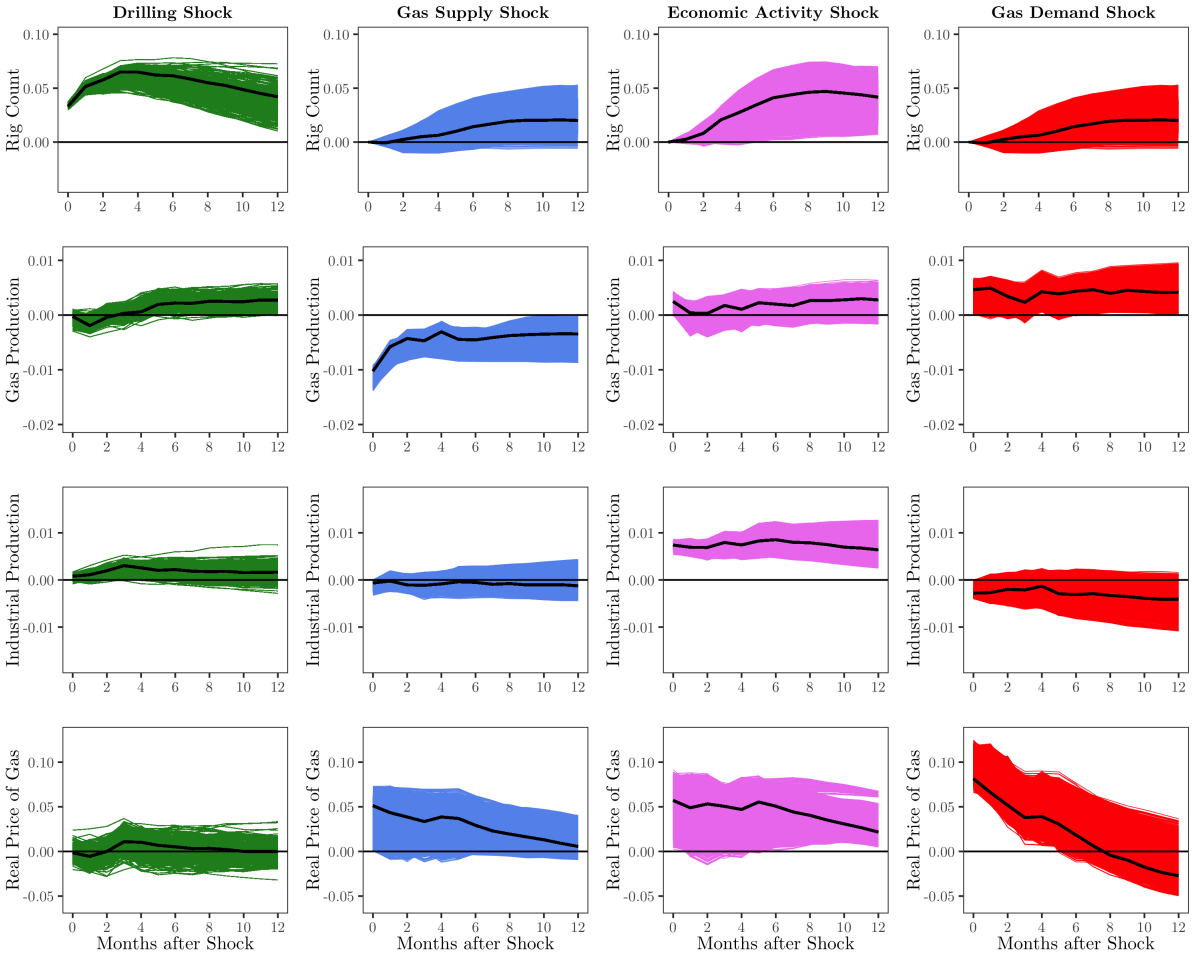


Figure 3: Structural impulse response functions.

Notes: Modal model and Bayesian 68% highest density region are constructed as defined in (8) and (9).

trajectory of the price effect looks completely different for the two shocks. This is why it was so important to retain the exact shape of the sampled IRFs when presenting point estimates and confidence regions, which would not have been the case if point-wise quantiles had been used. The price effect after an economic activity shock is persistent and only gets significant after about half a year. In contrast, after a gas specific demand shock prices rise high initially, but the effect starts to evaporate immediately afterwards. After six to eight months, the price effect is not significant any more.

As a next step, I present the response of drilling to the structural shocks. A drilling shock raises the rig count significantly and does so in a hump-shaped way. Moreover, at least according to the point estimates, any price increase seems to be followed by a positive reaction of drilling, regardless of whether this is due to a supply shortage or an increase in demand. This result is not a big surprise, as it is not only consistent with optimizing behaviour of companies but also with previous studies, which show that drilling is sensitive to price changes (Anderson et al.,

2018; Mason and Roberts, 2018). The size of the drilling response to a price change, however, seems to vary with the source of the shock. While the shock to gas specific demand triggers the strongest price increase, the rig count reacts only relatively weak. The economic activity shock seems to have the largest impact on drilling. Apparently the magnitude of the response of drilling to a change in the gas price not only depends on the initial size of this change but also on its source, i.e. on the structural shock that is responsible for the price change in the first place. This makes sense from a business point of view, considering that the shape of the price effects are completely different for the economic activity shock and the gas demand shock. While the former leads to a sustained price increase that can be exploited for quite some time, the price increase due to the latter disappears quickly. This is relevant for oil and gas producers since drilling takes quite some time and is an investment for years if not decades (Anderson et al., 2018; Mu, 2019). The following section will test the claim made here more formally.

4.2 The Price Elasticity of Drilling

The response of drilling to a change in the natural gas price has been studied in various papers such as Mason and Roberts (2018) or Khalifa et al. (2017). Often these studies set up a regression model of the form

$$r_t = \beta_h p_{t-h} + \gamma X_t + \eta_t, \tag{10}$$

where r_t is drilling, p_{t-h} is the fuel price studied lagged by h periods, X_t are other covariates, β_h and γ are regression coefficients and η_t is an error term. If the equation has a structural interpretation and the variables are transformed in log-levels, then the regression coefficient β_h can be interpreted as the price elasticity of drilling for the horizon h. As always, finding a plausible strategy to estimate β_h is hard. However, given some additional assumptions it is possible in my modelling framework. Suppose, the structural shock $\epsilon_{t-h,j}$ – which can be the gas supply shock, the economic activity shock or the gas specific demand shock – affects r_t only through p_{t-h}, then $\epsilon_{t-h,j}$ is a viable instrument to estimate β_h. In particular, one can then use the following identity:

$$\beta_h = \frac{\text{Cov}(\epsilon_{t-h,j}, r_t)}{\text{Cov}(\epsilon_{t-h,j}, p_{t-h})} = \frac{\theta_{1j,h}}{\theta_{4j,0}} =: \kappa_h^j, \tag{11}$$

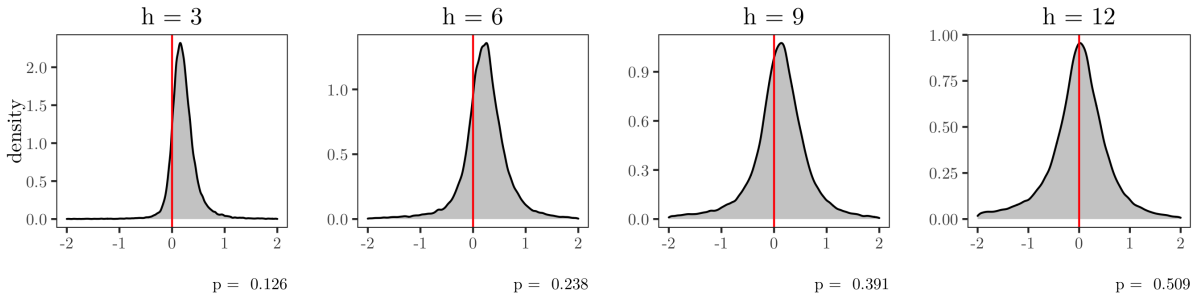
where $\theta_{ij,h}$ is the impulse response of variable i to shock j after h periods. Since it is

possible to sample from the posterior distribution of the IRFs, one can easily compute the posterior distribution of κ_h^j, which can be interpreted as the causal effect of p_{t-h} on r_t. If the price elasticity of drilling does not depend on the source of the structural price shock, as (10) would imply, then β_h should not depend on the fact whether the gas supply shock, the economic activity shock or the gas specific demand shock is used an instrument, i.e. κ_h^j should be the same for $j = 2, 3, 4$. Given the assumption placed above, one can test this equality to determine if the price elasticity of drilling is in fact independent of the shock that causes the price change.

