Statistisk sentralbyrå
Statistics Norway

U.S. tight oil supply flexibility - A multivariate dynamic model for production and rig activity

TALL SOM FORTELLER

Halvor Briseid Storrøsten

DISCUSSION PAPERS

997

Discussion Papers No. 997, February 2023
Statistics Norway, Research Department

Halvor Briseid Storrøsten

U.S. tight oil supply flexibility - A multivariate dynamic model for production and rig activity

Abstract:
This paper examines the supply of U.S. LTO from both a theoretical and empirical point of view. The theory model combines endogenous rig activity and stylized reservoir pressure mechanics with the classic Hotelling model for exhaustible resource extraction. The empirical section presents a vector error correction model for U.S. LTO production. Both models allow for simultaneous modeling of U.S. LTO supply and rig activity. A one percent shock to the oil price is estimated to increase LTO supply and rig activity with 0.3 and 0.8 percent, respectively. A one percent increase in rig activity leads to a 1.7 percent increase in oil production, but also a 0.1 percent increase in costs.

Keywords: Oil supply, rig activity, elasticity, tight oil, shale oil, vector error correction models.

JEL classification: Q3, Q4, L71, C32.

Acknowledgements: I am grateful to Terje Skjerpen, Thomas S. Gundersen, Lars Lindholt and Knut Einar Rosendahl for valuable comments to earlier drafts.

Address: Halvor Briseid Storrøsten, Statistics Norway, Research Department. E-mail: hbs@ssb.no

Discussion Papers comprise research papers intended for international journals or books. A preprint of a Discussion Paper may be longer and more elaborate than a standard journal article, as it may include intermediate calculations and background material etc.

The Discussion Papers series presents results from ongoing research projects and other research and analysis by SSB staff. The views and conclusions in this document are those of the authors

Abstracts with downloadable Discussion Papers
in PDF are available on the Internet:
https://www.ssb.no/discussion-papers
http://ideas.repec.org/s/ssb/dispap.html

ISSN 1892-753X (electronic)

Sammendrag

Siden 2010 har USA mer enn doblet sin andel av global oljeproduksjon. Årsaken er den kraftige veksten i amerikansk skiferoljeproduksjon, som har hatt stor innflytelse på hvordan det globale oljemarkedet fungerer. Denne artikkelen handler om tilbudet av amerikansk skiferolje og det tilhørende riggmarkedet.

Først presenteres en teorimodell som analyserer riggaktivitet og oljeproduksjon innenfor den klassiske Hotelling-modellen for utvinning av ikke-fornybare ressurser. En viktig konklusjon fra teoridelen er at, i kontrast med konvensjonell oljeproduksjon, kan skiferoljeproduksjon respondere på endringer i oljeprisen også på kort sikt. Forskjellen skyldes i hovedsak ulik produksjonsteknologi.

Deretter følger en økonometrisk dynamisk multivariat modell (VECM) hvor amerikansk skiferoljeproduksjon og riggaktivitet er sentrale endogene variabler. En oljeprisøkning på én prosent anslås å øke amerikansk skiferoljeproduksjon og riggaktivitet med henholdsvis 0,3 og 0,8 prosent. Videre medfører en økning på én prosent i riggaktiviteten at oljeproduksjonen øker med 1,7 prosent, mens produksjonskostnadene øker med 0,1 prosent

1 Introduction

The boom in United States (U.S.) light tight oil (LTO) production during the last decade or so has been touted by many as a game changer with potentially wide reaching consequences for the global oil market.[1] Since 2010, the U.S. has more than doubled its share of global oil production, from 9.1 percent in 2010 to 18.6 percent in 2020. According to the U.S. Energy Information Administration (EIA), about 7.76 million barrels per day of crude oil were produced directly from U.S. tight oil resources in 2019. This averages up to 2.83 billion barrels, or 63 percent of total U.S. crude oil production, in 2019.[2] In comparison, LTO accounted for 15 percent of US crude oil production in 2010. U.S. crude supply and its share of world oil production are graphed in Figure 1.

LTO is very light (API 45-50) and sweet (< 0.1 percent sulfur) crude oil produced from low permeability formations such as shale or tight sandstone.[3] LTO extraction requires hydraulic fracturing (fracking) and typically involves the same horizontal well technology as used in, e.g., production of shale gas. In contrast to most conventional oil production, tight oil production declines fast with the dominating part of cumulative production occurring within the first few years after investment. This production profile suggests that LTO supply may be more responsive to the oil price than oil from conventional wells.

This paper investigates U.S. supply of LTO using a combination of economic theory and econometrics. I first present the theoretical model for LTO production. The theory model combines endogenous rig activity and stylized reservoir pressure mechanics with the classic Hotelling model for exhaustible resource extraction. A key model prediction is that oil supply does not respond to changes in the oil price in the short run if reservoir and cost structures are similar to those typical for conventional petroleum extraction. This is consistent with the empirical literature on short-run conventional oil supply price elasticities (see, e.g., Pesaran, 1990; Dahl and Yücel, 1991; Ramcharran, 2002; Smith, 2009; Anderson et al. 2018; Kilian, 2020). For cost structures similar to LTO production, however, higher oil prices may very well increase both current and future oil production. This is consistent with the empirical results in the present paper, and the small but growing empirical literature on LTO production (see below). The theory model also suggests that economy wide capacity constraints may dampen the response in oil production and rig activity spurred by higher oil prices.[4]

The empirical section presents a vector error correction model (VECM)

[1]See, e.g., Fattouh and Sen (2013), and Wethe (2019)

[2]In the pandemic year 2021, U.S. LTO production was 7.28 million barrels per day, which accounted for 65 percent of total U.S crude oil production.

[3]LTO should not be confused with 'oil shale', which is shale rich in kerogen. Also, the term 'LTO' is broader than the term 'shale oil', because LTO can be extracted from not

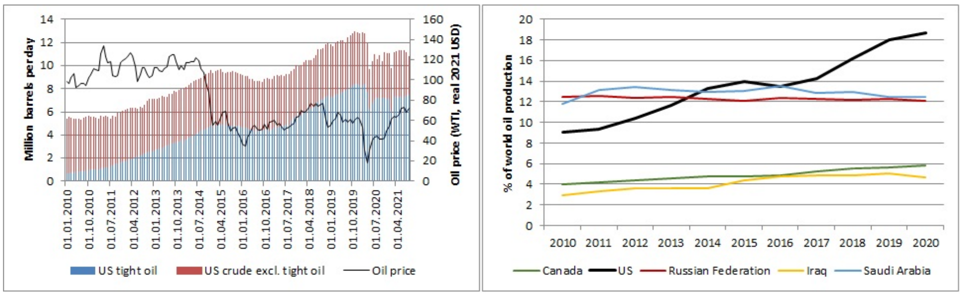

Figure 1: Left: U.S. oil production and the West Texas Intermediate oil price. Right: Percentages of world oil production from the five largest oil producing countries in 2020 (sources: EIA, BP Statistics and own calculations).

for U.S. LTO production using monthly data over the period Jan. 2010 to Apr. 2022. The (endogenous) variables included are U.S. LTO production, rig activity associated with U.S. LTO, the West Texas Intermediate (WTI) oil price, and a proxy for tight oil production costs. I find that a one percent positive transitory shock to the oil price causes a significant gradual increase in LTO supply, which stabilizes around 0.3 percent after around two years.[5] The VECM captures how the oil price affects rig activity, and how this in turn affects oil production. This may be even more important in the case of tight oil than for conventional oil, due to the large number of wells that must be drilled for this type of production. Indeed, around 1/3 of the world drilling rigs are working with tight oil in the U.S.[6] The results indicate, perhaps not surprisingly, that the rig market responds stronger to changes in oil prices than LTO production itself. The increase in rig activity following a transitory shock to the oil price tops out at 1.2 percent after one year, before it drops and stabilizes around 0.8 percent in the longer run. Further, a one percent positive shock to the number of active rigs leads to a 1.7 percent increase in oil production, and a 0.1 percent increase in costs. I find no significant effect from the U.S. LTO on the West Texas Intermediate (WTI) oil price, however.

just shale formations but also from sandstone and carbonates.

[4]The modern-day gold rush of oil companies and contractors converging on western Canada's oil-sands markets bogged down as high materials costs and outstripped labor resources forced project delays and budget overruns around the year 2007, see ENR. Osmundsen et al. (2010) find that oil well drilling speed tends to be negatively correlated with capacity utilization, due to capacity bottlenecks and lower drilling quality. Further, Osmundsen et al. (2015) and Skjerpen et al. (2018) find that increased capacity utilization in the rig market increases the rig rates and, hence, the cost of capacity construction in the Gulf of Mexico and on the Norwegian continental shelf, respectively.

[5]The effects are derived from the impulse response functions, see Figures 5 and 7 for details. Note that the transitory shock to the oil price causes a more than one percent increase in the oil price in the subsequent periods in the VECM, cf., Figure 6.

[6]This figure is based on Baker Hughes Rig Count for Nov. 2019 and does not include former Soviet Union and onshore China.

The theoretical analysis is based on Anderson et al. (2018), but expands their model with fixed and variable extraction costs, economy wide capacity constraints and technological progress, all of which are arguably particularly relevant for unconventional petroleum production. Notably, the model in Anderson et al. (2018) does not predict oil supply to respond to the oil price in the short run. The model in the present paper replicates that result for cost structures reasonable for conventional oil fields, but finds that unconventional and LTO production may respond to the oil price also in the short run.

Anderson et al. (2018) also examine drilling and (conventional) oil production in Texas for 1990-2007 empirically, using monthly and quarterly time series data for Texas over the period 1990-2007. They find that whereas production from existing wells does not respond to prices, the oil price elasticity of drilling is approximately 0.7. Newell and Prest (2019) examine the supply-side elasticity of drilling and oil production for conventional and unconventional production and drilling in Texas, North Dakota, California, Oklahoma and Colorado over the period 2000 to 2015. Their simulations, based on estimates derived from micro data, show that the U.S. supply responsiveness has increased substantially due to the shale revolution. They also find that, given a price of 80 USD per barrel, U.S. production could rise by 0.5 million barrels per day in 6 months, 1.2 million in 1 year, 2 million in 2 years, and 3 million in 5 years (prices adjusted for inflation to 2014 dollars). Bjørnland et al. (2020) consider a well-level monthly production data set covering more than 16,000 crude oil wells in North Dakota. They find the short run supply elasticity of shale wells to be in the range of 0.3–0.9, depending on well and firm characteristics. They find no such responses for conventional wells. Gundersen (2020) examines the role of the U.S. shale oil boom in driving global oil prices, using a structural vector autoregressive model that identifies separate oil supply shocks for the U.S. and OPEC. He finds that U.S. supply shocks can account for up to 13 percent of the oil price variation over the 2003-2015 period. Aastveit et al. (2022) find that shale oil producers respond positively and significantly to favourable oil price signals, and that the response is heterogenous across various shale wells. Vatter et al. (2022) model impacts on oil production of price, capital costs, technological progress, well-to-well interference of closely situated wells, and location. They find oil from Bakken to be more responsive to changes in the oil price than that of non-OPEC oil supply in general, and argue that the price response of shale oil tends to dampen the long-term price cycle and moderate the price shocks in the oil market. Kilian (2016) investigates the impact of the shale oil revolution on U.S. crude oil and gasoline prices. Kilian (2017) examines how the shale boom affected U.S. oil imports, Arab oil exports, and the global oil price. The results indicate that U.S. shale has a negative impact on oil prices, U.S. oil imports and Arab oil exports in general. Balke et al. (2020) estimate a dynamic, structural model of the world oil market in order to quantify the impact of the shale revolution. They find that oil prices in 2018 would have been roughly 36 percent higher had the shale revolution not occurred, and that the shale revolution implies a reduction in current and long-run oil price volatility of around 25 percent and over 50 percent, respectively. Bornstein et al. (2018) use micro data to compile some key facts about the oil market and estimate a structural industry equilibrium model that is consistent with these facts. Perhaps most relevant to the present paper, their model predicts that the advent of fracking will reduce oil price volatility. Kleinberg et al. (2018) discuss LTO development economics and breakeven points, and why they are often misunderstood. Last, Foroni and Stracca (2022) formulate a structural VAR model of the oil market and find that the shale oil boom has not fundamentally changed global oil supply, which remains close to vertical with a significant estimated short-run price elasticity around 0.05.

The present paper is, to the author's best knowledge, the first to construct a VECM model for U.S. LTO production and rig activity. This puts it apart, e.g. from the panel data analyses of U.S. LTO cited above, by providing long run oil price responses of U.S. LTO production. Perhaps most importantly, the VECM framework allows for simultaneous modeling of U.S. LTO supply and rig activity. Whereas the theoretical analysis is based on Anderson et al. (2018), the extensions including fixed and variable extraction costs, economy wide capacity constraints and technological growth, are all arguably particularly relevant for unconventional petroleum production.

2 Theoretical analysis

Let there be $i \in I = \{1, 2, ..., \bar{i}\}$ oil producing price-taking firms. The remaining *undeveloped* resource stock available to firm i at time $t \in T$ is given by $S_{it} = \bar{S}_i + \int_{t_0}^{t}(d_{it} - x_{it})dt$, where $\bar{S}_i$ (an exogenous constant) is the initial resource stock at time $t = t_0$, d_{it} is an exogenous increase in resources available for development, x_{it} is firm i's field development at time t, and the model begins at time t_0. The stock S_{it} denotes oil and gas trapped in reservoir rocks unavailable for extraction before field development has taken place. New undeveloped resources, $d_{it} \geq 0$, reflects, e.g., new areas opened for petroleum activity or technological change that allows for exploitation of resources not previously technically or economically recoverable.[7] Differentiating with respect to time, we get the state movement equation for the remaining undeveloped resource stock:

$$\dot{S}_{it} = -x_{it} + d_{it}, \tag{1}$$

where $\dot{S}_{it} \equiv \partial S_{it}/\partial t$ denotes the rate of change of the remaining undeveloped resource stock S_{it} with respect to time. I assume that new resources, d_{it}, if

[7]The model does not feature endogenous exploration of new resources.

positive, are sufficiently small to retain resource scarcity.

The *developed* reserve R_{it} refers to the resource that is available for extraction for firm i at time t. It is given by $R_{it} = \bar{R}_i + \int_{t_0}^{t}(x_{it} - q_{it})dt$, where $\bar{R}_i$ (an exogenous constant) is the initial developed resource stock at time $t = t_0$ and q_{it} is firm i's oil extraction.[8] Differentiating with respect to time we get the state movement equation for the developed resource stock:

$$\dot{R}_{it} = x_{it} - q_{it}. \tag{2}$$

I assume that field development is costly and that the resource that is cheapest to develop is developed first. Hence, field development costs decrease in the remaining resource stock S_{it}. The cost of field development is given by the function $c_i^x(x_{it}, S_{it}, k_t)$ with $c_i^x(0, S_{it}, k_t) = 0$ and derivatives satisfying $\partial c_i^x(\cdot)/\partial x_{it} \equiv c_{x_{it}}^x(\cdot) > 0$, $c_{S_{it}}^x(\cdot) < 0$, $c_{k_{it}}^x(\cdot) > 0$, $c_{x_{it}x_{it}}^x(\cdot) > 0$, $c_{S_{it}S_{it}}^x(\cdot) > 0$, $c_{k_tk_t}^x(\cdot) \geq 0$, $c_{x_{it}S_{it}}^x(\cdot) < 0$, $c_{k_tx_{it}}^x(\cdot) > 0$ and $c_{k_tS_{it}}^x(\cdot) \geq 0$. The variable k_t is a catch-all variable capturing other things that affect production costs, e.g., available technology and the cost of labor and equipment. I define k_t such that costs increase in k_t.

The maximum flow of oil from developed fields depends on the pressure in the well which, everything else equal, decreases as the resource is depleted. Following Anderson et al. (2018), I will assume that the maximum flow is proportional with a factor ω to the amount of oil that remains underground. The oil producer can adjust the oil production using the flow control $y_{it} \in [0, 1]$. For example, the flow of oil may be increased by injecting gas or water into the reservoir to replace produced fluids, and thus maintain or increase the reservoir pressure. Similarly, production of LTO must be stimulated using hydraulic fracturing to create sufficient permeability to allow the mature oil and/or natural gas liquids to flow at economic rates.[9]

Production of oil is given by:[10]

$$q_{it} = \omega R_{it} y_{it}. \tag{3}$$

The operating cost of oil extraction is given by $c_i^y(y_{it}, k_t)$, where $c_i^y(\cdot)$ is convex, increasing in the flow rate y_{it}, and increasing in the catch-all cost

[8]Petroleum field development involves issues like reservoir and production engineering, construction of infrastructure and surface facilities, well design and construction, completion design, environmental impact and risk assessment, and so forth.

[9]About 25 percent of the shale wells in the sample examined by Aastveit et al. (2022) have been refractured at least once.

[10]I abstract from the fact that most LTO wells produce a mix of oil and gas. Equation (3) is a reasonable approximation only for reservoirs where pressure is an important determinant of production, e.g. conventional oil and gas worldwide, LTO in the U.S., in situ extraction of bitumen in the oil sands of Alberta (Canada) and extra heavy oil in the Orinoco belt (Venezuela). The mining of shallow reserves of bitumen in Alberta is not adequately modeled by Equation (3).

variable k_t.[11] Specifically, we have $c^y_{y_{it}}(\cdot) > 0$, $c^y_{k_{it}}(\cdot) > 0$, $c^y_{y_{it}y_{it}}(\cdot) > 0$, $c^y_{k_{it}k_{it}}(\cdot) \geq 0$ and $c^y_{y_{it}k_{it}}(\cdot) \geq 0$. Last, extraction cost is zero when there is no extraction, $c^y_i(0, k_t) = 0$.

Petroleum extraction involves fixed costs that do not depend on the day-to-day production and drilling rates. Examples of such costs may be long-term contracts for hire of skilled labor or rental equipment, maintenance costs, and costs of regulatory compliance. These costs, denoted $c^f_i(y_{it}, x_{it})$, are incurred if and only if extraction is positive. I assume the fixed operating costs are twice differentiable and increasing in both arguments. Further, it satisfies $c^f_i(0, 0) = 0$, $c_i(y_{it} > 0, 0) = f^y_i$, $c_i(0, x_{it} > 0) = f^x_i$, $c_i(y_{it} > 0, x_{it} > 0) = f^{xy}_i$, with $f^{xy}_i > f^y_i$, $f^{xy}_i > f^x_i$, $f^x_i > 0$ and $f^y_i > 0$.[12]

We have the following market equilibrium relations, which by assumption are not internalized by the competitive individual firms:

$$p_t = p(\sum_{i \in I} q_{it}, \nu_t), \tag{4}$$

$$k_t = k(\sum_{i \in I} x_{it}, \sum_{i \in I} y_{it}, \kappa_t), \tag{5}$$

Equation (4) is the inverse (residual) demand function for the homogeneous oil the $i \in I$ firms produce. It gives the equilibrium price as a function of the aggregate quantity produced, and a catch-all variable that affects residual demand, denoted ν_t (e.g., gross national product, or production by other producers that are not members of the set I). I assume that there exists a choke price $\bar{p}$, such that demand is zero if $p_t \geq \bar{p}$.[13] The equilibrium price is convex and decreasing in aggregate production with derivatives $\partial p(\cdot)/\partial \sum_{i \in I} q_{it} \equiv p_{q_t} \leq 0$, $p_{q_t q_t} \geq 0$, $p_\nu > 0$ (by definition of ν) and $p_{\nu\nu} \geq 0$. The model allows for $p_{q_t} = 0$, which is approximately true if the set of firms I constitutes a sufficiently small part of global world oil supply. Equation (5) states that the catch-all cost variable k_t may increase in aggregate production or resource development, e.g. due to economy-wide capacity constraints like infrastructure limitations, refinery capacity, or shortage of skilled labour or equipment.[14] This is captured by

[11] I abstract from the fact that field development and oil extraction may depend on different exogenous cost variables k_t (the assumption does not affect the results in any relevant way).

[12] I assume that $c^f_i(x_{it}, y_{it})$ is differentiable so that the cost function $c_i(\cdot)$ is differentiable also in the presence of fixed costs. One example of such a fixed cost function, based on the cumulative Cauchy distribution, is $c^f_i(x_{it}, y_{it}) = (f^x_i/\pi)arctan((x_{it} - x_l)/f_l) + (f^y_i/\pi)arctan((y_{it} - y_l)/f_l) + 1$, where x_l, y_l and f_l are very small numbers, e.g., x_l and y_l are one barrel of oil equivalent and $f^l = 0.0001$, and f^y_i and f^x_i are the fixed costs of production and resource development, respectively.

[13] While $\bar{p}$ prevents the price from going towards infinity, so that the integral of the object function V_i in (6) is not infinite, $\bar{p}$ can be so high that it has no practical significance.

[14] Osmundsen et al. (2010) find that oil well drilling speed tends to be negatively cor-

$\partial k(\cdot)/\partial \sum_{i \in I} x_{it} \equiv k_{x_t} \geq 0$ and $\partial k(\cdot)/\partial \sum_{i \in I} y_{it} \equiv k_{y_t} \geq 0$. The exogenous variable κ_t is a catch-all variable that affects the cost variable k_t, e.g., technology or environmental regulation stringency. I assume $k_{\kappa_t} < 0$. For example, technological progress, or increased refinery capacity geared towards the type of oil produced by the $i \in I$ firms (e.g., LTO), is modeled as a negative shift in κ. I assume all the second-order derivatives of $k(\cdot)$ to be non-negative. Two empirical research questions in Section 3 are whether U.S. LTO production affects global oil prices (p_t) and cost levels in U.S. petroleum activities (k_t).

The theory section disregards uncertainty and I assume that the firms have perfect information about future oil prices and production costs. Firm $i \in I$ maximizes the present value of the stream of profits from resource extraction:

$$V_i = \max_{x_{it}, y_{it}, t_{i1}} \int_{t_0}^{t_{i1}} \pi_{it} e^{-\delta t} dt, \tag{6}$$

where $\pi_{it} = p_t q_{it} - c(x_{it}, y_{it}, S_{it}, k_t)$ is instantaneous profits and $0 < \delta < 1$ is the discount rate. I assume that the discounting is time-consistent and common to all companies. The profit maximization problem in (6) is subject to Equations (1), (2) and (3), with terminal conditions $R_{it_1} \geq 0$ and $S_{it_1} \geq 0$. Note that the time horizon is endogenous in (6). The relations (4) and (5) are not internalized by the competitive firms, but they must be upheld in the competitive partial equilibrium.

Lemma 1. *The competitive partial equilibrium solving (6) must satisfy Equations (1)-(5) and the following necessary conditions:*

$$\pi_{y_{it}}(\cdot) - \omega R_{it} \lambda_{it} - \eta_{it} \leq 0, \tag{7}$$

$$\pi_{x_{it}}(\cdot) + \lambda_{it} - \mu_{it} \leq 0, \tag{8}$$

$$\dot{\lambda}_{it} - \delta \lambda_{it} = -\pi_{R_{it}}(\cdot), \tag{9}$$

$$\dot{\mu}_{it} - \delta \mu_{it} = -\pi_{S_{it}}(\cdot), \tag{10}$$

$$\lambda_{it_{i1}} \geq 0,\ \mu_{it_{i1}} \geq 0, \tag{11}$$

$$\pi_{it_{i1}} + \lambda_{it_{i1}}(x_{it_{i1}} - q_{it_{i1}}) - \mu_{it_{i1}} x_{it_{i1}} = 0, \tag{12}$$

where λ_{it}, η_{it} and μ_{it} are shadow prices described below. Further, we have (i) $y_{it} \leq 1$, with $\eta_{it} = 0$ if $y_{it} < 1$ in Equation (7), (ii) strict equalities in Equations (7) and (8) if and only if $y_{it} > 0$ and $x_{it} > 0$, respectively, and (iii) $\lambda_{it_1} = 0$ or $\mu_{it_1} = 0$ in Equation (11) if and only if $R_{it_1} > 0$ or $S_{it_1} > 0$, respectively.

---

related with capacity utilization, due to capacity bottlenecks and lower drilling quality. Further, Osmundsen et al. (2015) and Skjerpen et al. (2018) find that increased capacity utilization in the rig market increases the rig rates and, hence, the cost of capacity construction in the Gulf of Mexico and on the Norwegian continental shelf, respectively. Last, the modern-day gold rush of oil companies and contractors converging on western Canada's oil-sands markets bogged down as high materials costs and outstripped labor resources forced project delays and budget overruns around the year 2007 (see ENR-Oilsands).

Proof: *See Appendix A.*

Equation (7) states that the shadow price λ_{it} on the developed resource is equal to the marginal profits of resource extraction (for an interior solution). Note that λ_{it} is multiplied by the term ωR_{it} to control for how the choice variable y_{it} controls production, cf., Equation (3). The Lagrange multiplier η_{it}, associated with the constraint $y_{it} \leq 0$, is zero unless the flow rate is at its limit $y_{it} = 1$, in which case production is physically constrained and cannot be increased unless new wells are drilled. Equation (8) states that the shadow price μ_{it} on the undeveloped resource stock S_{it} is equal to the marginal change in profits following a marginal increase in developed fields S_{it}. We have strict equalities in Equations (7) or (8) if and only if $y_{it} > 0$ or $x_{it} > 0$, respectively. The reason is that the firms do not produce oil or drill wells if it decreases total discounted profits V_i in Equation (6). The control variables are then at their lower bounds $y_{it} = 0$ and $x_{it} = 0$ (with strict inequalities in (7) and (8), respectively).

Equation (9) is the Hotelling rule for resource extraction. It is an intertemporal efficiency rule stating that the profits from resource extraction should rise at a rate equal to the discount rate, δ, along the profit maximizing path. Note that the firm could increase the present value of profits V_i by moving extraction across time if the Hotelling condition (9) did not hold. Equation (10) is the Hotelling rule for resource development, stating that the marginal profits from field development also must increase at the rate of discount if present value profits (6) is to be maximized. The inequalities in (11) are the transversality conditions for the non-negative state variables R_{it_1} and S_{it_1}, respectively. Equation (12) is the Maximum principle condition for problems with variable time. It can be shown that we have $x_{it_1} = y_{it_1} = 0$ and $p_{t_1} = \bar{p}$; i.e., we have zero production, zero field development and a price equal to the choke price at the terminal point in time t_1.[15] If the lifting cost $c_i^y(\cdot)$ is sufficiently low for all $i \in I$, such that the constraint $y_{it} \leq 1$ is binding, we have $\eta_{it} > 0$ and equal to the increased value of the objective criterion (V_i in 6) following a marginal slackening of the constraint $y \leq 1$.

The case with low extraction costs may be a reasonable approximation for several conventional oil fields.[16] There are at least two reasons why a corner solution with $y = 1$ is likely to occur whenever the extraction cost $c_i^y(y_{it} = 1, k_t)$ is low relative to the oil price p_t: First, the firms can reduce the present value of the fixed operating cost expenditures $\int_{t_0}^{t_{i1}} c_i^t(\cdot)dt$ by reducing the time horizon (lower t_{i1}). Second, firms discount future development

[15]The price may equal the choke price $\bar{p}$ over the whole time horizon $t \in T$ if the oil price is exogenous, i.e., if $\partial p(\cdot)/\partial \sum_{i \in I} q_{it} = 0$ $(\forall t)$ in Equation (4).

[16]Anderson et al. (2018, see especially p. 997 and their online Appendix C) do not include costs in their analysis, arguing that costs do not play a qualitatively important role for oil production in Texas for the period 1990-2007.

costs and, therefore, do not develop resources before they are needed. This also pulls in the direction of y_{it} being close or equal to one. This implies that current production from active wells with positive production do not respond much to changes in current prices, because the flow already is close to or at its maximum, and new wells must be drilled to increase production. As pointed out by Anderson et al. (2018), the prediction of price inelastic oil supply in the short run is consistent with the data for oil production in Texas for the period 1990 to 2007. It is also consistent with the low short-run oil supply elasticities found in the empirical literature referred to in the Introduction. As we will see in the econometric Section 3, however, this prediction appears to be less consistent with US LTO production. This is not really surprising, as *(i)* LTO production is cost intensive, and *(ii)* LTO producers can adjust the flow rate, albeit at a cost, by increased use of e.g. multi-stage hydraulic fracturing. Given the present paper's focus on LTO, I will henceforth assume that the lifting costs are sufficiently substantial to induce an interior solution for the flow rate $y_{it} \in [0, 1]$.[17] More specifically, I will assume that $lim_{y \to 1} c_i^y(\cdot) > \bar{p}/\omega R_{it}$, which ensures that we have $\eta_{it} = 0$.[18]

Appendix B presents a numerical illustration to ease understanding of the model in Lemma 1. The model is solved as a nonlinear programming (NLP) problem in GAMS (numerical software) using the CONOPT solver. I will present some figures from this illustration in the text, but refer to Appendix B for further details on the numerical model.

2.1 Selected implications of the necessary conditions for petroleum extraction

In this section I discuss some predictions from the model in Lemma 1. The topics are selected based on their relevance for U.S. LTO production.

2.1.1 The oil price

Suppose we have an increase in the exogenous demand parameter ν_t in the model in Lemma 1, such that the oil price at time $t' \in T$ increases from p_s to $p_s + \Delta$ for $s \in T$ and $s \geq t'$. Here $\Delta > 0$ is the constant and permanent change in the price trajectory occurring at time t'. The theory framework suggests that production from existing wells has two key responses to the increased oil price: (i) a short-run response caused by an increase in the

[17]Newell and Prest (2019) estimate the price elasticity of drilling of unconventional wells to be in the range of 1.2-1.9, and the price elasticity production from finished unconventional wells to be 0.12 (both statistically significant). Note that a positive price elasticity on production implies that the flow rate can be adjusted.

[18]It can also be verified that the constraint $y_{it} \in [0, 1)$ holds after the solution to (6) is derived.

flow control y_{it}, and (*ii*), a long-run response via increased field development x_{it}, which increases developed reserves R_{it} and thereby oil production. Mechanism (*i*) is present because, everything else equal, the optimal flow control y_{it} increases in current prices (cf., Equation, 7). Mechanism (*ii*) occurs because the shadow value of developed reserves λ_{it} increases in future oil prices (cf., Equation 9). This increases resource development (cf., Equation 8), which again increases future production (cf., Equation 3). The price increase typically implies that resource extraction is moved forward in time, implying more early extraction, and less late extraction.

The dynamics in the case of an announced *future* increase in oil prices is somewhat less straightforward, because of two opposing effects. The profits of future production increases, implying a larger shadow value λ_{it} (cf., Equation 9). This pulls in the direction of (*i*) less extraction today (cf., Equation 7) and (*ii*) more resource development (cf., Equation 8) and, hence, increased production (cf., Equation 3). Note that whereas mechanism (*i*) occurs very fast, there is a delay before increased field development cause oil production to increase, depending, e.g., on the extraction technology.[19] It follows that the current effects of a known future price increase is in general ambiguous.[20] Figure 2 illustrates the effects on optimal production following a known increase in the oil price in period $t' = 50$.[21] We see that production before the price increase in the future period $t' = 50$ decreases slightly in this numerical simulation.

2.1.2 New undeveloped resources

Suppose there is one single anticipated addition to the undeveloped resource S_{it} at some future time $t = t' > t_0$, e.g., because new areas with known resources will be made available to the petroleum industry.[22] The effect before the new resources are available (i.e., $t \in [t_0, t')$) is to reduce the current shadow price on the resource stock μ_{it} (cf., Equation 10). The isolated effect of this is to increase resource development (cf., Equation 8)

[19]The dynamic response of U.S. LTO production following increased drilling activity is examined in the Section 3; see Figure 5 C in particular.

[20]This result is relevant to the literature on intertemporal effects induced by future environmental policies in the presence of resource scarcity. In particular, Sinclair (1992) and Sinn (2008) caution against environmental policies that become more stringent with the passage of time, because such policies will accelerate resource extraction and, thereby, accelerate global warming.

[21]The references to the numerical illustrations are in a discrete time framework. When comparing Figures 1 and 2, it is important to remember that the time horizon in the theory model is the whole lifespan of positive production, whereas Figure 1 only graphs around twelve years. It appears reasonable to assume that U.S. LTO is still in its early phases, corresponding to the time with increasing production in Figure 2, i.e., before resource scarcity forces a production decline.

[22]We have $d_{it} = 0$ over the whole time horizon except for the single resource discovery at time $t = t'$.