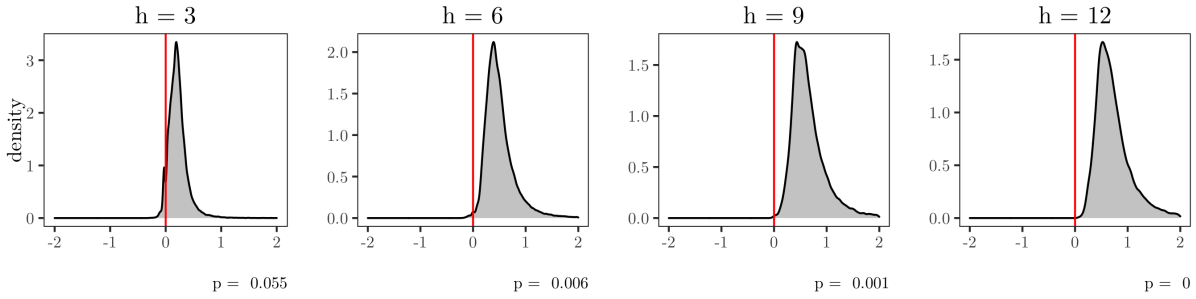
While for small horizons of two or three months the identifying assumption that $\epsilon_{t-h,j}$ only works through p_{t-h} might be plausible, since r_t is not able to react to $p_{t-h+1}, p_{t-h+2}...$ due to the sluggish nature of drilling[12], this is clearly not the case for large h. In this case – if one is willing to make the weaker assumption that $\epsilon_{t,j}$ only affects the rig count through the price channel, regardless of the timing – κ_h^j has to be interpreted as the causal effect of the sequence of price changes that is triggered by a structural shock in period $t-h$ on drilling in period t. In other words, it is the effect on r_t of changing $\{p_{t-h}, p_{t-h+1}, ..., p_t\}$ by $\{1, \theta_{4j,1}/\theta_{4j,0}, ..., \theta_{4j,h}/\theta_{4i,0}\}$. Therefore, a difference in κ_h^j for two different structural shocks does not need to be an indicator for β_h being different for the two shocks. Also differences in the dynamic price responses $\theta_{4j,1}/\theta_{4j,0}, \theta_{4j,2}/\theta_{4j,0}, ...$ – i.e. the shape of of the IRFs – can be responsible for differences in κ_h^j across shocks. After dropping the strong assumption that only p_{t-h} and no price change after $t-h$ has an effect on r_t, one does not compare elasticities in the sense as defined above. However, I think that κ_h^j is still a meaningful measure and maybe even be more informative than the elasticity β_h since it tells how much more or less drilling one should expect h months after observing a one percent price change that is caused by the structural shock $\epsilon_{t,j}$.

Figure 4 plots the posterior densities of the differences $\kappa_h^j - \kappa_h^k$ for two different structural shocks $\epsilon_{t,j}$ and $\epsilon_{t,k}$ for various horizons h. The value p denotes the fraction of IRFs that imply a non-positive value of the difference, i.e. a stronger drilling response after shock k than after shock j. The top panel compares the drilling response to price changes due to economic activity and gas supply shocks. One can see that after one quarter, more than 85 percent of all models imply a stronger drilling response after the former than after the latter shock. However, for a longer horizon this difference vanishes. This suggests that differences in the response of drilling to price changes are short-term at best for these two shocks. When comparing economic activity and gas specific demand shocks this is not the case. Here the

[12]This is why Kellogg (2014) lagged oil price volatility by three months when regressing drilling onto it.



(a) $\kappa_h^{\text{econ. activity}} - \kappa_h^{\text{gas supply}}$



(b) $\kappa_h^{\text{econ. activity}} - \kappa_h^{\text{gas demand}}$

Figure 4: Posterior distributions of the differences of κ_h by shock.

Notes: This figure plots the posterior distribution of $\kappa_h^j - \kappa_h^k = \theta_{1j,h}/\theta_{4j,1} - \theta_{1k,h}/\theta_{4k,0}$ for two structural shocks.

economic activity shock produces a stronger drilling response with a high probability for every plotted horizon h. After a year, almost the whole posterior distribution of the IRFs implies a stronger drilling response after an economic activity shock than after a gas demand shock. Since for a horizon of six months the price effect of a gas demand shock is already gone, most likely for large h some part of the difference in κ_h^j comes from the difference in the shape of the IRF. However, looking at the density plot for $h = 3$, one has reason to believe that also the causal effect of p_{t-3} on r_t is smaller after a gas demand shock than after an economic activity shock. This could be because oil and gas producers already know that a demand-side shock that is not underpinned by strong economic performance will lead to price effects that are short-lived. This is a possible explanation for why after the Russian invasion of Ukraine natural gas drilling in the US remained at a relatively low level despite the highest US natural gas prices since 2008.[13] Natural gas producers may have been uncertain whether prices would remain high for long, as they were not primarily driven by strong economic activity. In retrospect, these doubts would have been justified, because in the winter of 2022 – about half a year after the start of the war

[13]The number of active natural gas rigs, according to the Baker Hughes rig count, never reached 170 in a month after the Russian invasion of Ukraine, while until December 2015 and between May 2017 and July 2019 the number of active rigs was never below 170 despite sometimes way lower natural gas prices. As a comparison: Between 2004 and 2008, the rig count was above 1000 for every single month.

in Ukraine – natural gas prices in the USA fell to a lower level than before the war had begun.

4.3 Disentangling Historical Gas Price Fluctuations

The second big question this paper wants to address is what has historically been responsible for fluctuations in the price of natural gas in the US – especially for the huge price drop after 2008. To this end I present a historical decomposition of the natural gas price for the sample period. Given endogenous variables $\begin{bmatrix} y_{-p+1} & y_{-p+2} & ... & y_T \end{bmatrix}$, exogenous variables $\begin{bmatrix} x_1 & ... & x_T \end{bmatrix}$, an estimate of the reduced form slope coefficients $\hat{B}$ and the contemporaneous effects matrix $\hat{A}_0$ one computes estimates for the structural shocks $\hat{E} = \begin{bmatrix} \hat{\epsilon}_1 & ... & \hat{\epsilon}_T \end{bmatrix}$ as

$$\hat{E} = \hat{A}_0(Y - \hat{B}Z), \tag{12}$$

where $Y = \begin{bmatrix} y_1, ..., y_T \end{bmatrix}$ and column t of Z is defined as $Z_t := \begin{bmatrix} 1 & y'_{t-1} & ... & y'_{t-p} & x'_t \end{bmatrix}'$. Afterwards, one can define a vector $\hat{\epsilon}^i_t$ with $\hat{\epsilon}^i_{t,j} := \delta_{ij}\hat{\epsilon}_{t,i}$, where δ_{ij} is the Dirac delta and $i \in \{0, ..., n\}$. In words, $\hat{\epsilon}^i_t$ is the counterfactual vector of structural shocks, where all shocks other than $\epsilon_{t,i}$ have been set to zero and shock $\epsilon_{t,i}$ is as estimated from observed data. Note that $\hat{\epsilon}^0_t$ is the zero-vector. Afterwards, one can recursively compute the decomposition:

$$\hat{y}^i_t = \hat{d} + \sum_{j=1}^{p} \hat{B}_j \hat{y}^i_{t-j} + \hat{B}_x x_t + \hat{A}_0^{-1}\hat{\epsilon}^i_t, \tag{13}$$

with starting values $\begin{bmatrix} \hat{y}^i_{-p+1} & ... & \hat{y}^i_0 \end{bmatrix} := \begin{bmatrix} y_{-p+1} & ... & y_0 \end{bmatrix}$. This is not a decomposition in the sense known from stationary, mean-zero AR(p) systems (see for example Kilian and Lütkepohl, 2018, p. 116-122). Neither does $\hat{y}^i_t$ only contain past structural errors $\epsilon t, i$, nor does the sum of all decompositions equal the observed value y_t. The first point is due to the fact that the starting values contain all kind of structural shocks whose effects cannot be expected to vanish over time. Instead, $\hat{y}^i_t$ has to be interpreted as the counterfactual value of y_t for the case of all structural shocks other than shock $\epsilon_{t,i}$ having been zero *from* $t = 1$ *on*. The second point raises the question how to present the historical decomposition in a way that the relative contributions of different structural shocks can be interpreted qualitatively. For stationary, mean-zero AR(p) processes, a natural comparison is the constant zero-function, since in the absence of any shocks the process converges to this value. To get a series of meaningful comparison values in our case, one has to compute $\hat{y}^0_t$ using the recursive procedure given in (13). This is the counterfactual value of y_t for the case of no shocks from $t = 1$ on. If shock $\epsilon_{t,i}$ has a positive influence on