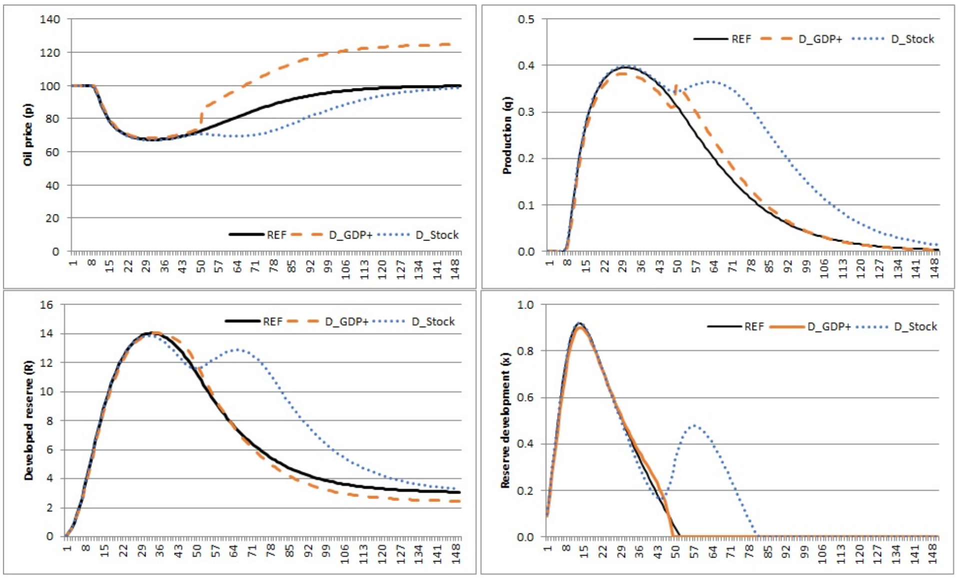

Figure 2: Numerical illustration of selected model variables in reference scenario (REF) and two scenarios identical to REF, except that the oil price (*D_GDP*) and remaining resource stock (*D_Stock*) increases in period $t' = 50$, respectively. Time periods along the horizontal axis

and thereby production (cf., Equation 3). On the other hand, if the resource is anticipated and the new resource is relatively cheap to extract, resource development may be postponed to take advantage of the cheaper extraction in the future. The relative sizes of these counteracting effects depends, e.g., on the waiting time before the new resource is available, the cost of extracting the new resource (as compared to the existing resource stock), and the size of current reserves (relative to production).[23] After the resource is made available (i.e., $t > t'$), the increase in the resource stock S_{it} will also reduce the cost of resource extraction. The effect may occur immediately after the discovery, or later on along the time trajectory (again depending on the relative cost of developing the newly discovered resources as compared with the old resource stock). The effects of a single resource discovery in period $t = 50$ in the numerical illustration is given in Figure 2 (*D_stock*) ($d_{it} > 0$ for $t = 50$ and $d_{it} = 0$ for $t \neq 50$), where the model parameters are such that some of the new resource is profitable to develop at once. In the period before the resource is made available, resource development first

[23]This allows for a theoretical 'green paradox type' argument where environmental policy decisions that close down areas for future petroleum activity may increase current resource development and, hence, current production (because the industry develops currently available reserves instead of waiting for the new and more promising area to be made available). Whereas this may be relevant, e.g., for the literature on supply side climate policies, the magnitude of this unintended effect, if present, may be very modest.

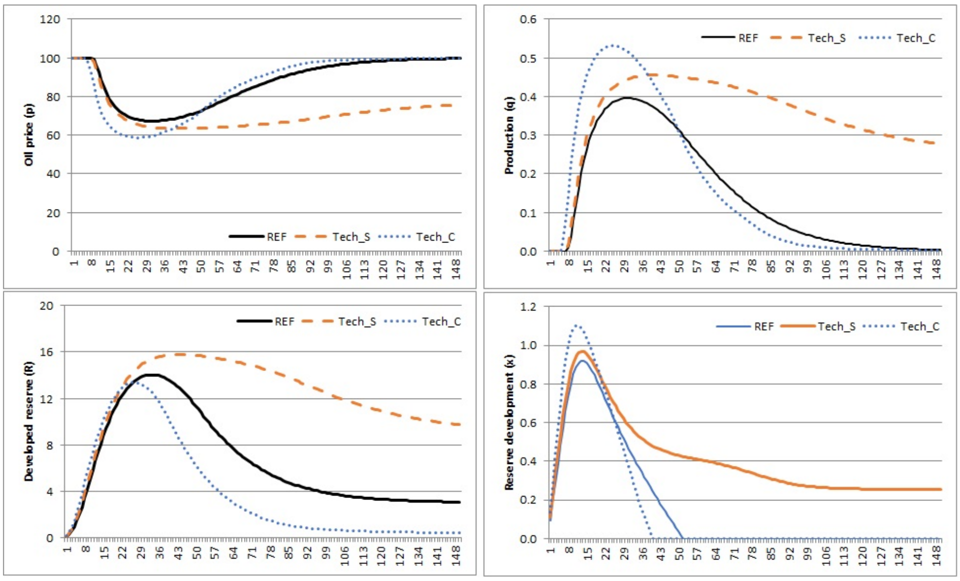

Figure 3: Numerical illustration of selected model variables in reference scenario (REF) and two scenarios identical to REF, except that either the field development and production costs decreases ($Tech\_C$), or that the undeveloped resource stock S_{it} grows a small amount each period(D_Stock)

increases and then decreases relative to the reference scenario (REF). This reflects the two counteracting mechanisms described above. In the longer run, the new resource increases extraction and reduces the petroleum price.

2.1.3 Technological change

In this section I discuss the role of technological development. This is highly relevant to unconventional petroleum production in general, and U.S. LTO in particular. I examine two types of technological change: *(i)* a gradual decrease in production and field development costs, $c_i(\cdot)$, and *(ii)* an exogenous growth in the resource stock, S_{it}, e.g., because of advances in horizontal drilling techniques that allow more resources to be exploited. Whereas this dichotomy highlights two key aspects related to technological change in the petroleum industry, the technological advances that have spurred e.g. the U.S. shale revolution feature a mixture of both.

A gradual exogenous decline in production costs $c_i(\cdot)$, caused by a decrease in k_i, has two key effects on production. First, field development and production will be cheaper, implying increased developed reserves (cf., Equation 8) and current production (cf., Equation 3 and 7). The isolated effect of this is to increase current production. On the other hand, with continuous technological progress it will be even cheaper to produce in the future, and the value of future reserves increases (cf., Equation 9). This pulls in the direction of delaying production (cf., Equation 7). The effect of continuously reduced extraction and field development costs are illustrated in Figure 3 ($Tech\_C$), where the first effect dominates, implying that extraction is pushed forward in time.

A gradual exogenous growth in undeveloped reserves, $d_{it} > 0$ for $\forall t \in T$, implies lower resource scarcity, and hence a lower shadow price on the undeveloped resource stock S_{it} (cf., Equation 10). This increases field development (cf., Equation 8) and thereby production (cf., Equation 3). This is illustrated in Figure 3 ($Tech\_S$), where the increase in stock each period is sufficiently small to retain resource scarcity ($\mu > 0$).

3 A vector error correction model for U.S. light tight oil production and rig activity

This section continues the examination of LTO supply, but from an empirical angle. An important difference between the theory model and the econometric model is that the theory features a long time horizon in which resources are depleted over time. In comparison, the econometric model is based on the covariation between non-stationary variables and a steadily rising U.S. LTO production (cf., Section 3.3 below). The key research objective of Section 3 is to quantify how U.S. LTO production and the associated rig activity respond to changes in the oil price, both in the short and longer run.

3.1 Variable selection and other model considerations

Key variables of interest in this paper are the oil price (p_t), oil production (q_{it}) and rig activity (x_{it}) (which corresponds to field development in the theory section). Further, the theory model indicates that it is really the difference between the oil price and production costs that is decisive for production and rig activity (cf., Equation (6)). This means that modeling of production costs is also important. There are several possible ways to operationalize costs in the empirical model, e.g., real interest rates (capital costs), wages (labor costs), rig rates and so forth. In this paper I use a cost index for equipment and capital in the oil and gas industry as a proxy for costs, see Section 3.2.

Whereas production and rig activity are clearly endogenous variables, the theory is ambiguous about whether the oil price is endogenous or not, depending, e.g., on the size of the petroleum industry that is to be examined (cf., Equation (4)). The theoretical model also indicates that both increased rig activity and increased oil production can increase marginal production costs, both due to convex cost functions and due to potential economy-wide capacity constraints (cf., Equation (5)). Further, exogenous shifts in the

Table 1: P-values from Granger-causality tests in a VAR model, specified on first-differences of log-transformed variables (see Table 2) with one lag.

	Not Granger-causing variable	**No instantaneous causality**
diff(oilprice)	2.2e-16***	0.0217**
diff(cost)	0.5459	0.0217**
diff(rigs)	0.0370**	0.6793
diff(oil)	0.0466**	0.0869*

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

cost functions, e.g., due to new technology that makes previously unavailable reserves economically attractive, affect the level of production (cf., Sections 2.1.2 and 2.1.3). Consequently, it is not entirely obvious which, if any, of the above-mentioned variables may be exogenous in the model. Granger causality tests have been carried out to shed some further light on the topic of variable endogeneity, see Table 1. The null-hypotheses in Table 1 are: (*i*) X do not Granger-cause Y, and (*ii*) no instantaneous causality between X and Y, where X is the relevant variable and Y is the set of remaining variables (see Appendix C for details). The Granger causality test does not really indicate endogeneity or causality, but rather whether one time series is useful for forecasting another.

Current production depends on both current and future oil prices, and rig activity and field development primarily depend on the companies' expectations about future oil prices, which are not observable. This is a challenge when modeling oil production, and extraction of other exhaustible resources in general. The present paper assumes the adaptive expectations hypothesis in the empirical modeling of U.S. LTO production. Under this hypothesis, expectations about future oil prices can be modeled as functions of past and present oil prices (and perhaps other variables too). Oil production and rig activity are functions of lagged oil prices in the econometric model. Hence, I (implicitly) assume that companies' price expectations are adaptive and continually updated in the modeling of production and rig activity decisions. Adaptive expectations about the future oil price may be an important reason why the lagged oil price is a significant explanatory variable for rig activity (see Table 3). The adaptive expectations hypothesis is fairly standard in the empirical literature, see e.g. Farzin (2001) Nguyen and Nabney (2010), Aune et al. (2010), Osmundsen et al. (2015) and Skjerpen et al. (2018).[24]

The theory suggests that the price response of oil production and rig activity will change over time. For example, the short-run response in LTO production to higher oil prices is likely to be smaller than the long-run response. The reason is that it takes time before changed rig activity, induced by the oil price change, leads to changes in oil production. This simple ob-

[24]See also Reitz et al. (2009) on the role of regressive expectations and oil price forecasting.

servation indicates that a dynamic model may be appropriate for adequately capturing how oil production responds to changes in oil prices.

The theory and the investigations of Granger-causality between the variables in Table 1 suggest that a dynamic model with several endogenous variables may be suitable. The point of departure for this paper is a vector autoregressive (VAR) framework. This allows modeling of several endogenous variables and a reasonably flexible lag structure. The following four endogenous (log-transformed) variables are included in the model: U.S. LTO production, rig activity, the oil price, and a proxy for U.S. LTO supply costs. Other variables that have been included in the VAR model selection process, but which do not enter the final model, are U.S. gas prices, the U.S. long term interest rate, industrial production indexes for OECD and the U.S., U.S. GDP, the U.S. wage rate, and a (noisy) measure for existing production capacity.[25] The choice of variables has been made with a focus on modeling oil production and rig activity, not the development in the oil price (which is beyond the scope of the present paper). The state variables for developed and undeveloped reserves play important roles in the theory model, but do not enter the econometric model directly. These variables influence production via the cost function and are therefore indirectly present via the cost variable. Remember that the time horizon in the theory model is significantly longer than the data basis for the empirical analysis, and that resource scarcity does not necessarily play a central role in the data sample period. Further, reserves and reservoir pressure also play a direct role in production (cf., Equation (3)), which is present with several lags in the VAR model. The lags of oil production may thus not only capture existing infrastructure and capacity, but also reservoir characteristics like pay zone thickness, rock permeability and pressure.

The author has not identified any changes in regulation of U.S LTO production over the relevant time span (Jan. 2010 to Apr. 2022, see Section 3.2 below) that needs to be controlled for in the econometric model. Two possible issues were (*i*) the U.S. environmental protection agency (EPA) issued new rules in 2012 to limit emissions of some air pollutants from fracking, and, (*ii*), in 2015, New York became the first state with significant natural gas reserves (the Marcellus Shale play) to prohibit fracking. Note that the Energy Policy Act of 2005 excluded fracking from the Safe Drinking Water Act's underground injection control's regulation, except when diesel fuel is used. President Joe Biden pledged a moratorium on new oil and gas leasing on federal lands and waters, but at the time of writing it appears unlikely that he will able to fulfill this until his first term ends.[26]

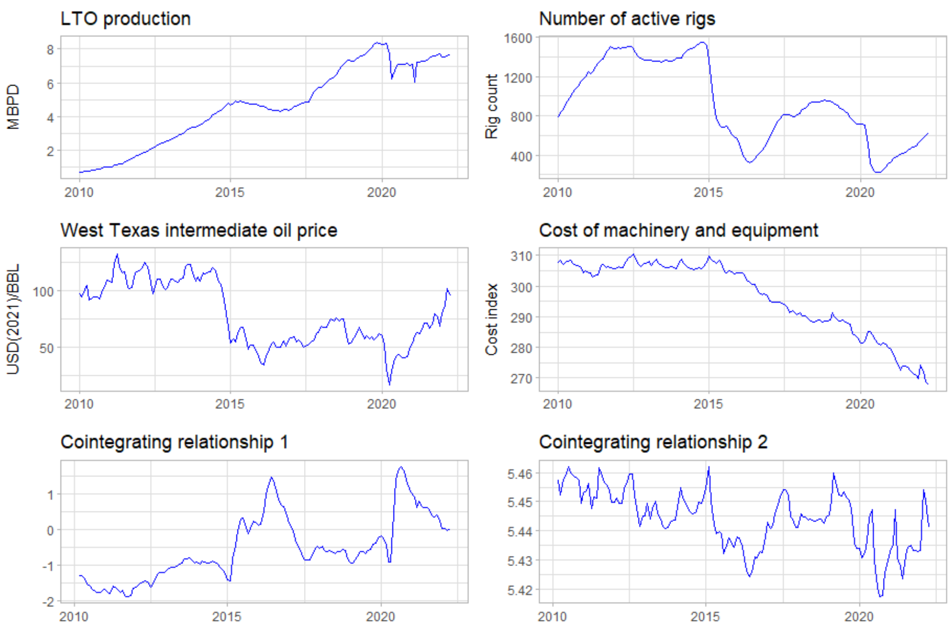

Figure 4: The endogenous variables included in the analysis and the cointegrating relationships. LTO production in million barrels per day (MBPD). Real WTI per barrel of oil in US 2021 dollars. The cointegrating relationships are generated from the log transformed variables (see Figure 9 in Appendix C).

3.2 The data

I use monthly data for the period Jan. 2010 to Apr. 2022 (148 months). The data series are graphed in Figure 4, which also includes the cointegrating relationships, see Section 3.3 below. An overview of the endogenous variables are given in Table 2.

The data for U.S. LTO production are fetched from the U.S. Energy Information Administration (EIA) and cover the whole U.S. The data for the number of active rigs cover the regions included in EIA's Drilling productivity report for key tight oil and shale gas regions: Andarko, Appalachia, Bakken, Eagle Ford, Haynesville, Niobrara and Permian. The data set does not distinguish between oil-directed and gas-directed rigs, because once a well is completed it may produce both oil and gas (more than half of the wells do that). Hence, the rig figures used in the analysis are rigs involved in both oil and gas extraction operations. I use the number of active rigs as a proxy for rig activity and field development.[27]

[25]See 'legacy oil production change' at EIA-drilling

[26]Washington Post

[27]Exploration wells play a less important role in the U.S. LTO extraction. Since extensive resources have already been discovered, extraction costs and technology have traditionally been the bottleneck.

Table 2: Overview of monthly time series used in the econometric model (Jan. 2010 to Apr. 2022).

Variable	Description	Underlying variable	Source	Denominator
oilprice	Log of real oil price	WTI spot price	EIA	USD(Oct. 2021)/barrel
cost	Log of proxy variable for cost	Cost index for oil and gas field machinery and equipment	FRED	Producer price index
rigs	Log of number of active rigs	Rig count	EIA	Number of active rigs
oil	Log of LTO production	Oil production	EIA	Barrels of oil/day

For the oil price I use the Cushing (Oklahoma) monthly West Texas Intermediate (WTI) FOB spot price. This price is deflated using the U.S. consumer price index.[28]

It is very difficult to find good data on costs related to oil production. Companies such as Rystad Energy and IHS Markit provide variables that may be used as proxies at a cost. In this paper I use the 'Producer Price (monthly) Index by Industry: Oil and Gas Field Machinery and Equipment Manufacturing', published by the U.S. Bureau of Labor Statistics, as a proxy for cost. It is available as monthly data and can be freely downloaded from the Federal Reserve Economic Data base (FRED).[29] This price index was also used as a measure for production costs by Golombek et al. (2018). The cost index is not significant in the final model, but if omitted the residual diagnostics are worsened, in particular regarding heteroscedasticity. The cost index was significant with expected signs in several preliminary and discarded model formulations. The cost index is deflated using the U.S. consumer price index (same as oil price above).

There are some outliers in the data, perhaps most notably associated with the Covid-19 pandemic and Russia's invasion of Ukraine. Note that oil production, the oil price and rig activity all drop sharply in April 2020 (see Figure 4).[30] To deal with this, I formulated a general model with monthly dummies from and including March 2020, and then removed the least important dummy variables. This involved a trade-off between limiting the total number of dummies, retaining the significant dummies, and maintaining acceptable model properties, specifically in terms of autocorrelation and heteroskedasticity. The model also includes seasonal dummies. The dummies and their significance are given in Table 4 in Appendix C.

[28] See EIA-spot-prices and FRED-CPI. The abbreviation FOB indicates that the price is for oil loaded onto a vessel and ready for shipping. So it includes the cost of purchasing and loading the oil, but not the cost to deliver it to its final destination.