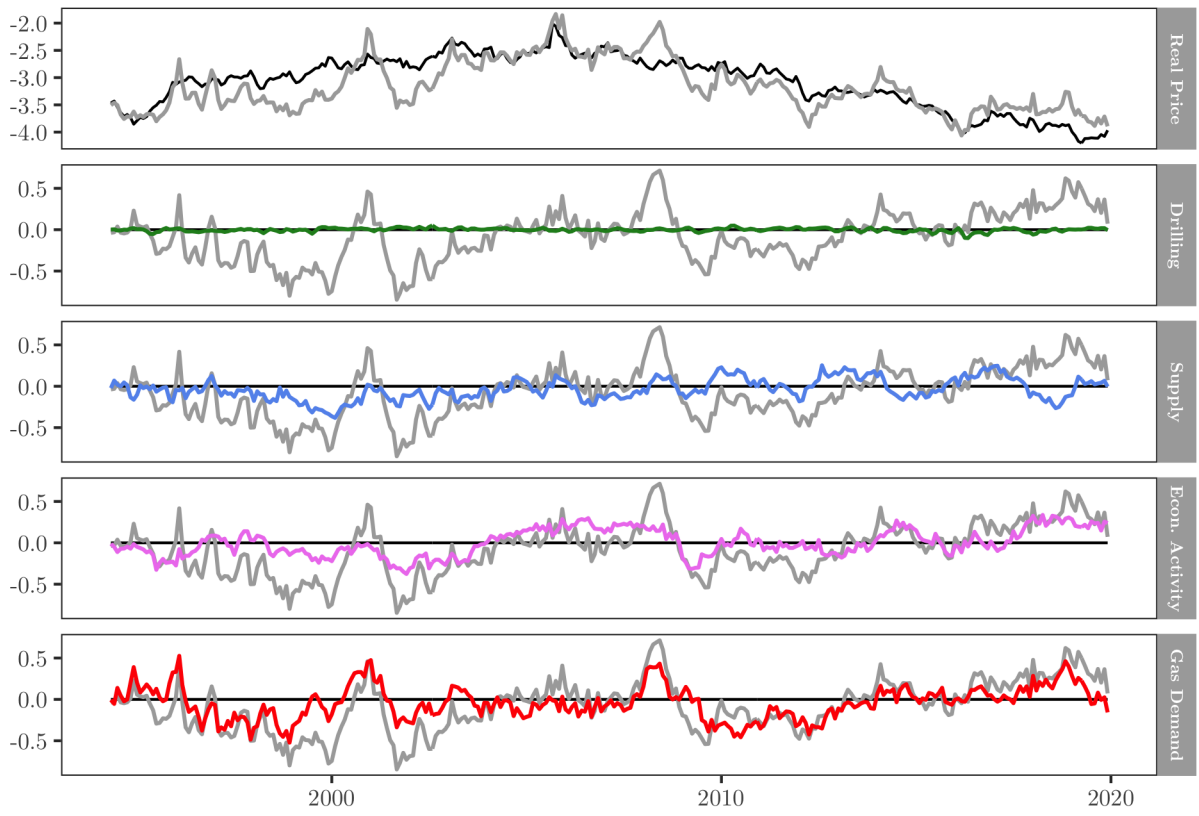


Figure 5: Historical decomposition of the natural gas price.

Notes: The comparison value $y^0_{4,t}$ is plotted in black, the grey line in the top panel is the true rig count $y_{4,t}$ as included in the model. The remaining panels present the historical decomposition series that were computed using (13) minus the comparison value, i.e. $(p^i_t - p^0_t)$ (color), as well as the true rig count minus the comparison value, i.e. $(p_t - p^0_t)$ (grey).

$y_{k,t}$ then $\hat{y}^i_{k,t}$ lies above $\hat{y}^0_{k,t}$. Further, this comparison value has the nice property that the differences $(\hat{y}^i_{k,t} - \hat{y}^0_{k,t})$ sum up to $(y_{k,t} - \hat{y}^0_{k,t})$, which somehow justifies the term decomposition.

Figure 5 shows the historical decomposition for the real price of natural gas that corresponds to the modal model presented in Figure 3. The top panel plots the true gas price (in log-levels, grey line) together with the comparison value $\hat{p}^0_t$ (black). One can see that this comparison value rises until 2005 and starts to decline afterwards, which can be attributed to the shale gas boom. However, there is much that cannot be explained by the deterministic part of this series. The lower four panels plot the difference between the gas price and the comparison value (grey) together with the difference $\hat{p}^i_t - \hat{p}^0_t$ for the shock i (colored line). If the latter value is positive for a given month, it means that the shock i pushed prices up at that time. The results nicely match the economic calendar: Before 2000, natural gas prices were unusually low, which was mainly due to gas specific demand shocks. This could be due to extremely

cheap oil after the Asian financial crisis.[14] In 2000 and 2001, natural gas prices spiked. The model correctly attributes the biggest share of this increase to rising gas demand. The reason for the surge in natural gas prices was the California energy crisis, where artificial tightening of electric energy supply and market manipulation led to a surge in demand for natural gas, since California relied heavily on this fuel for power generation (see Wilson, 2002, for details on this event). Shortly afterwards prices fell, most likely due to the economic crisis in the aftermath of the Dotcom bubble and falling oil prices. In the years before the financial crisis of 2008 economic activity shocks pushed the gas price up and shortly before the great recession, natural gas prices went up sharply. According to CRS (2009), a rise in demand from the power generation (+2.8%) and residential (+3%) sector in 2008 could have been responsible for this. Additionally, the price of oil was surging, with the sort WTI hitting a record high of almost 140 USD/Barrel in June 2008. After this peak, the natural gas price was in free fall from 2008 to 2012. Note that it are demand-side factors and not supply changes due to the shale boom that seem to be mainly responsible for this price drop. This is not self-evident, since economists and market analysts have attributed the drop of the natural gas price around the financial crisis often to the shale boom (Wiggins and Etienne, 2017). Hausman and Kellogg (2015) for example estimate that demand-side factors are responsible for less than 20 percent of the fall in natural gas prices between 2007 and 2013, while the shale gas revolution is argued to be responsible for the remainder. My model results challenge the predominant role of supply-side factors. The question of what was really responsible for the rapid fall in energy prices in the US after 2008 is highly relevant, as the US experience acted and still acts as a blueprint for many countries to develop shale gas resources themselves. Therefore it is worthwhile to look at this period in greater detail.

4.4 What Drove Prices Down After 2008?

To explicitly compute the contributions of different shocks and the deterministic component to a change in the natural gas price, one can can decompose the price change between t and $t+h$ as follows:

[14] In November 1998, the nominal price for one Barrel of WTI was the lowest since 1986. Source: FRED (2022) 'DCOILWTICO' (monthly): `https://fred.stlouisfed.org/series/DCOILWTICO`.

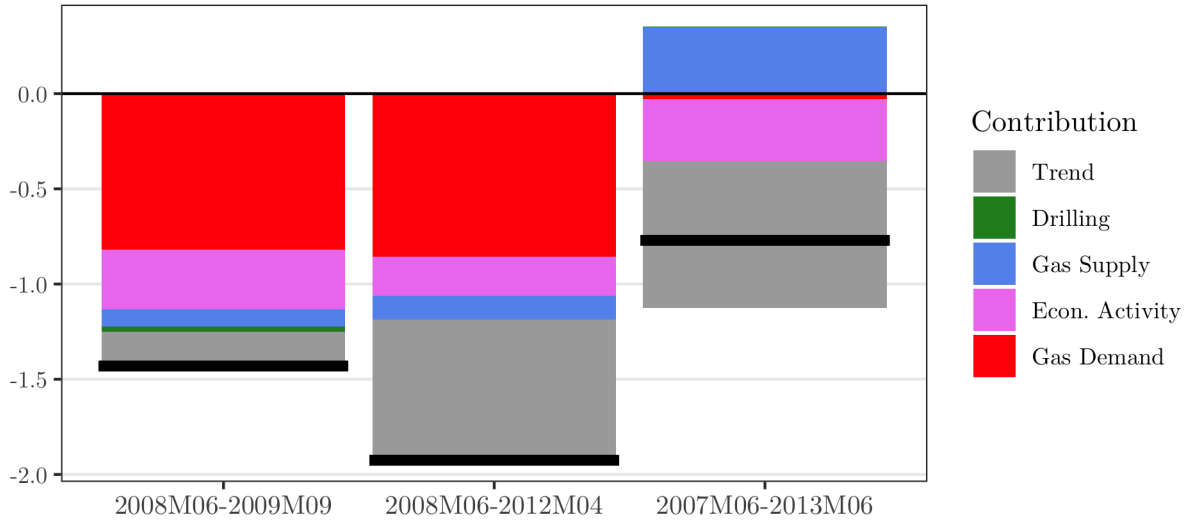


Figure 6: Contributions of trend and different shocks to natural gas price collapse around 2008.
Notes: Solid lines mark changes in log real natural gas prices in two time windows. Colored bars represent contributions of fitted trend and different structural shocks to this price change that were computed using (14).

$$\begin{aligned} p_{t+h} - p_t &= (p_{t+h} - \hat{p}^0_{t+h}) + (\hat{p}^0_{t+h} - \hat{p}^0_t) + (\hat{p}^0_t - p_t) \\ &= \underbrace{(\hat{p}^0_{t+h} - \hat{p}^0_t)}_{C_0^{t,h}} + \sum_{i=1}^{n} \underbrace{[(\hat{p}^i_{t+h} - \hat{p}^0_{t+h}) + (\hat{p}^0_t - \hat{p}^i_t)]}_{C_i^{t,h}}. \end{aligned} \tag{14}$$

Here, the $C_i^{t,h}$'s are the contributions of the n different shocks to the price change and $C_0^{t,h}$ is the contribution of the deterministic component. To be cautious, I interpret C_0 as effect of the shale gas boom, i.e. a supply-side factor.[15] Figure 6 plots the historical contributions to the gas price drop for three time-windows using the same structural model as in Figure 5. From its peak in June 2008, the natural gas price began to fall rapidly, reaching a local minimum in September 2009. My model attributes more than three quarters of this fall to gas demand and economic activity shocks. June 2008 to April 2012 is the time window in which the gas price fell by a dramatic 85 percent. Here, too, the effect of demand-side shocks is estimated to be bigger than that of the shale gas boom and other supply-side factors. Last but not least, I look at the time window June 2007 to June 2013 to establish some comparability with the results of Hausman and Kellogg (2015), even though they used annual data and therefore I do

[15] Apart from a broken trend component, C_0 also contains seasonal and temperature effects. Moreover, another long-term trend such as the secular stagnation, which would be a demand-side factor in my framework, could have contributed to the falling trend in natural gas prices. However, since I argue for the importance of demand-side factors in this paper, it is appropriate to consider C_0 entirely as a supply-side component.