[29] See FRED-costindex

[30] President Trump declared a nationwide US emergency on March 13, 2020, because of Covid-19, and U.S. states began to shut down to prevent the spread of Covid-19 on March 15. In May, 2020, the U.S. unemployment rate was 14.7 percent, the highest rate since the Great Depression.

3.3 The vector error correction model

Consider the VAR model with m endogenous variables, n exogenous variables (incl. dummies) and l lags:[31]

$$\mathbf{y}_t = \mathbf{c} + \delta_1 \mathbf{y}_{t-1} + \cdots + \delta_k \mathbf{y}_{t-l} + \phi \mathbf{x}_t + \epsilon_t, \quad t = 1, 2, \ldots, T, \tag{13}$$

where $\mathbf{y}_t$ is a $m \times 1$ vector of endogenous variables observed at time t, $\mathbf{x}_t$ is a $n \times 1$ vector of exogenous variables, ϵ_t is a $m \times 1$ vector of error terms, and δ_k and ϕ are $m \times m$ and $m \times n$ coefficient matrices.

As mentioned above, the final econometric model for LTO production has $\mathbf{y}_t = (oilprice_t, cost_t, rigs_t, oil_t)^\intercal$, $\mathbf{x}_t$ consists of the dummy variables, and we have $T = 148$ months of observations. The number of lags, l, remains to be determined. For our data and the specification given by Equation (13), the The Akaike Information Criterion, the Hannan-Quinn Criterion, the Schwarz Information Criterion and the Final Prediction Error Criterion all suggest using two lags in the VAR model (see Appendix C). Model experimentation shows that two lags yield a model with no significant autocorrelation in the residuals (see reported diagnostics later in this section). Hence, I specify the model in (13) with two lags, $l = 2$.[32]

Visual inspection of Figure 4 suggests that the variables in $\mathbf{y}_t$ may be non-stationary and possess unit roots. This is also suggested by the autocorrelation (ACF) and partial autocorrelation (PACF) plots associated with the four endogenous variables, which all decay slowly and remain well above the 95 percent significance range for the 24 months plotted (see Figure 11 in Appendix C).[33] To check more formally for the presence of unit-roots, I perform augmented Dickey-Fuller tests and Phillips-Perron unit-root tests. The test results indicate that the variables in levels have a unit root, and that the first-differences of the variables are stationary (see Appendix C). This indicates that estimation of a VAR in levels is problematic due the possibility of spurious or nonsense regressions. This issue can be ameliorated by estimating a VAR on the stationary first-differenced data. It is not ideal, however, to fit a VAR in differences if the system features cointegrating relationships, because the variables in levels then contain information that is useful for explaining the movement of the variables beyond that contained in a finite number of lagged differences alone (see, e.g., Johansen and Juselius, 1990, and Hamilton, 1994).[34]

[31] This paper follows somewhat conflicting conventions and use notation π and δ to denote profits and the discount rate in Section 2, and coefficient matrices in Section 3.

[32] Remember that the number of lags in the VAR determines the functional form of the lag structure, not the memory length of the process. For example, the simple AR1 model $y_t - \rho x_{t-1} = \phi_t$ is equivalent with $y_t = \sum_{n=1}^{\infty} \rho^n \phi_{t-n}$ (the Koyck transformation).

[33] As a rough rule of thumb, the ACF declines linearly for an I(1) series and exponentially for an I(0) series.

[34] The matrix polynomial associated with the moving average operator of the cointe-

Defining $\Delta = 1 - L$, where L is the lag operator, the VAR (13) can be rewritten in VECM form (with $l = 2$, see Appendix A):

$$\mathbf{\Delta y}_t = \gamma_0 + \gamma \Delta \mathbf{y}_{t-1} + \pi \mathbf{y}_{t-1} + \phi \mathbf{x}_t + \epsilon_t, \qquad t = 3, 4, \ldots, 148. \tag{14}$$

We observe that the model (14) is a standard VAR in (stationary) first-differences, except for the equilibrium correction term $\pi \mathbf{y}_{t-1}$. The matrices π and γ capture the long- and short-run impacts of shocks to the dynamic system (14), respectively.

The choice between a VAR in differences and the VECM (14) depends on whether the coefficient matrix π contains information about long-run relationships in the data vector $\mathbf{y}_t$ or not. This issue can be examined by analyzing the rank of π. As pointed out by, e.g., Johansen and Juselius (1990), there are three possible cases: Firstly, (i), $\text{rank}(\pi) = m = 4$, i.e., the matrix π has full rank, which indicates that the vector process $\mathbf{y}_t$ is stationary. This outcome is not consistent with the augmented Dickey-Fuller tests and Phillips-Perron unit-root tests mentioned above. Secondly, (ii), $0 < \text{rank}(\pi) = r < m = 4$, i.e., we have r cointegrating relationships and a VECM is appropriate. Thirdly, (iii), $\text{rank}(\pi) = 0$, indicating that π is the null matrix and the model in Equation (14) reduces to a standard VAR in first-differences.

A likelihood ratio test for no linear deterministic trend in the cointegrating relationship was conducted (assuming two cointegrating relationships, see below). This test rejected the null-hypothesis of not including a trend at one percent level of significance (see Appendix C).

I use the trace type of the Johansen test (Johansen and Juselius, 1990, Johansen, 1991; 1995), specified with two lags and a linear time trend, to examine $\text{rank}(\pi) = r$. The test procedure rejected the two null-hypotheses $r = 0$ and $r \leq 1$ at one percent level of significance. Further, the null-hypothesis $r \leq 2$ is not rejected at a level of significance equal to 5 percent or less. Hence, the test indicates that $\text{rank}(\pi) = 2$ (at a 5 percent level of significance or below). I proceed by specifying a VECM under the assumption that we have two cointegrating relationships in the data. The matrix product $\pi \mathbf{y}_{t-1}$ then consists of the first lag of two stationary linear combinations of the variables in levels (that are themselves non-stationary) and their coefficients. The estimated VECM (see Equation 14) is given in Table 3. Detailed output from the econometric software R, including the unrestricted VECM, is given in Appendix C. See Appendix C for details, including the Johansen test output, plots of the cointegrating relationships, and the estimated unrestricted VECM.

The main text graphs the endogenous variables in levels, but the log transformed endogenous variables in levels and first-differences, as included

grated system has a root at unity, implying that the moving average operator is noninvertible and thus no finite-order VAR can describe the process (Hamilton, 1994, p. 573).

in the econometric model, are graphed in Figure 9 in Appendix C, which also includes complete regression results including the 11 seasonal and 10 monthly dummy variables. Whereas Table 3 gives some indication of how the variables interact in the model, along with their statistical significance and how well the model explains the variation in the endogenous variables, some care should taken when interpreting the estimated coefficients. The reason is that this is a dynamic model with four endogenous variables that interacts with each other. Section 3.4 below discusses results and illuminates the dynamics of the system in the context of impulse response functions.

Table 3: The VECM model. Estimated equation by equation using OLS.

	Dependent variable:			
	diff(oilprice)	diff(cost)	diff(rigs)	diff(oil)
	(1)	(2)	(3)	(4)
Cointegrating relationship 1	0.053**	−0.003***	0.007	−0.016***
	(0.026)	(0.001)	(0.007)	(0.004)
Cointegrating relationship 2	4.406**	−0.340***	−0.867	−0.073
	(2.085)	(0.070)	(0.557)	(0.297)
constant	−23.977**	1.850***	4.721	0.407
	(11.345)	(0.378)	(3.031)	(1.615)
diff(oilprice).lagged	0.274**	−0.013***	0.137***	0.045***
	(0.107)	(0.004)	(0.029)	(0.015)
diff(cost).lagged	−0.911	0.158*	1.068	0.098
	(2.545)	(0.085)	(0.680)	(0.362)
diff(rigs).lagged	−0.163	0.004	0.861***	0.060***
	(0.153)	(0.005)	(0.041)	(0.022)
diff(oil).lagged	0.671	0.031	0.331**	0.077
	(0.575)	(0.019)	(0.154)	(0.082)
Observations	146	146	146	146
R^2	0.487	0.516	0.915	0.911
Adjusted R^2	0.366	0.401	0.895	0.890
Residual Std. Error (df = 118)	0.090	0.003	0.024	0.013
F Statistic (df = 28; 118)	4.008***	4.486***	45.374***	43.089***
Note:				*p<0.1; **p<0.05; ***p<0.01

The Johansen procedure decomposes the matrix π in equation (14) into two matrices α and β, defined such that $\alpha\beta^\intercal = \pi$. The cointegrating relations are then given by $\beta^\intercal \mathbf{y}_t$, whereas α is the loading (or adjustment) matrix. The cointegrating relations are graphed in Figure 4. They capture common trends that links the variables in the long-run. As pointed out by Johansen (1995, p. 41), the long run relations $\beta^{\intercal}\mathbf{y}_t$ are not relations that are satisfied in the limit as $t \to \infty$ (unless $\epsilon_t = 0$ for all sufficiently high values of t). Rather, they are relations in the economy, as described by the statistical model, which pulls the variables towards the attractor set defined by the cointegrating relations. The speed of which the variables are pulled towards the attractor set is determined by the loading matrix α. The estimated loading matrix α and the estimated eigenvectors β are specified in Appendix C.[35]

Below is a summary from a suite of tests on the model in Table 3 (see Appendix C for details).

Stationary model residuals: Philips-Perron unit root tests conducted on the residuals from the VECM reject the hypothesis of unit root for all four residual time series (all p-values are equal to 0.01). This corroborates the result from the Johansen-test on the presence of cointegrating relationships and indicates that the choice of a VECM was appropriate.

Autocorrelation: The Portmanteau test, the Breusch-Godfrey LM (BG) test, and the Edgerton-Shukur F test, which generalize the BG test to systems of equations, all reject the presence of serially correlated error terms in the VECM model at 5 percent level of significance. This indicates that the specification with two lags (in the VAR form) is sufficient for this model.

Heteroscedasticity: A multivariate ARCH-LM test rejected the presence of autoregressive conditional heteroscedasticity in the VECM at 5 percent level of significance.

Normally distributed residuals: Multivariate Jarque-Bera tests and multivariate skewness and kurtosis tests for the residuals in the VECM indicate that the model residuals are not normally distributed.

The Gauss Markov Theorem states that the OLS estimator is the best (i.e., smallest variance) linear unbiased estimator. This result is contingent on the absence of autocorrelation and heteroscedasticity, but does not require normally distributed error terms.[36] Without normally distributed disturbances, the exact distributions of the F, t and chi-squared statistics depend on the data and the parameters, and are not exactly F, t and chi-squared, however. The Central Limit Theorem (CLT) implies that, as a

[35]The parameters in the matrixes α and β are not uniquely identified. The reason is that the matrixes are derived from π and, for any choice of α and β and a non-singular $m \times m$ matrix φ, $\alpha\varphi$ and $\beta(\varphi)^{\top}$ will give the same matrix π. In this paper normalize the cointegrating relations to the first column, as suggested by Johansen (1995) and default in the the R urca package (see Appendix C). The normalization is not important for the results presented in this paper (any choice of normalization gives the same VECM).

[36]See Carter Hill et al. (2001, p. 77) for more on the conditions of the Gauss Markov Theorem. In the particular context of cointegration, Johansen points that 'The methods derived are based upon the Gaussian likelihood but the asymptotic properties of the methods only depend on the i.i.d.assumption of the errors' (Johansen, 1995, p. 29).

large sample approximation, the standard normal distribution can be used to approximate the true distribution of the t-test statistic. The CLT also implies that the Wald statistic is asymptotically normal, even in the absence of normally distributed disturbances. The implication of this is that use of the conventional t and F test statistics is a reasonable approach in large samples (see, e.g, Greene, 2003, p. 106 and 108). The final model uses 146 observations and has 118 degrees of freedom (we have 148 observations and lagged first-differences for each equation in Table 3, and several dummy variables - see Table 4 in Appendix C). Whereas this sample is arguably sufficiently large to rely on the CLT, the main empirical results in this paper is based on bootstrapped impulse response functions and, hence, do not rely on the normal distribution. Furthermore, an alternative to the standard errors reported in Table 3 is standard errors estimated by bootstrapping. Such standard errors are reported in Appendix C (they are of similar magnitude).

Coefficient stability: A test based on cumulative ordinary least squares residuals for structural change was done to assess the stability of coefficients. Under the null hypothesis of coefficient constancy, values of the sequence outside an expected range suggest structural change in the model over time. The test results do not indicate model instability, see Figure 14 in Appendix C.

3.4 Model results

Figure 5 presents selected impulse response functions (IR) and 95 percent bootstrapped confidence intervals (CI) from the model in Table 3.[37] The IRs graphs the estimated effects on the relevant variable following a one unit transitory shock. Because all variables enter the model in natural logarithms, we can read Figure 5 as percentage changes in the 'dependent' variable following a one percent shock to the 'impulse' variable. According to the estimated VECM, a one percent (positive) transitory shock to the oil price gives a long-run increase of 0.3 percent in U.S. LTO production.[38] We also find a positive effect on rig activity, which tops out at 1.2 percent after 12 months and then declines and stabilizes around 0.8 percent. The VECM indicates that a one percent shock to rig activity increases oil production with 1.7 percent in the longer run. We further observe that increased rig activity significantly increases the cost of oil production, which indicates the presence of capacity constraints as conjectured in the theory section (cf., Equation 5). There is also a somewhat smaller but significant effect from LTO production on costs, see Figure 10A in Appendix C. Last, I find no significant effects on the oil price following a one percent transitory shock to U.S. LTO production, see Figure 10 B in Appendix C. The lack of signif-

[37]The effects in Figure 5 are also significant at the 1% level of significance.

[38]In terms of the theory model, this could be modeled as a shock to the exogenous ν in Equation (4), calibrated such that the oil price increases with one percent.

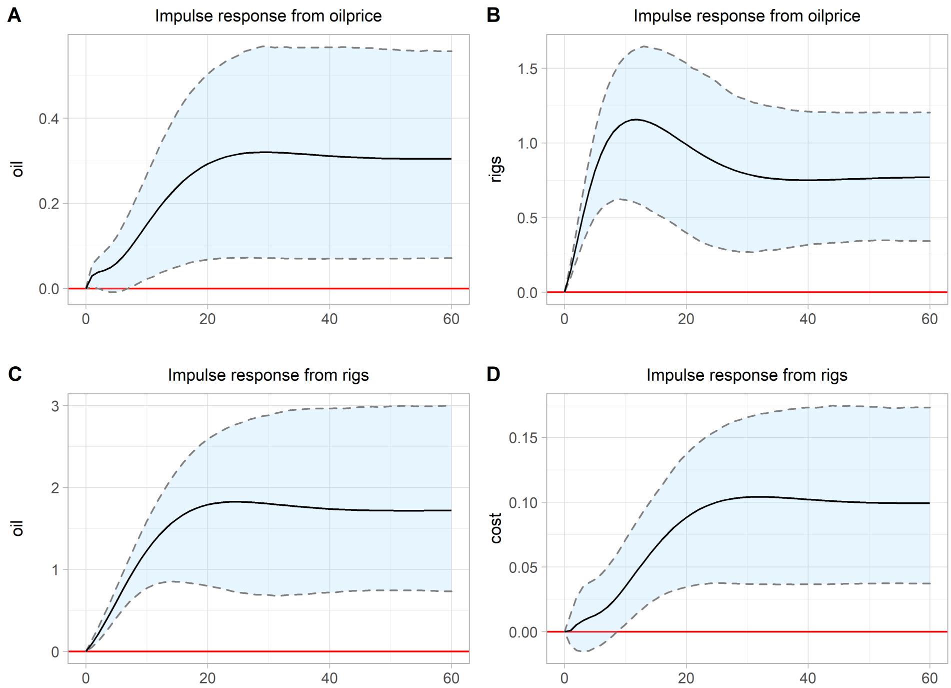

Figure 5: Selected impulse response functions. 95% bootstrap CI, 5000 runs. Generated from the model in Table 3.

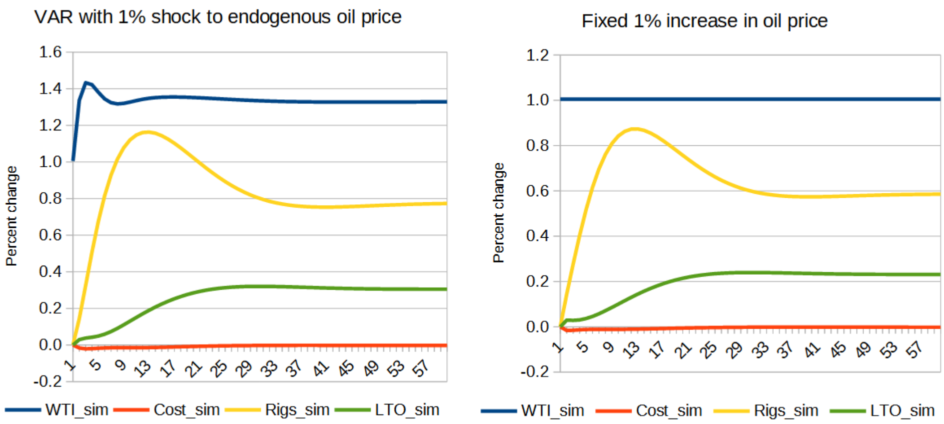

Figure 6: Left: Simulation of effect following a one percent transitory shock to the oil price in period 1. Right: Simulation of model with a fixed one percent increase in oil price. Changes are given in the original variables (not logarithms).

icance may not be that surprising, given the high volatility of oil prices (see Figure 4), the many factors that affects the it, and the global nature of the oil price.[39]

We also observe that the effects of the transitory shocks do not disappear over time, but rather that the model converges towards a new equilibrium. This is because we have a system with cointegrated non-stationary variables.

As pointed out by Anderson et al. (2018), most of the Hotelling style resource economics literature neglects the role of pressure dynamics and drilling activity. Panels B and C in Figure 5 highlights the importance of accounting for drilling activity and rig markets when examining petroleum extraction, both in theoretical and empirical models.

It is not straightforward to compare the results in Figure 5 with the previous literature on crude oil production. One reason is that the literature often focuses on the supply elasticity of oil; i.e., the percentage increase in oil supply induced by a one percent permanent increase in the oil price. Such an elasticity cannot be directly derived in the VECM model. Specifically, the oil price is not an exogenous variable in the VECM, and a change in the oil price in some period t will induce changes in the oil price in period $t+1$, $t+2$ and so forth.

The left graph in Figure 6 illustrates the effects in the VECM of a one percent transitory shock to the oil price in period $t = 1$ (disregarding un-

[39] Besides OPEC and OPEC plus policy decisions, and the Covid-19 pandemic, political and financial issues like the uprisings in Egypt and Libya in 2011 and the Syrian conflict have undoubtedly affected the oil price development during the last decade or so. The U.S. LTO share of global oil production has increased much during the sample period, see Figure 1. It is conceivable that one would have found a significant effect from the U.S. LTO on oil prices if the U.S. LTO's share of global oil production had been at, for example, the 2020 level throughout the data period.

certainty). The figure is based on two simulations over the time horizon $t = \{-1, 0, ..., 60\}$. In the baseline scenario, I set all variables to their mean values in periods $t = -1$ and $t = 0$, and then forecast the values for the remaining periods $t = \{1, 2, ..., 60\}$ (i.e., five years). Then I run a scenario which is identical, except that the oil price is hit by a transitory shock that increases the oil price in period $t = 1$ with one percent. The effects of the shock is measured as the difference between the two scenarios. The results, which are graphed in the left part of Figure 6, correspond to the standard impulse response functions. Note that the increase in oil prices following the transitory shock is greater than one percent after the first period.[40]

The right hand side of Figure 6 is obtained by the same procedure, except that the oil price is fixed at its mean value in the preceding periods $t = -1, 0$, and at its mean value times 1.01 in periods $t = 1, 2, \ldots, 60$. Whereas the results from this exercise may be somewhat easier to compare with the results on price elasticities in the previous literature, the right hand side in Figure 6 does not represent the true VECM model in Table 3. The reason is that the fixed oil price compromise the model dynamics. The approximation may nevertheless have some value, because the oil price appears to be only modestly dependent on the other variables in the VECM, see, e.g., Figures 8 and 10, and the test for a weakly exogenous oil price below.

The results in Figure 6 suggest that a one percent increase in the oil price causes a 0.3 percent increase in U.S. LTO production, or 0.2 percent if we restrain the model to feature an exogenous oil price. In comparison, Bjørnland et al. (2020) find the short run supply elasticity of shale wells to be positive and in the range of 0.3–0.9, depending on wells and firms characteristics. Anderson et al. (2018) estimate a price elasticity of approximately 0.7 for drilling in Texas during 1990-2008, whereas Newell and Prest (2019) find a cumulative drilling response of 1.6 percent for unconventional wells. We observe from Figure 6 that the VECM indicates that the response in rig activity varies markedly over time.

The error terms in the different equations of the VECM in the Table 3 can be correlated with each other. The covariance matrix of the error terms represents the contemporaneous (i.e., within the same month) effects. These contemporaneous effects are not captured in Figures 5 and 6, so a shock to one variable in period t cannot cause an effect in any other variable before the next period $t + 1$. This is somewhat problematic, as the shocks to the variables are indeed correlated, in this case especially oil and rigs, and rigs and oil price; see the correlation matrix for the residuals given in Appendix C.

To ameliorate this, Figure 7 presents orthogonal impulse response func-

[40]Remember the presence of a unit root in the variables; cf., Section 3.3. That is, whereas the exogenous shock itself is transitory, its effects on the endogenous variables are not.

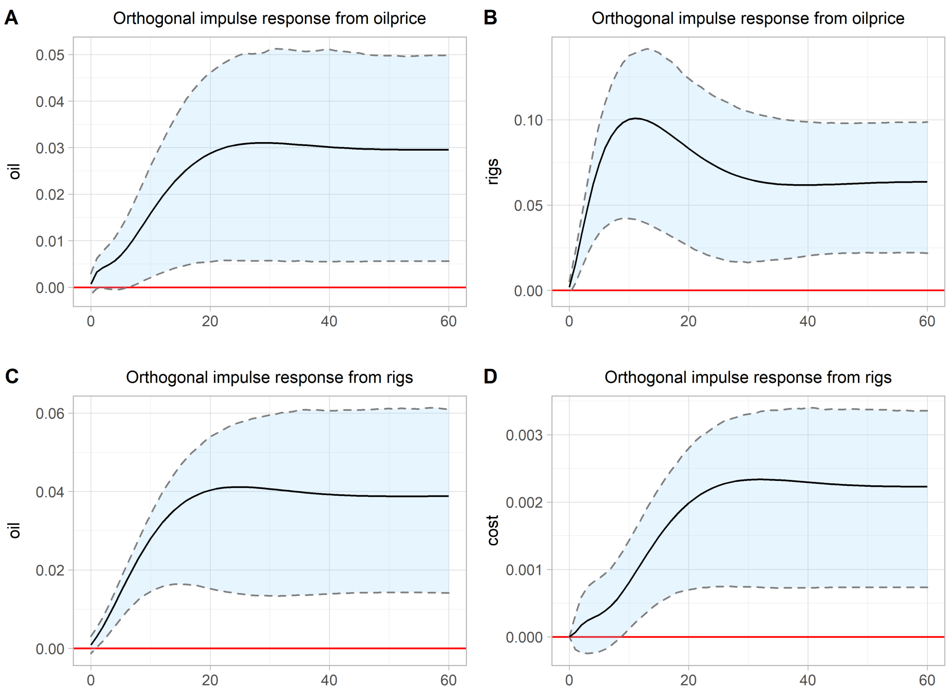

Figure 7: Orthogonal impulse response functions. 95% bootstrap CI, 5000 runs.

tions (OIR) from the VECM. These are obtained using the Cholesky decomposition. That is, the variance-covariance matrix Σ is decomposed such that $\Sigma = PP^{\intercal}$, where P is a lower triangular matrix with positive diagonal elements. A caveat with this approach is that the results are dependent on the ordering of the variables in the VECM. The causality chain assumed in the Cholesky decomposition is oilprice $\rightarrow$ cost $\rightarrow$ rigs $\rightarrow$ oil. So, for example, the oil price will never be sensitive to a contemporaneous shock in any other variable, whereas oil will be sensitive to shocks of all other variables. The estimated Cholesky matrix is given in Appendix C. Fortunately, model experimentation suggests that the OIRs are not very sensitive to the ordering of the variables. Specifically, model formulations where the control variables enters first, i.e., rigs, oil, oilprice/cost, only result in small differences from those presented in Figure 7. One reason for this may be the use of monthly data. Note that the magnitude of the exogenous transitory shocks are not equal in Figures 5 and 7, because of the Cholesky decomposition.

Figure 8 graphs the forecast error variance decomposition (FEVD) of the model. The FEVD, which is also based on the orthogonalized impulse response coefficient matrices in this case, indicates how much of each the endogenous variables contribute to the forecast error variance of the other variables in the VECM. We see that the oil price is a key determinant. Remember that this paper does not try to explain the development of the

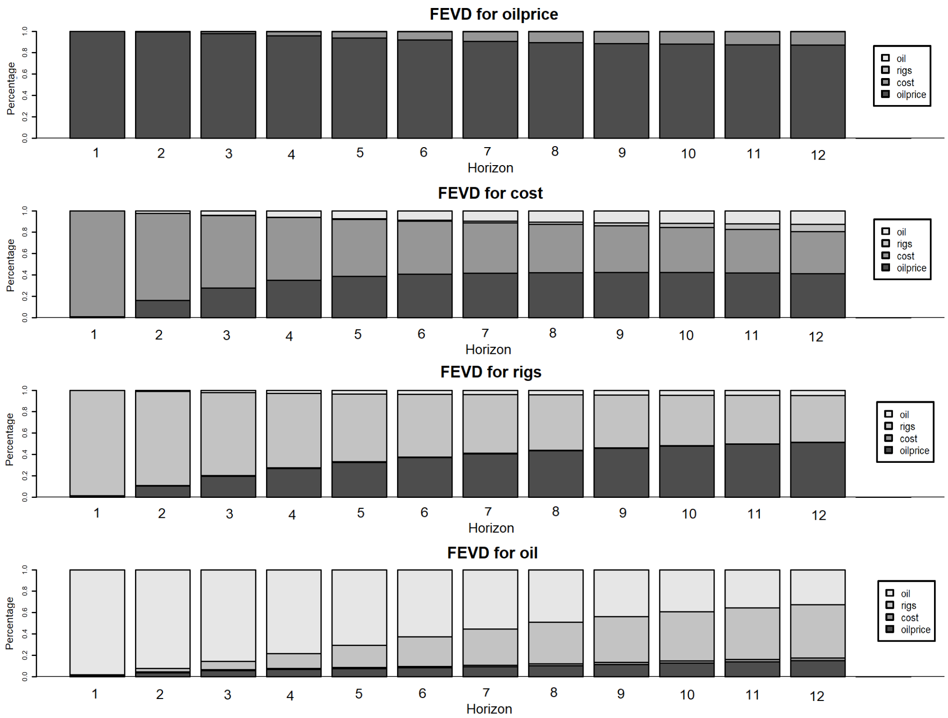

Figure 8: Forecast error variance decomposition (FEVD) of the model.

oil price. The oil price reflects a plethora of factors that indirectly affect the residual demand for U.S. LTO production, e.g., GDP fluctuations, OPEC supply decisions and various other international events.

The theory suggests that changes in costs and the oil price should have a similar effects in the long-run. One way to examine this is to test whether the absolute values on the coefficients for cost and oil price are equal in the cointegrating relations matrix β. This test is performed in R using the blrtest. The test rejects the null hypothesis of equal absolute values on the coefficients with a p-value below 0.01 (see Appendix C for details). This result may not be that surprising, given the estimated VECM in Table 3 and Figure 8. Nevertheless, the fact that costs are so unimportant compared to the oil price is perhaps a little surprising in light of the theory model. This may, e.g., be because our proxy for costs does not capture the whole cost picture, because of the role expectations play in the oil market (e.g. the expected persistence of current changes in the oil price versus current changes in costs), or because the theory model exaggerates the role costs have relative to oil prices.

I also test for weak exogeneity, i.e., whether a particular variable is independent of the cointegrating relations in the long run, by testing whether its coefficient in the loading matrix α is zero. This is performed in R using the alrtest. The tests reject the null-hypothesis of weak exogeneity for the variables cost, rigs and oil at the 5% level of significance (p-values below 0.01), but the null-hypothesis that the oil price is weakly exogenous is not rejected (the p-value is 0.12).[41] It is reasonable to assume that the weak exogenity of the oil price reflects all the other variables that affect global oil prices not captured in this model. That is, because this paper focuses on the supply of LTO and how it responds to the oil price, key determinants for the oil price itself is not included in the model (e.g., variables capturing demand, or OPEC policy).

4 Concluding remarks

This paper examined the supply of U.S. LTO both from a theoretical and empirical point of view. The implications from the theory model are essentially consistent with the results from the empirical analysis. We note, however, that resource scarcity does not seem to be a driving factor for U.S. LTO in the period covered by the data base. This is not so surprising considering the large U.S. LTO resource base, and that large-scale LTO production is a fairly young industry.

The theoretical model emphasized that the effect of changed oil prices on oil production levels largely depends on changes in rig activity. This is supported by the results from the econometric model. Specifically, the rig market reacts both faster and stronger to changes in the oil price than oil production itself. Furthermore, oil production depends positively on rig activity. The results in the present paper hence highlights the importance of seeing oil production and rig activity in context. We also observe that the response to changes in the oil price varies markedly over time.

A key research question in the literature is whether the rise of LTO and unconventional oil can contribute to stabilizing oil prices (Bornstein et al., 2018; Balke et al., 2020; Vatter et al., 2022)). The estimated VECM model indicates that LTO production is responsive to price changes, which implies that LTO can dampen price volatility given that conventional oil is less responsive. On the other hand, I did not find a significant effect from U.S. LTO production on the WTI oil price. This is perhaps not very surprising, given the volatility in oil prices over the data period. In this context, it is worth noting that the share of LTO in global oil production has grown

[41] See Appendix C for details. The final model includes oil price as an endogenous variable, but models with an exogenous oil price have been tested. Oil price endogeneity is often tackled using instrument variables in the literature (Davis and Kilian, 2011; Coglianese et al., 2017; Newell and Prest, 2019). Note that the Granger tests in Table 1 indicate that none of the (endogenous) variables, including the oil price, are strongly exogenous (strong exogeneity requires weak exogeneity and no Granger-causality, see Engle et al., 1983). An auxiliary test using a bivariate VAR consisting of the first-differences of oilprice and oil rejects the null-hypothesis that oil does not Granger-cause oilprice at a 5% level of significance (p-value 0.02, see Appendix C).

significantly over the (pre-Covid-19) data sample period.

Besides various combinations of dummy variables and exogenous explanatory variables, I have tested for models where *(i)* the oil price is exogenous, *(ii)* cost is exogenous, and *(iii)* cost and oil price are exogenous. Neither of these models performed as well as the final model, nor did models with a larger number of lags.

Last, as a final caveat, there may have been a change in the behavior of U.S. LTO production after the Covid-19 pandemic. As pointed out by e.g. the International Energy Agency (IEA, 2021) and Wood Mackenzie (2020), LTO operators today are under extreme duress from banks and shareholders simply to generate free cash flow, and it is not obvious that LTO will return to growth paths like we have seen the last decade.[42] Hence, it is conceivable that the future price elasticity of U.S. LTO will turn out to be lower than during the sample period used in this paper.

Acknowledgements

I am grateful to Terje Skjerpen, Thomas S. Gundersen, Lars Lindholt and Knut Einar Rosendahl for valuable comments to earlier drafts.

References

Aastveit, K.A., Bjørnland, H.C., and Gundersen, T. (2022): The price responsiveness of shale producers: Evidence from micro data, Norges Bank research working paper 10/2022. (https://www.norges-bank.no/aktuelt/nyheter-og-hendelser/Signerte-publikasjoner/Working-Papers/2022/wp102022/)

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

Aune, F.R., Mohn, K., Osmundsen, P., and Rosendahl, K.E. (2010): Financial market pressure, tacit collusion and oil price formation, Energy Economics 32(2), 389–398.

Balke, N.S., Jin, X., and Yüce, M. (2020): The shale revolution and the dynamics of the oil market, Federal Reserve Bank of Dallas Working Paper no. 2021, https://doi.org/10.24149/wp2021.

Bjørnland, H.C., Nordvik, F.M., and Rohrer, M. (2020): Supply flexibility in the shale patch: Evidence from North Dakota, Journal of Applied Econometrics, 36, 273–292.

[42]'In the Stated Policies Scenario, tight oil operators choose to prioritize returns over aggressive production growth, even as annual average prices rise to 2030. Tight oil production satisfies around 20% of global oil demand growth between 2020 and 2030 (compared with the 2010-2019 period when it provided 70%)' (IEA, 2021, p. 217). See also oilprice.com and energyintel.com.

Bornstein, G., Krusell, P., and Rebelo, S. (2018): Lags, costs and shocks: An equilibrium model of the oil industry, NBER Working Paper 23423, (http://www.nber.org/papers/w23423)

Carter Hill, R., Griffiths, W.E., and Judge, G.G. (2002): Undergraduate Econometrics (2rd ed.), John Wiley and Sons, Inc.

Coglianese, J., Davis, L.W., Kilian, L. and Stock, J.H. (2017): Anticipation, tax avoidance, and the price elasticity of gasoline demand, Journal of Applied Econometrics, 32(1), 1–15.

Dahl, C. and Yücel, M. (1991): Testing alternative hypotheses of oil producer behavior, The Energy Journal, 12(4), 117–138.