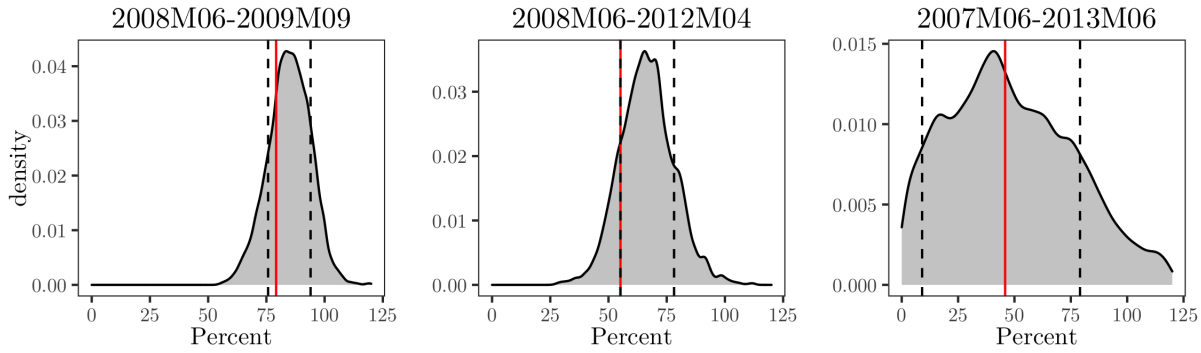


Figure 7: Posterior distribution of the fraction of the natural gas price drop that can be attributed to demand-side shocks (econ. activity and gas demand shocks) in three time intervals. *Notes:* Red line represents fraction that corresponds to the modal IRF, dotted lines are endpoints of the 68% equal-tailed confidence intervals. Plotted are distributions of the random variable defined in (15).

not estimate exactly the same quantity. Still, my model attributes more than 40 percent of the fall of the gas price to demand-side shocks, which is more than double the estimate of Hausman and Kellogg (2015). For this time window, however, the shale gas boom might have been the most important driver behind the fall of natural gas prices.

So far, I have only considered point estimates for the historical decomposition. To determine the uncertainty in the share of the gas price decline that is due to demand-side factors, one can directly calculate the posterior distribution of the following estimate:

$$100 \cdot \frac{C_3^{t,h} + C_4^{t,h}}{p_{t+h} - p_t}, \tag{15}$$

where $C_3^{t,h}$ and $C_4^{t,h}$ are the contributions of the economic activity shock and the gas specific demand shock, respectively. Figure 7 plots the posterior distribution of this estimate for the three time windows examined above. The graph also shows the value corresponding to the model in Figure 6 (red) and a 68% equal-tailed confidence interval (black, dashed). The confidence interval is above 75% for, i.e. the time window from June 2008 to September 2009, which is the immediate price drop after the financial crisis, and still above 50% for the longer time window until April 2012, when natural gas prices fell by 85%. From this, one can conclude that for the rapid price decline after 2008 demand-side factors were more important than the shale gas boom. For the six-year window June 2007 until June 2013, however, the estimate is so imprecise that it is not possible to reject any meaningful hypothesis about the role of supply and demand-side shocks in this period. In particular, the estimate of Hausman and Kellogg (2015), which also has a high variance, is covered by my confidence interval.

5 Robustness Checks

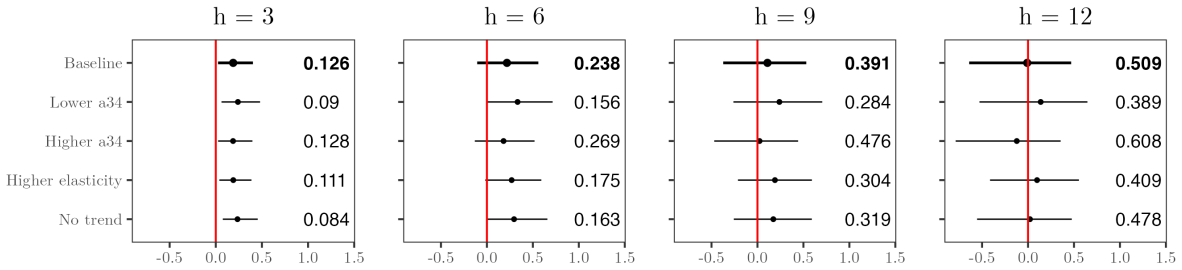
Key results in structural vector autoregressive models often rely heavily on certain prior assumptions. Baumeister and Hamilton (2019), for example, report a greater importance of oil supply shocks in explaining oil price fluctuations than Kilian and Murphy (2012, 2014), once the short-run price elasticity of oil supply is allowed to take on values greater than 0.0258, which is the bound of Kilian and Murphy. Also in this paper's model, some parts of the forecast error variance decomposition (not reported) depend heavily on the influence that a gas demand shock is allowed to have on industrial production. While especially in the case of set-identification it is the responsibility of the researcher to employ strong prior information in order to not let economically implausible models dominate the posterior distribution, it is important to know how sensitive the results are to slight changes in prior assumptions. Therefore I conduct the following robustness checks:

- *Lower a34.* As described above, the maximum impact a gas specific demand shock is allowed to have on industrial production in the first month is neither strongly rooted in theory nor empirics. Instead of -0.004, which is half the standard deviation of the reduced form error of industrial production times minus one, I use -0.0027 as a lower bound, which is a third of this standard deviation times minus one.

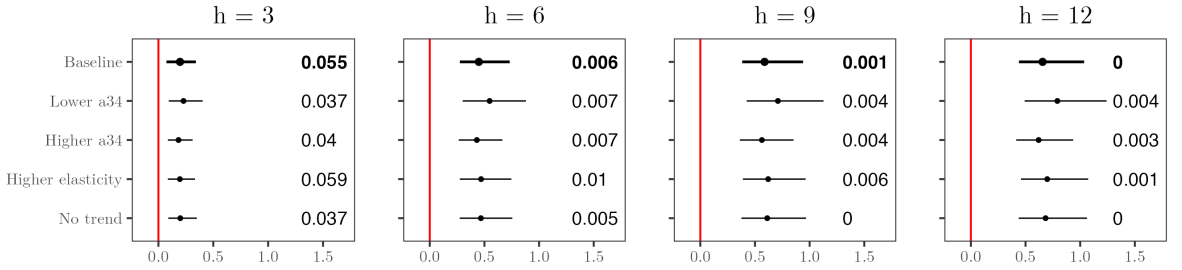
- *Higher a34.* As a next step, I use -0.0053 as a lower bound of the effect of a gas demand shock on industrial production, which is two thirds of the reduced form error of industrial production times minus one.

- *Higher elasticity.* My upper bound for the short-run price elasticity of natural gas supply was chosen rather high. It is the point estimate for the most productive quarter of all active rigs, as reported by Mason and Roberts (2018), which is more than double the estimated average value for all rigs. Nevertheless, I re-run my analysis using 0.085 as an upper bound for this elasticity, which is two standard errors above the point estimate 0.065.

- *No trend.* In my opinion, the broken trend in this paper's model is an appropriate way to deal with the shale gas boom. To show that this modelling choice is not responsible for my main findings, I re-estimate the model using no such trend component.



(a) $\kappa_h^{\text{econ. activity}} - \kappa_h^{\text{gas supply}}$



(b) $\kappa_h^{\text{econ. activity}} - \kappa_h^{\text{gas demand}}$

Figure 8: Robustness of the posterior distributions of the differences of κ_h by shock.

Notes: This figure plots the median as well as the equal-tailed 68 percent confidence interval for the posterior distribution of $\kappa_h^j - \kappa_h^k = \theta_{1j,h}/\theta_{4j,1} - \theta_{1k,h}/\theta_{4k,0}$ for the baseline model and various specifications. The numbers in the panels denote the fraction of all models where $\kappa_h^j \leq \kappa_h^k$.

Figure 8 plots the median as well as the equal-tailed 68 percent confidence band for the posterior distributions of Figure 4. Also depicted are the fractions of all models for which the response of drilling to a price change is smaller after an economic activity shock than after a gas supply or gas specific demand shock, respectively. For every specification, price changes due to economic activity shocks lead to a stronger reaction of drilling than price changes due to other demand-side influences, especially for large horizons. The second result is also remarkably robust: Figure 9 shows that for all specifications the model attributes the majority of the natural gas price collapse after June 2008 to demand-side factors. For the time window from June 2007 to June 2013, the median of the distribution is also rather robust. The only visible difference for the longer time window is the confidence region of the estimate for the specification without trend, which is way longer than in the baseline model. However, this only confirms my conclusion that my model does not allow an accurate estimate of the demand-side contribution for this time period. All in all, my two main results are very robust with respect to the specific values used for the identifying assumptions and the modelling of a broken trend.