Davis, L.W., and Kilian, L. (2011): Estimating the effect of a gasoline tax on carbon emissions, Journal of Applied Econometrics, 26(7), 1187–1214.

Engle, R.F., Hendry, D.F., and Richard, J-F. (1983): Exogeneity, Econometrica 51(2), 277–304.

Farzin, Y.H. (2001): The impact of oil prices on additions to US proven reserves, Resource and Energy Economics 23(3), 271–291.

Fattouh, B. and Sen, A. (2013): The US tight oil revolution in a global perspective, Oxford Energy Comment, The Oxford Institute for Energy Studies. (https://www.oxfordenergy.org/publications/the-us-tight-oil-revolution-in-a-global-perspective/)

Foroni, C., and Stracca, L. (2022): The shale oil revolution and the global oil supply curve, forthcoming in Journal of Applied Econometrics.

Golombek, R., Irarrazabal, A.A., and Ma, L. (2018): OPEC's market power: An empirical dominant firm model for the oil market, Energy Economics, 70, 98–115.

Greene, W.H. (2003): Econometric analysis (5th ed.), Prentice Hall, Pearson Education International.

Gundersen, T.S. (2020): The impact of U.S. supply shocks on the global oil price, The Energy Journal 41(1), 151–174.

Hamilton, J.D. (1994): Time series analysis, Princeton University Press, Princeton, New Jersey.

IEA (2021): World Energy Outlook 2021, International Energy Agency. https://iea.blob.core.windows.net/assets/4ed140c1-c3f3-4fd9-acae-789a4e14a23c/WorldEnergyOutlook2021.pdf.

Johansen, S., and Juselius, K. (1990): Maximum likelihood estimation and inference on cointegration – with applications to the demand for money, Oxford Bulletin of Economics and Statistics, 52(2), 169–210.

Johansen, S. (1991): Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models, Econometrica 59 (6), 1551–1580.

Johansen, S (1995): Likelihood-based inference in cointegrated vector autoregressive models, Oxford University Press.

Kilian, L. (2016): The impact of the shale oil revolution on US oil and gasoline prices, Review of Environmental Economics and Policy 10(2), 185–205.

Kilian, L. (2017): The impact of the fracking boom on Arab oil producers, The Energy Journal 38(6), 137–160.

Kilian (2020): Understanding the estimation of oil demand and oil supply elasticities, CFS Working Paper, No. 649, 2020.

Kleinberg, R.L., Paltsev, S., Ebinger, C.K.E, Hobbs, D.A., and Boersma, T. (2018): Tight oil market dynamics: Benchmarks, breakeven points, and inelasticities, Energy Economics, 70, 70–83.

Newell, R.G., and Prest, B.C. (2019): The unconventional oil supply boom: Aggregate price response from microdata, The Energy Journal, 40(3), 1–30.

Nguyen, H.T., and Nabney, I.T. (2010): Short-term electricity demand and gas price forecasts using wavelet transforms and adaptive models, Energy 35(9), 3674–3685.

Osmundsen, P., Roll, K.H. and Tveterås, R. (2010): Drilling speed - the relevance of experience, Energy Economics 34 (3), 786–794.

Osmundsen, P., Rosendahl, K.E., and Skjerpen, T. (2015): Understanding rig rate formation in the Gulf of Mexico, Energy Economics 49, 430–439.

Pesaran, M.H. (1990): An econometric analysis of exploration and extraction of oil in the UK continental shelf, The Economic Journal, 100(401) 367–390.

Ramcharran, H. (2002): Oil production responses to price changes: An empirical application of the competitive model to OPEC and non-OPEC countries, Energy Economics, 24(2), 97–106.

Reitz, S., Rülke, J.-C., and Stadtmann, G. (2009): Are oil price forecasters finally right? Regressive expectations toward more fundamental values of the oil price, Discussion Paper Series 1: Economic Studies 2009,32, Deutsche Bundesbank. (https://ideas.repec.org/p/zbw/bubdp1/200932.html)

Sinclair, P. (1992): High does nothing and rising is worse: carbon taxes should be kept declining to cut harmful emissions, Manchester School 60, 41–52.

Sinn, H.W. (2008): Public policies against global warming: A supply side approach, International Tax and Public Finance. Vol. 15(4), 360–94.

Skjerpen, T., Rosendahl, K.E., Osmundsen, P., and Storrøsten, H.B. (2018). Modelling and forecasting rig rates on the Norwegian Continental Shelf, Resource and Energy Economics 53, 220–239.

Smith, J. L. (2009): World oil: Market or mayhem? The Journal of Economic Perspectives, 23(3), 145–164.

Sydsæter, K., Hammond, P., Seierstad, A., and Strøm, A. (2008): Further mathematics for economic analysis, Pearson Education Limited, ISBN: 9780273713289.

Vatter, M.H., VanVactor, S.A., and T.C. Coburn (2022): Price responsiveness of shale oil: A Bakken case study, Natural Resources Research, 31(1), 713–734.

Wethe, D. (2019, Dec. 27): The Shale Revolution, Bloomberg Quicktake, https://www.bloomberg.com/quicktake/fracking?leadSource=uverify

Wood Mackenzie (2020): Five constraints on US tight oil production. Can the sector return to a growth path? Wood Mackenzie, Global Energy Research. https://www.woodmac.com/news/the-edge/five-constraints-on-us-tight-oil-production/.

Appendices

A Appendix: Proofs and derivations

Proof of Lemma 1: We first observe that the integral in (6) converges under our assumptions of a finite choke price $\bar{p}$. The shadow prices λ and μ cannot exceed the finite choke price $\bar{p}$ (in fact, they are below $\bar{p}$ whenever costs are positive). Further, the stocks R_{it} and S_{it} are finite. The discount factor $r > 0$ is essential for a unique time profile.

The Lagrangian associated with the mixed constraints problem (6) is:

$$H_{it}(x_{it_1}, y_{it_1}, R_{it_1}, S_{it_1}, k_t) = H_i(x_{it_1}, y_{it_1}, R_{it_1}, S_{it_1}, k_t) + \eta_{it}(1 - y_{it}), \quad \text{(A.1)}$$

where $H_i(x_{it_1}, y_{it_1}, R_{it_1}, S_{it_1}, k_t) = \pi_{it} + \lambda_{it}(x_{it} - q_{it}) - \mu_{it}x_{it}$ is the current value Hamiltonian associated with the maximization problem (6). Here λ_{it} and μ_{it} are the shadow prices (or adjoint/co-state variables) on the state variables R_{it} and S_{it}, respectively. The cost function is convex; i.e., the Hessian matrix associated with $c_i(\cdot)$ is negative definite, and all variables are non-negative. It follows that the Lagrangian is a sum of linear and concave functions, and therefore itself concave.[43] The necessary conditions for optimum are given by (see Sydsæter et al., 2008, pp. 360–366):

$$H_{y_{it}} = \pi_{y_{it}}(\cdot) - \omega R_{it}\lambda_{it} - \eta_{it} \leq 0, \quad \text{(A.2)}$$

$$y_{it} \leq 1, \quad \text{and } \eta_{it} = 0 \text{ if } y_{it} < 1, \quad \text{(A.3)}$$

$$H_{x_{it}} = \pi_{x_{it}}(\cdot) + \lambda_{it} - \mu_{it} \leq 0, \quad \text{(A.4)}$$

$$\dot{\lambda}_{it} - \delta\lambda_{it} = -H_{R_{it}}(\cdot) = -\pi_{R_{it}}(\cdot), \quad \text{(A.5)}$$

$$\dot{\mu}_{it} - \delta\mu_{it} = -H_{S_{it}}(\cdot) = -\pi_{S_{it}}(\cdot), \quad \text{(A.6)}$$

$$\lambda_{it_1} \geq 0, \ \mu_{it_1} \geq 0, \quad \text{(A.7)}$$

$$H_{it_1}(\cdot) = 0. \quad \text{(A.8)}$$

[43] Neither the Mangasarian nor the Arrow theorem applies to variable final time problems like (6). Nevertheless, any optimal path must satisfy the necessary conditions given by the system of equations (7)-(12).

This system of equations is equivalent with Equations (1)-(5) in Lemma 1.[44]

Derivation of Equation (14): Let $i, j \in I = (oilprice, cost, rigs, oil)$. Then the j'th equation in the VAR (13) can be written:

$$y_{j,t} = c_j + \sum_{i \in I} (\delta_{ji} y_{i,t-1} + \mu_{ji} y_{i,t-2}) + \epsilon_{j,t} \ , \forall j \in I, \tag{A.9}$$

where δ_{ji} and μ_{ji} denote the coefficients for the first and second order lags, respectively. The term with exogenous variables, $\phi \mathbf{x}_t$, enters Equations (13) and (14) identically and is omitted for simplicity. Note that the long run equilibrium, as characterized by $y_j = y_{j,t} = y_{j,t-1} = y_{j,t-2}$, satisfies:

$$y_j = \frac{1}{(1 - \delta_{jj} - \mu_{jj})} \left(c_j + \sum_{i \in I/\{j\}} (\delta_{ji} y_{i,t-1} + \mu_{ji} y_{i,t-2}) \right) \tag{A.10}$$

for given values on the other $i \in I/\{j\}$ variables. Equation (A.9) can be rewritten to find the j'th equation in the VECM model:

$$\Delta y_{j,t} = \sum_{i \in I} \gamma_{ji} \Delta y_{i,t-1} - \lambda_j \left(y_{j,t-1} - \pi_{j0} - \sum_{i \in I/\{j\}} \pi_{ji} y_{i,t-1} \right) + \epsilon_t, \tag{A.11}$$

where $\gamma_{ji} = -\mu_{ji}$, $\lambda_j = (1 - \delta_{jj} - \mu_{jj})$, $\pi_{j0} = \frac{c_j}{\lambda}$ and $\pi_{ji} = \frac{\delta_{ji} + \mu_{ji}}{\lambda}$. The matrices in Equation (14) follows directly from Equation (A.11).

Note that the term in parenthesis is the deviation from the long run value given in Equation (A.10). The estimation of the VECM model which results are reported in Table 3 does the following: (i) Estimate the four equations in (A.10) by OLS (we do not loose information doing this equation by equation, because the same variables enters in all equations), then (ii) estimate the four equations in (A.11) using OLS, where the square parenthesis is replaced by the residuals from step (i). Note that we found two cointegrating relationships in this model (cf., Section 3.3). Note that a VAR in levels with two lags corresponds to a VECM with one lag (and a Johansen test for reduced rank with two lags), because the VECM features the first-differences of the lagged variables.

[44] The case of a free flow control variable in Anderson et al. (2018), which is arguably a good approximation to much conventional oil production, can be approximated by a cost function $c_i^y(\cdot)$ that is close to zero for $y < 1$, and then jumps steeply as y approaches 1, e.g. the cumulative Cauchy distribution function variation $c_i(y_{it}, k_t) = ((2.01\bar{p}/\pi) arctan((y - 1)/0.001) + 1/2)$.

B Appendix: The numerical illustration

The numerical illustration uses the theory framework from Section 2 with quadratic cost functions and one representative resource extracting firm. The model is formulated as a non-linear programming problem and solved using the Conopt solver in GAMS (GAMS). The GAMS code is supplied in the separate attachment 'Appendix B - GAMS code'.

C Appendix: The econometric model

This appendix presents output from the econometric software R (R) and selected figures. The output is supplied in the separate attachment 'Appendix C: Selected output from R'.

Table 4: The VECM model. Estimated equation by equation using OLS.

	Dependent variable:			
	diff(oilprice)	diff(cost)	diff(rigs)	diff(oil)
	(1)	(2)	(3)	(4)
Cointegrating relationship 1	0.053**	−0.003***	0.007	−0.016***
	(0.026)	(0.001)	(0.007)	(0.004)
Cointegrating relationship 2	4.406**	−0.340***	−0.867	−0.073
	(2.085)	(0.070)	(0.557)	(0.297)
constant	−23.977**	1.850***	4.721	0.407
	(11.345)	(0.378)	(3.031)	(1.615)
diff(oilprice).lagged	0.274**	−0.013***	0.137***	0.045***
	(0.107)	(0.004)	(0.029)	(0.015)
diff(cost).lagged	−0.911	0.158*	1.068	0.098
	(2.545)	(0.085)	(0.680)	(0.362)
diff(rigs).lagged	−0.163	0.004	0.861***	0.060***
	(0.153)	(0.005)	(0.041)	(0.022)
diff(oil).lagged	0.671	0.031	0.331**	0.077
	(0.575)	(0.019)	(0.154)	(0.082)
sd1	0.028	0.003**	−0.018*	−0.009*
	(0.039)	(0.001)	(0.010)	(0.005)
sd2	0.027	0.002	−0.023**	0.009
	(0.042)	(0.001)	(0.011)	(0.006)
sd3	−0.009	0.0003	−0.008	0.015***
	(0.039)	(0.001)	(0.011)	(0.006)
sd4	0.015	0.002*	−0.003	0.002
	(0.041)	(0.001)	(0.011)	(0.006)
sd5	−0.013	0.002	−0.021*	0.010*
	(0.040)	(0.001)	(0.011)	(0.006)
sd6	−0.030	0.002*	0.007	0.005
	(0.041)	(0.001)	(0.011)	(0.006)
sd7	0.009	0.001	−0.001	0.012**
	(0.040)	(0.001)	(0.011)	(0.006)
sd8	−0.039	0.0004	−0.004	0.015***
	(0.039)	(0.001)	(0.011)	(0.006)

Table 4 cont.

	Dependent variable:			
	diff(oilprice)	diff(cost)	diff(rigs)	diff(oil)
	(1)	(2)	(3)	(4)
sd9	0.019	−0.0001	−0.020*	0.014**
	(0.039)	(0.001)	(0.010)	(0.006)
sd10	0.012	−0.0001	−0.008	0.016***
	(0.038)	(0.001)	(0.010)	(0.005)
sd11	−0.017	−0.00004	−0.013	0.010*
	(0.038)	(0.001)	(0.010)	(0.005)
D2011_02	0.002	−0.001	−0.020	−0.049***
	(0.097)	(0.003)	(0.026)	(0.014)
D2020_04	−0.360***	−0.001	−0.229***	−0.071***
	(0.115)	(0.004)	(0.031)	(0.016)
D2020_05	0.761***	−0.005	−0.113***	−0.187***
	(0.131)	(0.004)	(0.035)	(0.019)
D2020_06	0.185	0.015**	0.144***	0.102***
	(0.181)	(0.006)	(0.048)	(0.026)
D2020_07	−0.153	0.002	0.043	0.053***
	(0.122)	(0.004)	(0.033)	(0.017)
D2020_12	0.131	0.00000	−0.052*	−0.008
	(0.098)	(0.003)	(0.026)	(0.014)
D2021_02	0.059	−0.001	−0.005	−0.162***
	(0.099)	(0.003)	(0.026)	(0.014)
D2021_03	0.082	0.008*	0.005	0.174***
	(0.132)	(0.004)	(0.035)	(0.019)
D2021_04	−0.157	−0.010**	−0.064*	−0.002
	(0.137)	(0.005)	(0.037)	(0.019)
D2021_06	0.141	0.003	−0.047*	0.001
	(0.098)	(0.003)	(0.026)	(0.014)
Observations	146	146	146	146
R^2	0.487	0.516	0.915	0.911
Adjusted R^2	0.366	0.401	0.895	0.890
Residual Std. Error (df = 118)	0.090	0.003	0.024	0.013
F Statistic (df = 28; 118)	4.008***	4.486***	45.374***	43.089***
Note:				*p<0.1; **p<0.05; ***p<0.01

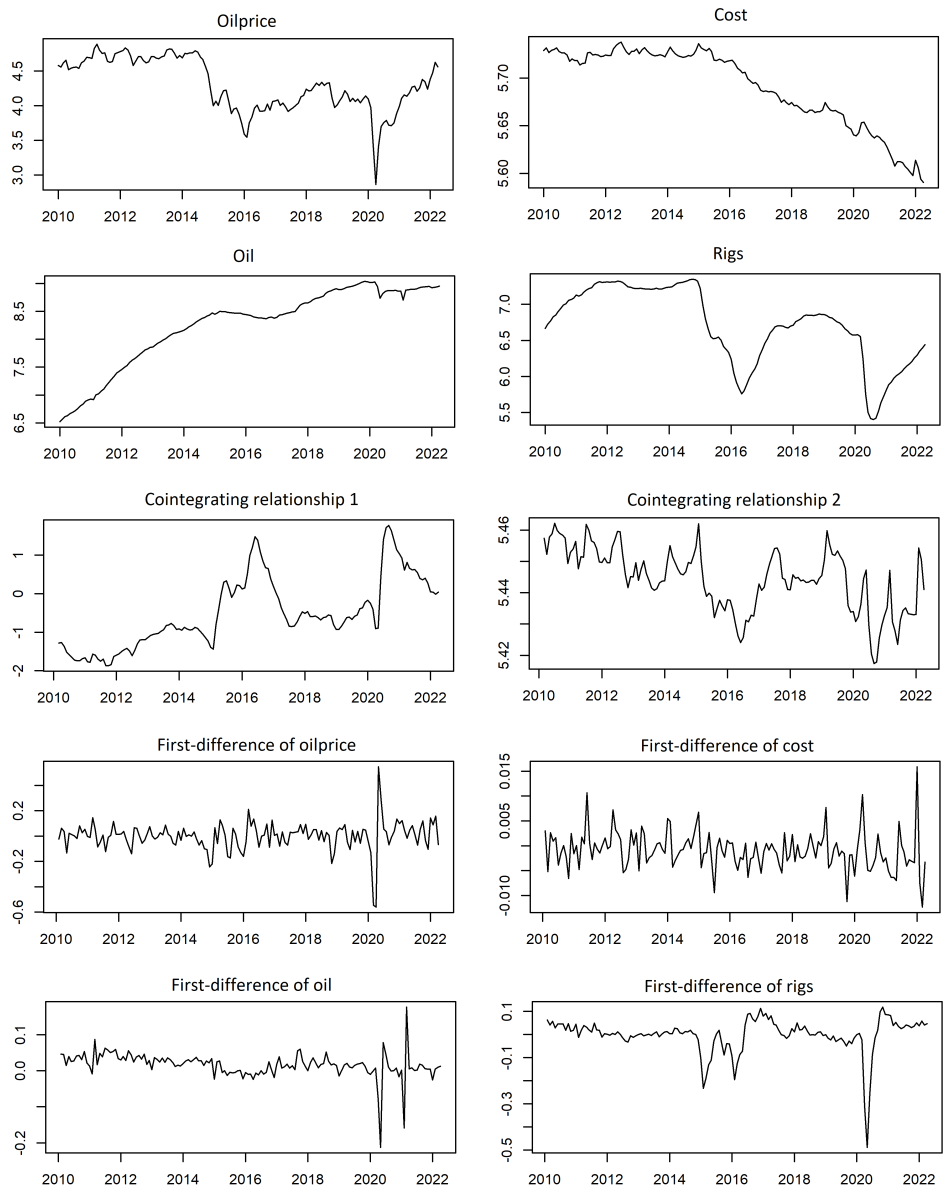

Figure 9: The cointegrating relationships and the log-transformed endogenous variables in levels and first-differences.