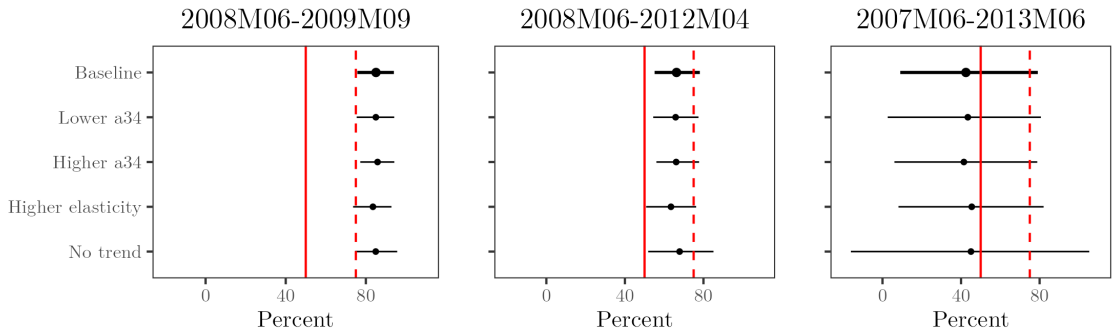


Figure 9: Robustness of the posterior distribution of the fraction of the natural gas price drop that can be attributed to demand-side shocks (econ. activity and gas demand shocks) in three time intervals.

Notes: Plotted are the median as well as the 68 percent equal-tailed confidence band of the distribution of the random variable defined in (15) for the baseline model and various specifications. The solid and dashed red lines mark 50 and 75 percent, respectively.

6 Conclusion

To study natural gas market dynamics in the United States I set up a structural vector autoregressive model that includes the number of active drilling rigs, natural gas production, industrial production and the real price of gas. To identify the structural form I employ a set of both, exclusion and sign restrictions. For conducting inference I adapt the procedure of Inoue and Kilian (2013, 2018) that was originally designed for purely sign identified models.

As a new contribution, I present evidence for the fact that the response of drilling to unexpected changes in the price of gas depends on the source leading to the price change. In particular, a price change due to an economic activity shock leads to a much stronger response in drilling than a price change due to different demand-side factors such as changing preferences or spillover effects from other energy markets. One explanation is that economic activity shocks lead to persistent price changes, while the effects of the latter shock on the price of gas disappear quickly, so higher prices cannot be exploited for long. This may be the reason why, after the Russian invasion of Ukraine, natural gas drilling in the U.S. remained relatively low, despite the highest U.S. natural gas prices since 2008. Perhaps natural gas producers were skeptical about the duration of the price increase, as it was not primarily driven by strong economic activity. Since previous studies such as Anderson et al. (2018) and Mason and Roberts (2018) have found that oil and gas producers increase their supply primarily through drilling, this finding is highly relevant for policymakers seeking to expand domestic energy production.

Expanding natural gas production is usually not an end in itself. Although national security

of supply and geopolitical autonomy have recently become the focus of renewed attention, governments typically hope that greater supply will lead to cheaper energy for national businesses and the electorate. It is therefore worthwhile to disentangle the reasons for the U.S. natural gas price collapse around the financial crisis. Hausman and Kellogg (2015) attribute this fall to the shale gas revolution – a view they seem to share with most economists and analysts (Wiggins and Etienne, 2017) – and estimate that demand side factors are responsible for less than 20 percent of the total price drop between 2007 and 2013. Although the variance of my estimate for the period 2007 to 2013 is too high to reject the estimated value of Hausman and Kellogg (2015), my model suggests that demand-side shocks were extremely important for explaining the drop in natural gas prices in this time. This is especially true for the immediate period after the financial crisis, when natural gas prices fell by an impressive 85 percent between June 2008 and April 2012. I show that demand-side factors were more important than the shale gas boom in explaining the price decline during this time window. This evidence of the importance of demand side shocks after 2008 should dampen the hopes of governments working on their own shale gas booms. As follows from Wang and Krupnick (2015) and Mu (2019), most countries lack one or many of the peculiarities that allowed for the massive boom in shale gas extraction in the United States.[16] However, even with a comparatively rapid expansion of natural gas production, no one should expect a price decline similar to that seen in the United States after 2008.

References

S. T. Anderson, R. Kellogg, and S. W. Salant. Hotelling under pressure. *Journal of Political Economy*, 126(3):984–1026, June 2018.

J. E. Arias, J. F. Rubio-Ramírez, and D. F. Waggoner. Inference based on structual vector autoregressions identified with sign and zero restrictions: Theory and applications. *Econometrica*, 86(2):685–720, March 2018.

V. Arora and J. Lieskovsky. Natural gas and u.s. economic activity. *The Energy Journal*, 35 (3):167–182, July 2014.

J. Bai and P. Perron. Estimating and testing linear models with multiple structural changes. *Econometrica*, 66(1):47–78, January 1998.

C. Baumeister and J. D. Hamilton. Sign restrictions, structural vector autoregressions, and useful prior information. *Econometrica*, 83(5):1963–1999, September 2015.

C. Baumeister and J. D. Hamilton. Structural interpretation of vector autoregressions with incomplete identification. *American Economic Review*, 109(5):1873–1910, May 2019.

S. P. A. Brown and M. K. Yücel. What drives natural gas prices? *The Energy Journal*, 29(2): 45–60, 2008.

[16] Examples for this are water availability, deep capital markets, a dynamic service-sector industry and the specific market structure of the oil and gas industry in the United States.

CRS. Natural gas markets: An overview of 2008. Congressional Research Service Report for Congress, 2009.

DOE. Impact of the 2005 hurricanes on the natural gas industry in the gulf of mexico region. United States of America Department of Energy Report, 2006.

J. D. Hamilton. What is an oil shock? *Journal of Econometrics*, 113(2):363–398, April 2003.

J. D. Hamilton. *Handbook on Energy and Climate Change*, chapter Oil Prices, Exhaustible Resources, and Economic Growth, pages 29–63. Edward Elgar Publishing, 2013.

C. Hausman and R. Kellogg. Welfare and distributional implications of shale gas. *Brookings Papers on Economic Activity*, pages 71–125, 2015.

C. Hou and B. H. Nguyen. Understanding the us natural gas market: A markov switching var approach. *Energy Economics*, 75:42–53, 2018.

IGU. Wholesale gas price survey: A global review of price formation mechanisms. Technical report, International Gas Union, 2021.

A. Inoue and L. Kilian. Inference on impulse response functions in structural var models. *Journal of Econometrics*, 177(1):1–13, 2013.

A. Inoue and L. Kilian. Corrigendum to "inference on impulse response functions in structural var models". *Journal of Econometrics*, 209(1):139–143, March 2018.

A. Inoue and L. Kilian. Joint bayesian inference about impulse responses in var models. *Journal of Econometrics*, 231(2):457–476, December 2022.

R. Kellogg. Learning by drilling: Interfirm learning and relationship persistence in the texas oilpatch. *Quarterly Journal of Economics*, 126(4):1961–2004, November 2011.

R. Kellogg. The effect of uncertainty on investment: Evidence from texas oil drilling. *American Economic Review*, 104(6):1698–1734, June 2014.

A. Khalifa, M. Caporin, and S. Hammoudeh. The relationship between oil prices and rig counts: The importance of lags. *Energy Economics*, 63:213–226, 2017.

L. Kilian. Not all oil price shocks are alike: Disentangling demand and supply shocks in the crude oil market. *American Economic Review*, 99(3):1053–1069, June 2009.

L. Kilian and H. Lütkepohl. *Structural vector autoregressive analysis*. Themes in modern econometrics. Cambridge University Press, 2018.

L. Kilian and D. P. Murphy. Why agnostic sign restrictions are not enough: understanding the dynamics of oil market var models. *Journal of the European Economic Association*, 10(5): 1166–1188, October 2012.

L. Kilian and D. P. Murphy. The role of inventories and speculative trading in the global market for crude oil. *Journal of Applied Econometrics*, 29(3):454–478, April/May 2014.

L. Kilian and X. Zhou. The econometrics of oil market var models. *Federal Reserve Bank of Dallas Working Paper*, March 2020.

C. A. León, J.-C. Massé, and L.-P. Rivest. A statistical model for random rotations. *Journal of Multivatiate Analysis*, 97(2):412–430, 2006.

R. Li, R. Joyeux, and R. D. Ripple. International natural gas market integration. *The Energy Journal*, 35(4):159–179, October 2014.

H. Lütkepohl, A. Staszewska-Bystrova, and P. Winker. Constricting joint confidence bands for impulse response functions of var models - a review. *Econometrics and Statistics*, 13:69–83, 2020.

J. R. Magnus. *Linear structures*. Oxford University Press, 1988.

C. F. Mason and G. Roberts. Price elasticity of supply and productivity: An analysis of natural gas wells in wyoming. *The Energy Journal*, 39(1):79–100, 2018.

X. Mu. *The Economics of Oil and Gas*. Agenda Publishing, 2019.

S. Nick and S. Thoenes. What drives natural gas prices? a structural var approach. *Energy Economics*, 45(C):517–527, 2014.

M. Rubaszek and G. S. Uddin. The role of underground storage in the dynamics of the us natural gas market: A threshold model analysis. *Energy Economics*, 87, 2020.

M. Rubaszek, K. Szafranek, and G. S. Uddin. The dynamics and elasticities on the u.s. natural gas market. a bayesian structural var analysis. *Energy Economics*, 103, 2021.

S. Shakya, B. Li, and X. Etienne. Shale revolution, oil and gas prices, and drilling activities in the united states. *Energy Economics*, 108, 2022.

H. Uhlig. What are the effects of monetary policy on output? results from an agnostic identification procedure. *Journal of Monetary Economics*, 52:381–419, 2005.

Z. Wang and A. Krupnick. A retrospective review of shale gas development in the united states: What led to the boom? *Economics of Energy and Environmental Policy*, 4(1):5–17, 2015.

S. Wiggins and X. L. Etienne. Turbulent times: Uncovering the origins of us natural gas price fluctuations since deregulation. *Energy Economics*, 64:197–205, 2017.

J. F. Wilson. High natural gas prices in california, 2000-2001: Causes and lessons. *Journal of Industry, Competition and Trade*, 2(1-2):39–57, June 2002.

Appendix A. On Rotations and Sub-Rotations

This section presents proofs and additional intuition for statistical claims made in the paper. These ideas are not new (see Kilian and Lütkepohl, 2018, p. 473-475 for an example of sub-rotations), I rather give them in a form that helped me understand in depth how the modelling approach of this project actually works.

A.1 The Role of Rotations

For the sake of completeness, I proof a central and well known result for sign identified SVARs (see Uhlig, 2005, Proposition A.1).

Proposition 1 *Suppose $A, \Sigma \in \mathbb{R}^{n\times n}$ are of full rank. Then $AA' = \Sigma$ if and only if $A = CQ$, where $C = chol(\Sigma)$ and $Q \in \mathcal{O}(n)$ is an arbitrary rotation matrix.*

Proof. Suppose $A = CQ$, then one can write

$$AA' = CQQ'C' = CC' = \Sigma. \tag{16}$$

Now let A satisfy $AA' = \Sigma$ and calculate

$$A = C\overbrace{C^{-1}A}^{\tilde{Q}:=}, \tag{17}$$

where $\tilde{Q} \in \mathcal{O}(n)$ because

$$\begin{aligned}\tilde{Q}\tilde{Q}' &= C^{-1}AA'(C^{-1})' \\ &= C^{-1}\Sigma(C^{-1})' \\ &= C^{-1}CC'(C')^{-1} = I_n,\end{aligned}$$

and similarly $\tilde{Q}'\tilde{Q} = I_n$. ■

This property enabled Arias et al. (2018) to use an alternative parametrization of the structural model (1), which they call the Orthogonal Reduced-Form Parametrization:

$$y_t = d + \sum_{j=1}^{p} B_j y_{t-j} + B_x x_t + CQ\epsilon_t, \quad (\epsilon_t) \overset{iid}{\sim} N(0, I_4), \tag{18}$$

which is convenient for sampling. When using this parametrization it is obvious that

$$e_t = CQ\epsilon_t \Leftrightarrow C^{-1}e_t = Q\epsilon_t, \tag{19}$$

where e_t is the reduced form error and $C^{-1}e_t$ is the 'structural' shock one would identify when using a recursive identification strategy. Thus one can always obtain a rotated version of the structural shocks using just observable moments, independent of the identifying assumptions one can place. This is intuitive when interpreting zero-mean random variables with finite variance as vectors in a Hilbert space, where the length of a vector is its standard deviation and the inner product of two vectors is their covariance. Viewed through this lense, identifying structural shocks amounts to finding *the* coordinate system (in terms of linear algebra *the* orthonormal basis) that spans the observed vectors $e_{t,i}$. But since there exist infinitely many such coordinate systems in the n-dimensional space, the desired set of vectors cannot be found without additional information. It is clear, however, that every linear function that maps one coordinate system into another must preserve the length of and the angle between all vectors – in other words,

it has to be a rotation. Therefore it is also geometrically intuitive that the identification of structural shocks can be broken down into two steps, orthonormalization and rotation, which amounts to multiplying the reduced form errors by C^{-1} and then Q.

A.2 Why Sub-Rotations Work

Consider now a setting in which one has some more information about the nature of $A_0^{-1} \in \mathbb{R}^{n\times n}$ in the structural model (1). In particular, it is known that this matrix has the following form:

$$A_0^{-1} = \begin{bmatrix} A_{11} & 0_{n_1\times n_2} & \cdots & 0_{n_1\times n_k} \\ A_{21} & A_{22} & \cdots & 0_{n_2\times n_k} \\ \vdots & \vdots & \ddots & \vdots \\ A_{k1} & A_{k2} & \cdots & A_{kk} \end{bmatrix}, \tag{20}$$

where $A_{ij} \in \mathbb{R}^{n_i\times n_j}$ and $n = n_1 + ... + n_k$. In other words, A_0^{-1} is block lower-triangular. Unless $n_1 = ... = n_k = 1$, in which case one can just apply the Cholesky identification scheme, it is still not possible to exactly identify the matrix A_0^{-1} or equivalently the structural shocks. However, the set of admissible models can be restricted, which is shown by the next Proposition:

Proposition 2 *Suppose $A, \Sigma \in \mathbb{R}^{n\times n}$ are of full rank. Further, A is block lower triangular, i.e. there exist $n_1 + ... + n_k = n$ such that A consists of k^2 block matrices $A_{ij} \in \mathbb{R}^{n_i\times n_j}$ with $A_{ij} = 0_{n_i\times n_j}$ for $j > i$. Then $AA' = \Sigma$ if and only if $A = CQ$, where $C = chol(\Sigma)$ and $Q = diag(Q_1, ..., Q_k)$ is block-diagonal with $Q_i \in \mathcal{O}(n_i)$.*

Proof. Following Proposition 1 it is clear that A must have the form

$$A = CQ, \tag{21}$$

with $Q \in \mathcal{O}(n)$. Since C is lower triangular, C^{-1} is lower triangular and in particular block lower triangular as defined above. Therefore $Q = C^{-1}A$ is the product of two block lower triangular matrices. Using the same partition as for A, one can compute each of the k^2 blocks of Q as

$$Q_{ij} = (C^{-1}A)_{ij} = \sum_{s=1}^{k} (C^{-1})_{is} A_{sj}. \tag{22}$$

Now suppose $j > i$. Then for $s = 1, ..., j-1$ the matrix A_{sj} consists only of zeros and for $s = j, ..., k$ the same is true for matrix $(C^{-1})_{is}$. Therefore in this case the right side of (22) is a sum of zero matrices and Q_{ij} consists of zeros only. Thus Q is also block lower triangular. As a next step I will prove that using the same partition, every orthonormal matrix Q that is block lower triangular is in fact block diagonal with orthonormal matrices as diagonal entries. For $k = 1$ this statement is trivial. Now suppose it holds for a certain k and take an arbitrary matrix $Q \in \mathcal{O}(n)$ with a partition corresponding to $n_1 + ... + n_{k+1} = n$ such that Q is block lower triangular. One can then partition this matrix into

$$Q = \begin{bmatrix} Q_{11} & 0_{n_1\times n_\bullet} \\ Q_{\bullet 1} & Q_{\bullet\bullet}, \end{bmatrix}, \tag{23}$$

where Q is of dimension $n_1 \times n_1$, $Q_{\bullet 1}$ is of dimension $n_\bullet \times n_1$ with $n_\bullet = n_2 + ... + n_{k+1}$ and $Q_{\bullet\bullet}$ is of dimension $n_\bullet \times n_\bullet$. Note that $Q_{\bullet\bullet}$ is block lower triangular with k^2 blocks. One can then use the fact that $Q \in \mathcal{O}(n)$:

$$QQ' = \begin{bmatrix} Q_{11} & 0_{n_1\times n_\bullet} \\ Q_{\bullet 1} & Q_{\bullet\bullet} \end{bmatrix} \begin{bmatrix} Q'_{11} & Q'_{\bullet 1} \\ 0_{n_\bullet\times n_1} & Q'_{\bullet\bullet} \end{bmatrix} = \begin{bmatrix} Q_{11}Q'_{11} & Q_{11}Q'_{\bullet 1} \\ Q_{\bullet 1}Q'_{11} & Q_{\bullet 1}Q'_{\bullet 1} + Q_{\bullet\bullet}Q'_{\bullet\bullet} \end{bmatrix} \overset{!}{=} \begin{bmatrix} I_{n_1} & 0_{n_1\times n_\bullet} \\ 0_{n_\bullet\times n_1} & I_{n_\bullet} \end{bmatrix}. \tag{24}$$

One can derive similar conditions from $Q'Q \stackrel{!}{=} I_n$. It follows immediately that $Q_{11} \in \mathcal{O}(n_1)$. Thus left-multiplying $Q_{11}Q'_{\bullet 1}$ by Q'_{11} yields $Q_{\bullet 1} = 0_{n_\bullet \times n_1}$ which implies $Q_{\bullet\bullet} \in \mathcal{O}(n_\bullet)$. Therefore $Q_{\bullet\bullet}$ is an orthonormal, block lower triangular matrix with k^2 blocks which, by induction assumption, is block diagonal with orthonormal matrices as diagonal entries. It follows that the statement holds also for $k+1$ and therefore for all $k \geq 1$, which finishes the proof. ■