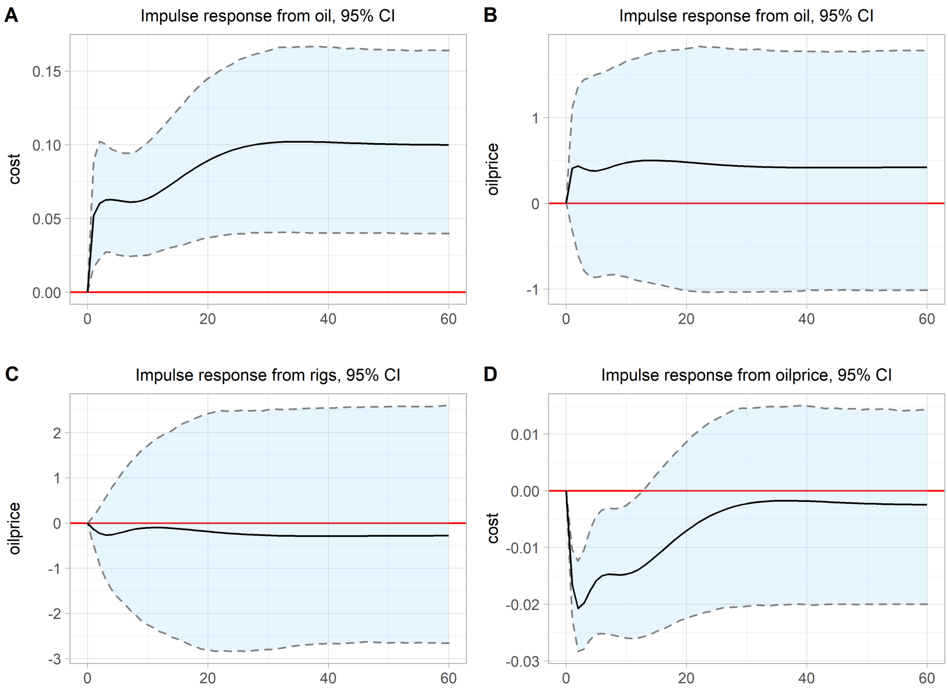

Figure 10: Selected impulse responses. Bootstrapped confidence intervals (5000 runs).

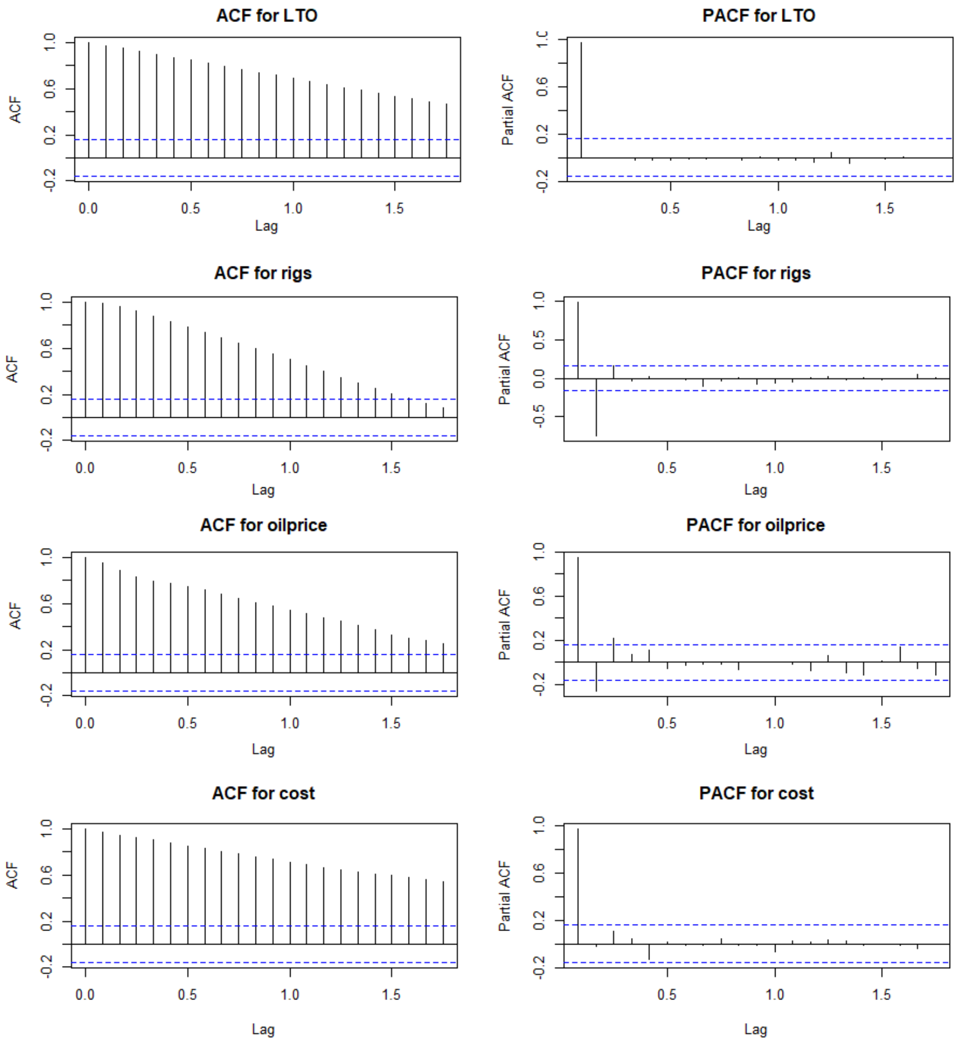

Figure 11: Autocorrelation and partial autocorrelation plots.

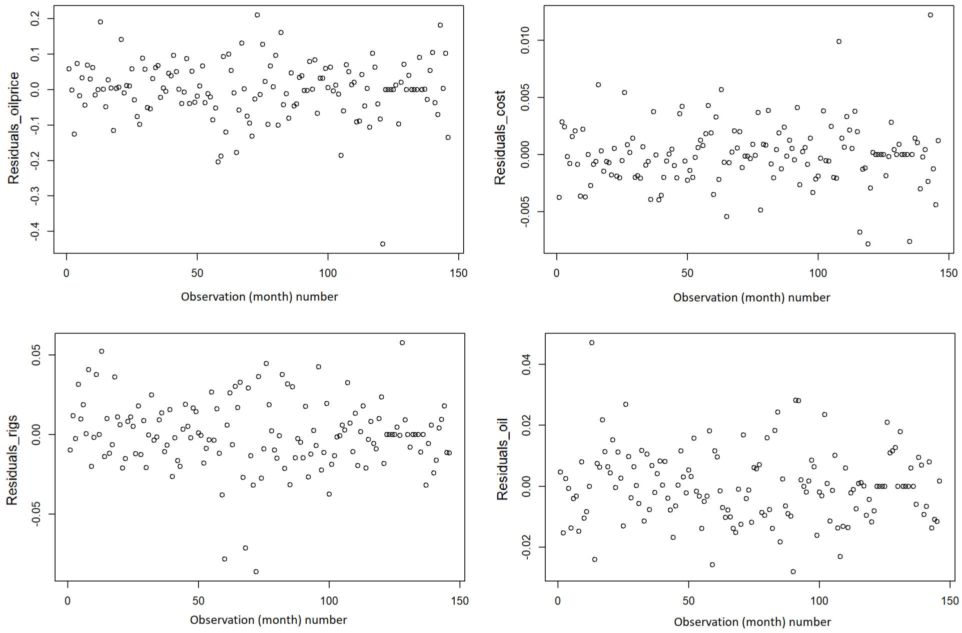

Figure 12: Regression residuals.

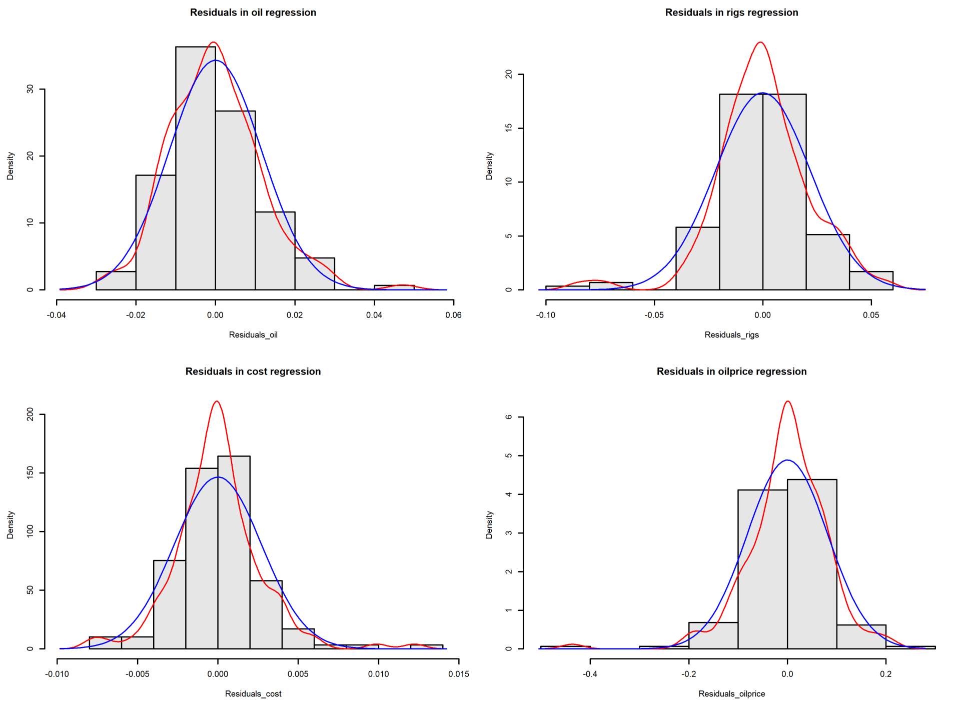

Figure 13: Histograms and kernel densities of residuals plotted against normal distributions with same mean and variance.

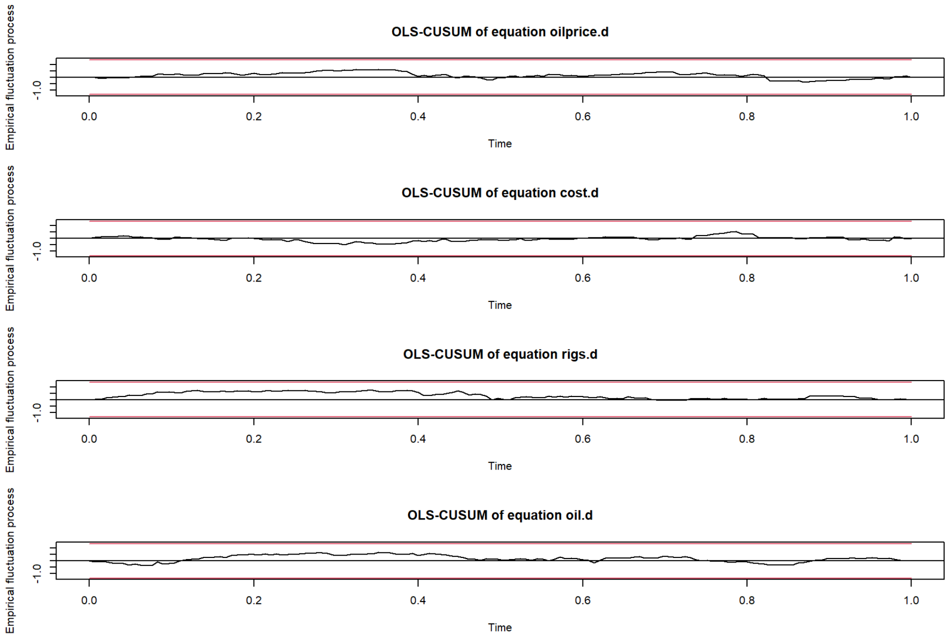

Figure 14: OLS-based CUSUM test for structural stability.