Setting $n_1 = 1$, $n_2 = 3$ and $k = 2$ it follows from Proposition 2 that for the model of this paper it must hold that $A_0^{-1} = CQ$ with $C = chol(\Sigma)$ and Q being as given in (7). This result, however, also has a nice geometric interpretation. Let us re-write (19), using the special properties of Q in the case of A_0^{-1} being block lower-diagonal:

$$\begin{bmatrix} u_{t,m_1:n_1} \\ u_{t,m_2:n_2} \\ \vdots \\ u_{t,m_k:n} \end{bmatrix} := C^{-1}e_t = Q\epsilon_t = \begin{bmatrix} Q_1\epsilon_{t,m_1:n_1} \\ Q_2\epsilon_{t,m_2:n_2} \\ \vdots \\ Q_k\epsilon_{t,m_k:n} \end{bmatrix}, \tag{25}$$

where $m_s := n_1 + ... + (n_{s-1} + 1)$, $m_1 := 1$ and $\epsilon'_{t,u:v} := [\epsilon_u \quad \epsilon_{u+1} \quad ... \quad \epsilon_v,]$ for $v > u$. Note that to obtain the true structural shocks ϵ_t one does not have to rotate $C^{-1}e_t$ in the whole n-dimensional space any more. Since $u_{t,m_j:n_j}$ only consists of $n_j \leq n$ of the structural shocks one has already identified sub-spaces of structural shocks and the rotation Q is restricted to not map between these sub-spaces, which is represented by the zero-blocks off the diagonal of Q. Note that in the case of $n_i = 1$ for a block $i \leq k$ it follows from (25) that $\epsilon_{t,i}$ is a scaled version of $u_{t,i}$ and therefore this shock is point-identified up to sign, which is the case with the drilling shock in this paper's model. Figure 10 gives a geometric example of the role that sub-rotations play in narrowing down the set of admissible rotations: Since the first axis of the coordinate system is identified through zero restrictions, rotations are not allowed to move it. This severely restricts the directions in which the system can be rotated. In summary, if one can plausibly impose block lower-triangularity on A_0^{-1}, sub-rotations allow one to seperate the space of all structural shocks into multiple smaller blocks of structural shocks, for which strategies such as sign or elasticity restrictions can be used to further narrow down the set of admissible rotations Q.

Appendix B. Derivation of the IRF Posterior Density

In the case of classical sign restrictions, Inoue and Kilian (2013, 2018) show that in a Bayesian framework one can explicitly derive the posterior density of the structural impulse response functions Θ as

$$f(\Theta) = \left| \frac{\partial vec(\Theta)}{\partial [vec(B)' \; vech(C)' \; s']} \right|^{-1} \left| \frac{\partial vech(\Sigma)}{\partial vech(C)} \right| f(B|\Sigma) f(\Sigma) f(s), \tag{26}$$

where $S = I_n - 2(I_n + \tilde{Q})^{-1}, s = veck(S)$, $C = \text{chol}(\Sigma)$ and the operators vec, $vech$ and $veck$ stack the columns, lower triangular and sub-diagonal elements of a matrix, respectively. $\tilde{Q}$ is Q manipulated to satisfy $|\tilde{Q}| = 1$. For the remainder of the section, when I write Q I will refer to $\tilde{Q}$. The matrix B is defined as $[B_1, ..., B_p]$. León et al. (2006) show that for a uniform distribution on $\{Q \in \mathcal{O}(n) || Q| = 1\}$ the variable s has the density function

$$f(s) = \left(\prod_{i=2}^{n} \frac{\Gamma(i/2)}{\pi^{i/2}} \right) \frac{2^{(n-1)(n-2)/2}}{|I_n + S|^{n-1}}. \tag{27}$$

The specific form of the density in Inoue and Kilian (2013), however, depends on the change-of-variable method which holds because there is a one-to-one mapping between $[vec(B)', vech(C)', s']$ and the first $p+1$ structural IRFs, $\Theta = [\theta_0, \theta_1, ..., \theta_p] = [A_0^{-1}, \Phi_1 A_0^{-1}, ..., \Phi_p A_0^{-1}]$, where Φ_i are

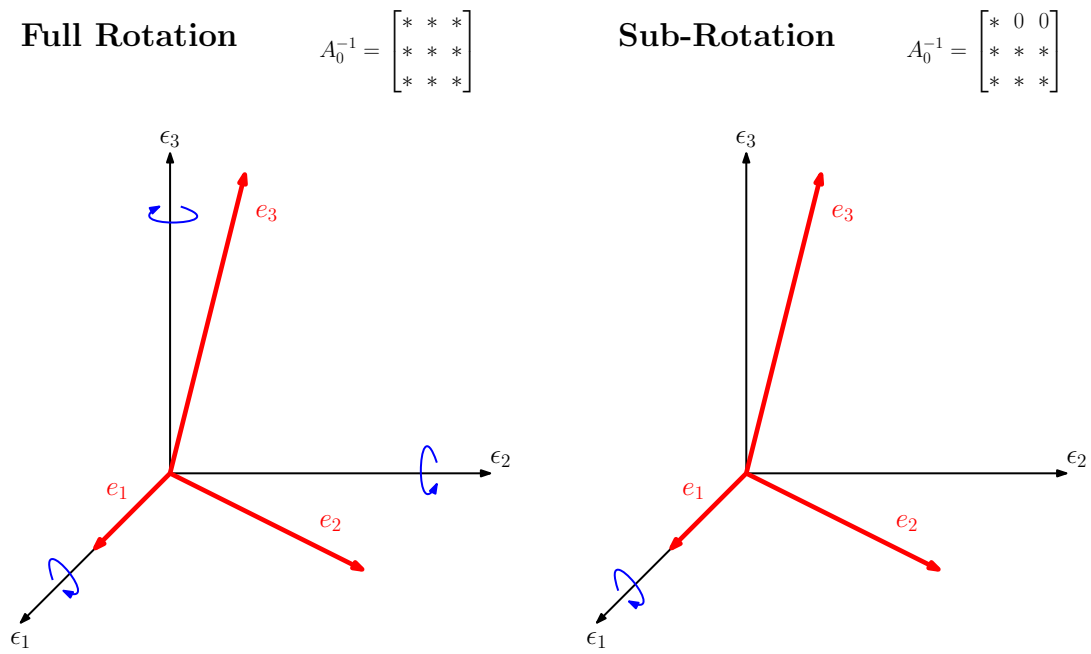


Figure 10: Difference between Rotations and Sub-Rotations.

Notes: The plotted coordinated system is the set of structural shocks obtained using the 'Cholesky-strategy'. Possible rotations of the coordinate system that are allowed under the identification assumptions are drawn on the axes.

the reduced form impulse responses at horizon i. In the case of sub-rotations, however, one has to adapt the formula above. This is because one does not draw from $\mathcal{O}(n)$ but from smaller dimensional subspaces $\mathcal{O}(n_i)$. One can, however, derive the posterior density of the IRFs as

$$f(\Theta) = \overbrace{\left|\frac{\partial vec(\tilde{\Theta})}{\partial[vec(B)'\ vech(C)'\ (s^*)']}\right|^{-1}}^{I} \overbrace{\left|\frac{\partial vech(\Sigma)}{\partial vech(C)}\right|}^{II} \overbrace{f(B|\Sigma)}^{III} \overbrace{f(\Sigma)}^{IV} \overbrace{f(s^*)}^{V}. \tag{28}$$

where $s^* = veck(S^*)$ and $S^* = I_{n-1} - 2(I_{n-1} + Q^*)^{-1}$. The matrix Q^* is the one used to construct Q like in (7). Terms II-IV are exactly like in Inoue and Kilian (2018), term V is given as (27), where one just has to substitute S^* for S and $n-1$ for n. Only deriving I needs some tedious computations. Before doing that, however, I should introduce some notation: Let D_n be the $n^2 \times n(n+1)/2$ duplication matrix such that $vec(M) = D_n vech(M)$ for any $n \times n$ symmetric matrix M. Further, $\tilde{D}_{n-1}$ is the $(n-1)^2 \times (n-1)(n-2)/2$ matrix designed to satisfy $vec(S^*) = \tilde{D}_{n-1} s^*$. Similarly L_n is $n(n+1)/2 \times n^2$ such that $vec(M) = L_n' vech(M)$ for any lower triangular matrix M. The $n \times 1$ vector e_i^n consists of zeros with only the ith element being equal to one, c_{ii} is the i-th diagonal element of C. Lastly, let us define the $n \times (n-1)$ matrix K as

$$K = \begin{bmatrix} 0_{1\times(n-1)} \\ I_{n-1} \end{bmatrix}. \tag{29}$$

Now I can proof the following Proposition:

Proposition 3 *Suppose an SVAR-model with n endogenous variables and n structural shocks, where only the first of n structural shock is allowed to have an immediate impact on the first variable and one draws from an uniform distribution on $\mathcal{O}(n-1)$ to identify the model via sub-rotations as outlined above. Then the joint posterior density of the structural impulse response*

functions can be computed as

$$f(\Theta) = 2^{(n^2-n+2)/2}|C|^{-np}\prod_{i=1}^{n} c_{ii}^{n-i+1}|I_{n-1}+Q^*|^{-(n-2)}$$
$$|\tilde{D}'_{n-1}(K'\otimes K'C')(I_{n^2}-L'_nL_n)(K\otimes CK)\tilde{D}_{n-1}|^{-1/2}f(B|\Sigma)f(\Sigma)f(s^*). \quad (30)$$

Proof. As first and major step, one has to take the derivative of $vec(\Theta)$ with respect to $vec(\Phi)'$, $vech(C)'$ and s'. Note that $Q^* = 2(I_{n-1}-S^*)^{-1} - I_{n-1}$ and therefore

$$dQ^* = 2(I_{n-1}-S^*)^{-1}dS(I_{n-1}-S^*)^{-1}. \quad (31)$$

It follows that

$$\begin{aligned} vec(dQ^*) &= 2(((I_{n-1}-S^*)^{-1})'\otimes(I_{n-1}-S^*)^{-1})vec(dS) \\ &= 2(((I_{n-1}-S^*)^{-1})'\otimes(I_{n-1}-S^*)^{-1})\tilde{D}_{n-1}ds. \end{aligned} \quad (32)$$

Now comes the step where I depart from Inoue and Kilian (2018). One can express Q as

$$Q = e_1^n(e_1^n)' + KQ^*K'. \quad (33)$$

It follows that

$$\begin{aligned} d(CQ) &= d(C)Q + Cd(KQ^*K') \\ &= d(C)Q + 2CK(I_{n-1}-S^*)^{-1}dS^*(I_{n-1}-S^*)^{-1}K' \quad (34) \\ d(\Phi CQ) &= d(\Phi)CQ + \Phi d(C)Q + 2\Phi CK(I_{n-1}-S^*)^{-1}dS^*(I_{n-1}-S^*)^{-1}K', \quad (35) \end{aligned}$$

and further

$$\begin{aligned} d(vec(CQ)) &= (Q'\otimes I_n)L'_n d(vech(C)) + 2(I_n\otimes C)vec(K(I_{n-1}-S^*)^{-1}dS^*(I_{n-1}-S^*)^{-1}K') \\ &= (Q'\otimes I_n)L'_n d(vech(C)) + 2(I_n\otimes C)J_Q ds^* \quad (36) \\ d\ vec(\Phi CQ) &= (Q'C'\otimes I_{np})d(vec(\Phi)) + (Q'\otimes\Phi)L'_n d(vech(C)) + (I_n\otimes\Phi C)J_Q ds^*, \quad (37) \end{aligned}$$

where $J_Q := (K((I_{n-1}-S^*)^{-1})'\otimes K(I_{n-1}-S^*)^{-1})\tilde{D}_{n-1}$. Note that the final results for the differentials are just as in Inoue and Kilian (2018), with the only difference being the definitions of J_Q and s^*. While in Inoue and Kilian (2018) J_Q is a square matrix in our case it is not, since one has to match the different dimensions of S^* and C. This difference in dimensions will remain the major challenge in the following proof.

Remembering that $\Theta = [(CQ)'\ (\Phi CQ)']'$ it follows immediately from (36) and (37) that

$$J_1 := \frac{\partial vec(\Theta)}{\partial[vech(C)'\ (s^*)'\ vec(\Phi)]} = \begin{bmatrix} (Q'\otimes I_n)L'_n & (I_n\otimes C)J_Q & 0_{n^2\times n^2p} \\ (Q'\otimes\Phi)L'_n & (I_n\otimes\Phi C)J_Q & (Q'C'\otimes I_{np}) \end{bmatrix}. \quad (38)$$

Because this matrix is block lower triangular its determinant is equal to the product of the determinants of the matrices on the diagonal:

$$\begin{aligned} |J_1| &= |(Q'\otimes I_n)L'_n\ (I_n\otimes C)J_Q||(Q'C'\otimes I_{np})| \\ &= |(Q'\otimes I_n)L'_n\ (I_n\otimes C)J_Q||C|^{np}. \quad (39) \end{aligned}$$

Now one can re-write

$$[(Q'\otimes I_n)L'_n\ (I_n\otimes C)J_Q] = (Q'\otimes I_n)Z, \quad (40)$$

with

$$Z = [L_n' \; 2(F \otimes G)\tilde{D}_{n-1}], \tag{41}$$

$$F = QK(I_{n-1} - (S^*)')^{-1} = \frac{1}{2}K(I_{n-1} + Q^*) = K(I_{n-1} - S^*)^{-1}, \tag{42}$$

$$G = CK(I_{n-1} - S^*)^{-1} = CF. \tag{43}$$

As a next step, the following equality can be used:

$$|(Q' \otimes I_n)Z| = |Z| = \sqrt{|Z|^2} = \sqrt{|Z'Z|}. \tag{44}$$

Because of

$$\begin{aligned} Z'Z &= \begin{bmatrix} L_n \\ 2\tilde{D}_{n-1}'(F' \otimes G') \end{bmatrix} \begin{bmatrix} L_n' & 2(F \otimes G)\tilde{D}_{n-1} \end{bmatrix} \\ &= \begin{bmatrix} I_{n(n+1)/2} & 2L_n(F \otimes G)\tilde{D}_{n-1} \\ 2\tilde{D}_{n-1}'(F' \otimes G')L_n' & 4\tilde{D}_{n-1}'(F' \otimes G')(F \otimes G)\tilde{D}_{n-1} \end{bmatrix} \\ &=: \begin{bmatrix} \mathcal{A} & \mathcal{B} \\ \mathcal{C} & \mathcal{D} \end{bmatrix} \end{aligned} \tag{45}$$

the squared determinant of Z can be computed as

$$\begin{aligned} |Z'Z| &= |\mathcal{A}||\mathcal{D} - \mathcal{C}\mathcal{A}^{-1}\mathcal{B}| \\ &= |\mathcal{D} - \mathcal{C}\mathcal{B}| \\ &= |4(\tilde{D}_{n-1}'(F' \otimes G')(I_{n^2} - L_n'L_n)(F \otimes G)\tilde{D}_{n-1})| \\ &= 2^{(n-1)(n-2)}|\tilde{D}_{n-1}'(F' \otimes G')(I_{n^2} - L_n'L_n)(F \otimes G)\tilde{D}_{n-1}|. \end{aligned} \tag{46}$$

Now, I define $F^* = (I_{n-1} - S^*)^{-1}$ and $\tilde{D}_{n-1}^+ = (\tilde{D}_{n-1}'\tilde{D}_{n-1})^{-1}\tilde{D}_{n-1}'$. Using a result from Magnus (1988) one can write

$$\begin{aligned} &\tilde{D}_{n-1}'(F' \otimes G')(I_{n^2} - L_n'L_n)(F \otimes G)\tilde{D}_{n-1}) \\ =&\tilde{D}_{n-1}'((F^*)' \otimes (F^*)')(K' \otimes K')(I_n \otimes C')(I_{n^2} - L_n'L_n)(I_n \otimes C)(K \otimes K)(F^* \otimes F^*)\tilde{D}_{n-1} \\ =&\mathcal{M}'\tilde{D}_{n-1}'(K' \otimes K')(I_n \otimes C')(I_{n^2} - L_n'L_n)(I_n \otimes C)(K \otimes K)\tilde{D}_{n-1}\mathcal{M}, \end{aligned} \tag{47}$$

where $\mathcal{M} = (\tilde{D}_{n-1}^+(F^* \otimes F^*)\tilde{D}_{n-1})$. Again, using Magnus (1988) this leads to

$$|\mathcal{M}| = |F^*|^{n-2} = |\frac{1}{2}I_{n-1} + Q^*|^{n-2} = 2^{-(n-2)(n-1)}|I_{n-1} + Q^*|^{n-2}. \tag{48}$$

Using (39), (40), (44), (46), (47) and (48) one obtains

$$\begin{aligned} |J_1| &= |C|^{np}|(Q' \otimes I_n)L_n' \; (I_n \otimes C)J_Q| \\ &= |C|^{np}|(Q' \otimes I_n)Z| \\ &= |C|^{np}|Z| \\ &= |C|^{np}2^{(n-1)(n-2)/2}|\tilde{D}_{n-1}'(F' \otimes G')(I_{n^2} - L_n'L_n)(F \otimes G)\tilde{D}_{n-1}|^{1/2} \\ &= |C|^{np}2^{(n-1)(n-2)/2}|\mathcal{M}||\tilde{D}_{n-1}'(K' \otimes K'C')(I_{n^2} - L_n'L_n)(K \otimes CK)\tilde{D}_{n-1}|^{1/2} \\ &= 2^{-(n-1)(n-2)/2}|C|^{np}|I_{n-1} + Q^*|^{n-2}|\tilde{D}_{n-1}'(K' \otimes K'C')(I_{n^2} - L_n'L_n)(K \otimes CK)\tilde{D}_{n-1}|^{1/2}. \end{aligned} \tag{49}$$

For the remaining derivatives the results from Inoue and Kilian (2018) can be used:

$$|J_2| := \left| \frac{\partial vec(\Phi)}{\partial vec(B)'} \right| = 1, \tag{50}$$

$$|J_3| := \left| \frac{\partial vech(\Sigma)}{\partial vech(C)'} \right| = 2^n \prod_{i=1}^{n} c_{ii}^{n-i+1}. \tag{51}$$

As a last step, (26) can be re-written as

$$\begin{aligned} f(\Theta) &= |J_1|^{-1}|J_2||J_3| f(B|\Sigma) f(\Sigma) f(s^*) \\ &= 2^{(n^2-n+2)/2} |C|^{-np} |I_{n-1} + Q^*|^{-(n-2)} \prod_{i=1}^{n} c_{ii}^{n-i+1} \\ &\quad |\tilde{D}'_{n-1}(K' \otimes K'C')(I_{n^2} - L'_n L_n)(K \otimes CK)\tilde{D}_{n-1}|^{-1/2} f(B|\Sigma) f(\Sigma) f(s^*). \end{aligned}$$

■