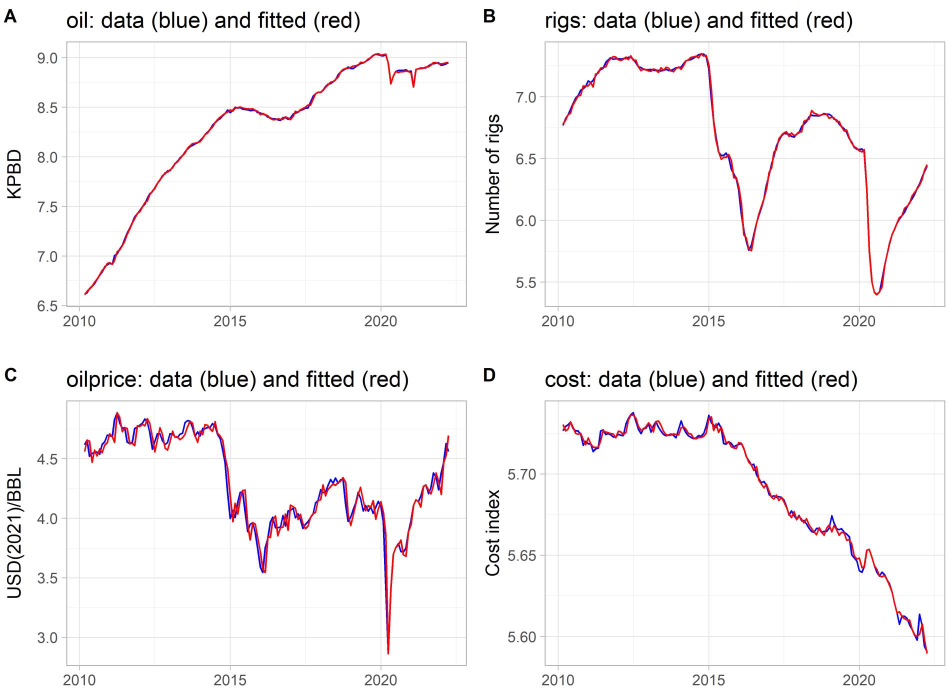

Figure 15: Model fitted and actual data.

```
*   APPENDIX B: GAMS CODE

*   NUMERICAL EXAMPLE OF NON-RENEWABLE RESOURCE MODEL WITH EXTRACTION
    AND RESOURCE DEVELOPMENT

********************************************************************************
* Code by HBS march 2022
* Model is for illustrative purposes only
********************************************************************************

*        SETS AND PARAMETERS

Sets
t                 Projection years                                          /0*200/
t0(t)             Time dummy positive at t=1
tS1(t)            Time dummy for resource discovery for first discovery
tS2(t)            Time dummy for resource discovery for second discovery
tGDP(t)           Time dummy for shift in demand
;
t0(t)=yes$(ord(t) eq 1);
tS1(t)=yes$(ord(t) = 25);
tS2(t)=yes$(ord(t) = 50);
tGDP(t) = yes$(ord(t) > 50);

display t0, tS1, tS2, tGDP;

Scalars
pK                Choke price given GDP=1 #100                              /100/
pp                How fast price declines in quantum #1                     /1/
rr                Discount factor #0.95                                     /0.95/
Cy1               Extraction cost parameter 1 (intercept) #15               /15/
Cy2               Extraction cost parameter 2 (sqr) #2                      /2/
Cy3               Extraction cost parameter 3 (resource R to Q) #0.1        /0.1/
Cx1               Resource development cost parameter 1 #10                 /10/
Cx2               Resource development cost parameter 2 (exp) 1 #2          /2/
Cx3               Development cost parameter 3 (resource S to R) #1         /1/
S0                Initial undeveloped res. stock #25                        /25/
R0                Initial developed resource stock #0                       /0/
DGDP              Shift in demand parameter #0 (#0.25)                      /0.0/
SS                Resource discovery addition to S at time tS  #0 (#10)     /10/
TSS               Techn adds TSS units to S each time period #0 (#0.25)     /0.0/
TCC               Techn reduces cost by 1 div 1+TCC each year #0 (0.05)     /0.0/
omega             Max flow from a unit mass if newly drilled wells #0.05    /0.05/
Cxadj             Adjustment cost change drilling  #10                      /10/
Cyadj             Adjustment cost change production rate #10                /10/
CF                Fixed cost element to drilling and prod #0 (#0.5)         /0.5/
;
* # refer to REF values, (#) refer to sensitivity values

Positive Variables
Q(t)              Production of firm i in period t
S(t)              Undeveloped resourtce stock (not available for production)
R(t)              Developed resourtce stock (available for production)
P(t)              Equilibrium price in period t
x(t)              Resource development of firm i in period t
y(t)              Extraction flow rate
;

Variables
vv(t)             Instantaneous profits
V                 Objective criterion profits over time horizon
C(t)              Cost of production
FCy(t)            Fixed cost flow rate u
FCx(t)            Fixed cost drilling x
;

Parameter
SS0(t)            Stock S in period 1
SR0(t)            Stock R in period 1
```

```
GDP(t)            Demand parameter
DS(t)             Shift in stock S (one time)
TTSS(t)           Shift in stock from tech (each period)
;
SS0(t) = t0(t)*S0;
SR0(t) = t0(t)*R0;
GDP(t) = 1 + tGDP(t)*DGDP;
DS(t) = SS*(tS2(t))
;
Display SS0, SR0, GDP, DS;
;
y.lo(t) = 0.00001; y.up(t) = 0.999;
x.lo(t) = 0.00001;
* limit extraction flow between 0 and 1 and resource development non-negative

*******************************************************************************

*        EQUATIONS AND MODEL FORMULATION

Equations
EqS(t)            State movement equation for undeveloped stock S
EqR(t)            State movement equation for developed stock R
EqP(t)            Equilbrium price inverse demand function
EqQ(t)            Equation for production
Cost(t)           Cost function
Cost_Fy(t)        Fixed cost production
Cost_Fx(t)        Fixed cost drilling
Profit(t)         Instantaneos profits
Sumvv             Objective criterion profits over time horizon
;
EqS(t)..          S(t) =e= SS0(t) + S(t-1) - x(t-1) + DS(t) + TSS;
EqR(t)..          R(t) =e= SR0(t) + R(t-1) - Q(t-1) + x(t-1);
EqP(t)..          P(t) =e= ( pk*exp(-pp*Q(t)) ) * GDP(t);
EqQ(t)..          Q(t) =e= omega * y(t) *R(t);
Cost(t)..         C(t) =e= ( Cy1*y(t) + Cy2 * (y(t)**2) + Cy3 * ( y(t)/(0.01+R(t)) )
    + Cx1*x(t) + Cx2 * (x(t)**2) +  Cx3 * ( x(t)/(0.01+S(t)) ) +
    Cyadj*power(y(t)-y(t-1),2)
    + Cxadj*power(x(t)-x(t-1),2) ) / (1+ord(t)*TCC) +FCy(t) + FCx(t);
Cost_Fy(t)..      FCy(t) =e= CF * ( (1/3.14159265359) * arctan( ((y(t)-0.001)/0.0001)
) + 0.5);
Cost_Fx(t)..      FCx(t) =e= CF * ( (1/3.14159265359) * arctan( ((x(t)-0.001)/0.0001)
) + 0.5);
Profit(t)..       vv(t) =e= (P(t)*Q(t)-C(t)) * 1;
Sumvv..           V =e= sum(t, (rr**(ord(t)))*vv(t));
;

*create GDX point file with the marginals and levels for the variables and equations.
option Savepoint=1;
*load GDX file with marginals and levels from previous model run
execute_loadpoint 'LTO_p';

Model LTO /EqS, EqR, EqP, EqQ, Cost, Cost_Fy, Cost_Fx, Profit, Sumvv/;
solve LTO using nlp max V;

*Error checking
Display LTO.modelstat, LTO.solvestat;
ABORT$(LTO.modelstat <> 2) "Model not normally completed", LTO.modelstat;
ABORT$(LTO.solvestat <> 1) "No optimum found", LTO.solvestat;

Display P.l, Q.l, x.l, y.l, C.l, R.l, S.l, vv.l, FCy.l, FCx.l;

*******************************************************************************

Parameter
AQ                Accumulated production
AR                Accumulated development
Mcy(t)            Marginal cost of oil flow rate y
Mcx(t)            Marginal cost of drilling x
lambda(t)         Shadow price R
```

```
mu(t)             Shadow price S
lambda_q(t)       Shadow price R normalized from y to q
mu_q(t)           Shadow price S normalized from y to q
test(t)           Fixed cost production
;
AQ = sum(t, Q.l(t));
AR = sum(t, x.l(t));
Mcy(t) = ( Cy1 + Cy2 * 2 * y.l(t) + Cy3 * ( 1/(0.01+R.l(t)) ) +
Cyadj*2*(y.l(t)-y.l(t-1)) )
    / (1+ord(t)*TCC);
Mcx(t) = ( Cx1 + Cx2 * 2 * x.l(t) + Cx3 * ( 1/(0.01+S.l(t)) ) +
Cxadj*2*(x.l(t)-x.l(t-1)) )
    / (1+ord(t)*TCC);
lambda(t) = ( p.l(t) - Mcy(t) )/(R.l(t)*omega+0.001) ;
mu(t) = -Mcx(t) + lambda(t);
lambda_q(t) = lambda(t)*(R.l(t)*omega+0.001);
mu_q(t) = -Mcx(t) + lambda_q(t);
;
Display
AQ, AR, Mcy, Mcx, lambda, mu, lambda_q,  mu_q;
;
```

```
##################################################################################

# APPENDIX C: SELECTED OUTPUT FROM R

##################################################################################


###########################
# SUMMARY STATISTICS
###########################

      WTI           Cost_index      Rig_count          LTO
 Min.   : 17.51   Min.   :267.9   Min.   : 222.0   Min.   : 681
 1st Qu.: 55.63   1st Qu.:288.5   1st Qu.: 585.8   1st Qu.:2665
 Median : 71.09   Median :303.2   Median : 880.0   Median :4648
 Mean   : 79.06   Mean   :296.5   Mean   : 924.7   Mean   :4634
 3rd Qu.:106.66   3rd Qu.:306.4   3rd Qu.:1361.0   3rd Qu.:7076
 Max.   :132.44   Max.   :310.3   Max.   :1549.0   Max.   :8390


###########################
# CORRELATION MATRIX
###########################

                  WTI Cost_index  Rig_count        LTO
WTI         1.0000000  0.5121388  0.8271627 -0.6835030
Cost_index  0.5121388  1.0000000  0.6827883 -0.8387674
Rig_count   0.8271627  0.6827883  1.0000000 -0.6233021
LTO        -0.6835030 -0.8387674 -0.6233021  1.0000000


###########################
# TEST FOR UNIT ROOT
###########################

# OILPRICE

Phillips-Perron Unit Root Test
data:  oilprice
Dickey-Fuller Z(alpha) = -9.8863, Truncation lag parameter = 4, p-value = 0.5468
alternative hypothesis: stationary

Phillips-Perron Unit Root Test
data:  diff(oilprice)
Dickey-Fuller Z(alpha) = -88.232, Truncation lag parameter = 4, p-value = 0.01
alternative hypothesis: stationary

# COST

Phillips-Perron Unit Root Test
data:  cost
Dickey-Fuller Z(alpha) = -2.0614, Truncation lag parameter = 4, p-value = 0.966
alternative hypothesis: stationary

Phillips-Perron Unit Root Test
data:  diff(cost)
Dickey-Fuller Z(alpha) = -114.31, Truncation lag parameter = 4, p-value = 0.01
alternative hypothesis: stationary

# RIGS

Phillips-Perron Unit Root Test
data:  rigs
Dickey-Fuller Z(alpha) = -10.137, Truncation lag parameter = 4, p-value = 0.5324
alternative hypothesis: stationary

Phillips-Perron Unit Root Test
data:  diff(rigs)
Dickey-Fuller Z(alpha) = -34.919, Truncation lag parameter = 4, p-value = 0.01
```

```
alternative hypothesis: stationary

# OIL

Phillips-Perron Unit Root Test
data:  oil
Dickey-Fuller Z(alpha) = -2.73, Truncation lag parameter = 4, p-value = 0.9454
alternative hypothesis: stationary

Phillips-Perron Unit Root Test
data:  diff(oil)
Dickey-Fuller Z(alpha) = -158.1, Truncation lag parameter = 4, p-value = 0.01
alternative hypothesis: stationary

# Comment: Augmented Dickey-Fuller test results are available from the author on request

###############################################
# lag selection criteria #
###############################################

$selection
AIC(n)  HQ(n)  SC(n) FPE(n)
     2      2      2      2

$criteria
                 1             2             3         4
AIC(n) -3.072481e+01 -3.187764e+01 -3.185368e+01 -3.174357e+01
HQ(n)  -3.017313e+01 -3.118804e+01 -3.102616e+01 -3.077812e+01
SC(n)  -2.936724e+01 -3.018068e+01 -2.981733e+01 -2.936782e+01
FPE(n)  4.552030e-14  1.443066e-14  1.487053e-14  1.674418e-14
          5             6             7
AIC(n)  -3.162733e+01 -3.162219e+01 -3.150056e+01
HQ(n)   -3.052396e+01 -3.038091e+01 -3.012136e+01
SC(n)   -2.891219e+01 -2.856766e+01 -2.810665e+01
FPE(n)   1.902618e-14  1.941332e-14  2.234576e-14
                 8             9            10
AIC(n) -3.157133e+01 -3.146017e+01 -3.150853e+01
HQ(n)  -3.005421e+01 -2.980512e+01 -2.971556e+01
SC(n)  -2.783802e+01 -2.738747e+01 -2.709643e+01
FPE(n)  2.131831e-14  2.452633e-14  2.420268e-14


###############################################
# LR-test for no linear trend #
###############################################

H0: H*2(r<=2)
H1: H2(r<=2)
Test statistic is distributed as chi-square
with 2 degress of freedom
        test statistic p-value
LR test          12.87       0
        test statistic p-value
LR test          12.87       0


###########################
# GRANGER CAUSALITY TESTS
###########################

# Comment: Below the term 'dX.l1' refers to the first-differrence of
# the lag of X

# THE VAR IN FIRST-DIFFERENCES USED FOR GRANGER CAUSALITY TEST
#(AIC used as lag criterion):

=====================================================================
                                        Dependent variable:
```

```
                                  ----------------------------------------
                                                    y
                                   (1)       (2)       (3)       (4)
----------------------------------------------------------------------
doilprice.l1                       0.236*** -0.013***  0.285***  0.140***
                                  (0.083)   (0.003)   (0.024)   (0.025)

dcost.l1                           2.298     0.026     -0.344    0.558
                                  (2.485)   (0.081)   (0.728)   (0.735)

drigs.l1                          -0.287**  -0.001     0.815***  0.082**
                                  (0.128)   (0.004)   (0.037)   (0.038)

doil.l1                           -0.367     0.012    0.167**    0.081
                                  (0.270)   (0.009)   (0.079)   (0.080)

const                              0.008   -0.001***   -0.004    0.016***
                                  (0.010)  (0.0003)   (0.003)   (0.003)

----------------------------------------------------------------------
Observations                        146       146       146       146
R2                                 0.135     0.170     0.828     0.211
Adjusted R2                        0.111     0.147     0.823     0.189
Residual Std. Error (df = 141)     0.108     0.003     0.031     0.032
F Statistic (df = 4; 141)         5.519*** 7.245***  169.186*** 9.438***
======================================================================
Note:                                         *p<0.1; **p<0.05; ***p<0.01

# THE GRANGER CAUSALITY TESTS:

    Granger causality H0: doilprice do not Granger-cause dcost drigs doil
data:  VAR object GVAR
F-Test = 60.675, df1 = 3, df2 = 564, p-value < 2.2e-16
    H0: No instantaneous causality between: doilprice and dcost drigs doil
data:  VAR object GVAR
Chi-squared = 9.6602, df = 3, p-value = 0.02169

    Granger causality H0: drigs do not Granger-cause doilprice dcost doil
data:  VAR object GVAR
F-Test = 2.8481, df1 = 3, df2 = 564, p-value = 0.03694
    H0: No instantaneous causality between: drigs and doilprice dcost doil
data:  VAR object GVAR
Chi-squared = 1.5127, df = 3, p-value = 0.6793

    Granger causality H0: doil do not Granger-cause doilprice dcost drigs
data:  VAR object GVAR
F-Test = 2.6733, df1 = 3, df2 = 564, p-value = 0.04662
    H0: No instantaneous causality between: doil and doilprice dcost drigs
data:  VAR object GVAR
Chi-squared = 6.5704, df = 3, p-value = 0.08693

    Granger causality H0: dcost do not Granger-cause doilprice drigs doil
data:  VAR object GVAR
F-Test = 0.71064, df1 = 3, df2 = 564, p-value = 0.5459
    H0: No instantaneous causality between: doilprice and dcost drigs doil
data:  VAR object GVAR
Chi-squared = 9.6602, df = 3, p-value = 0.02169

# A VAR WITH JUST OIL AND OILPRICE (in first-differences)
# check if oil Granger-cause oilprice which is not rejected as weakly exogenous
# below

===========================================================
                                    Dependent variable:
                                  ----------------------------
                                               y
                                      (1)            (2)
-----------------------------------------------------------
doilprice.l1                        0.227***      0.131***
```

```
                                   (0.080)        (0.024)

doil.l1                            -0.601**       0.130*
                                   (0.258)        (0.076)

const                               0.010       0.014***
                                   (0.010)        (0.003)

------------------------------------------------------------
Observations                         146           146
R2                                  0.096         0.184
Adjusted R2                         0.083         0.172
Residual Std. Error (df = 143)      0.109         0.032
F Statistic (df = 2; 143)         7.586***      16.075***
============================================================
Note:                            *p<0.1; **p<0.05; ***p<0.01

    Granger causality H0: doil do not Granger-cause doilprice
data:  VAR object GVAR2
F-Test = 5.4453, df1 = 1, df2 = 286, p-value = 0.02031
    H0: No instantaneous causality between: doil and doilprice
data:  VAR object GVAR2
Chi-squared = 6.5905, df = 1, p-value = 0.01025


###########################
# JOHANSEN PROCEDURE
###########################

# Comment: Below the term 'X.l1' refers to the lag of X, and 'X.d' denotes
# the first-difference of X

Test type: trace statistic , with linear trend in cointegration

Eigenvalues (lambda):
[1]  4.528142e-01  2.094245e-01  1.284459e-01  2.719755e-02 -2.071168e-17

Values of teststatistic and critical values of test:

           test 10pct  5pct  1pct
r <= 3 |   4.03 10.49 12.25 16.26
r <= 2 |  24.10 22.76 25.32 30.45
r <= 1 |  58.41 39.06 42.44 48.45
r = 0  | 146.44 59.14 62.99 70.05

Eigenvectors, normalised to first column:
(These are the cointegration relations)

             oilprice.l1    cost.l1     rigs.l1      oil.l1      trend.l1
oilprice.l1   1.00000000  1.0000000  1.00000000  1.00000000    1.00000000
cost.l1     -14.64780676 89.0461572 14.11820884 17.04316786 -107.48403615
rigs.l1      -2.78430360  0.6902855 -0.40802995 -0.14837463    3.06952576
oil.l1        2.59010914 -5.3681636 -1.12565526 -0.49509446   -2.62417359
trend.l1     -0.05233468  0.1725679  0.02940773  0.02676193    0.01934258

Weights W:
(This is the loading matrix)

             oilprice.l1      cost.l1      rigs.l1        oil.l1      trend.l1
oilprice.d  0.0032975072  0.050022848  0.058323763 -0.0759358168  1.455157e-14
cost.d      0.0002894651 -0.003774961 -0.001573414 -0.0011808289 -2.374105e-15
rigs.d      0.0141476131 -0.007409814  0.036569301  0.0040043215 -3.436670e-14
oil.d      -0.0130960220 -0.002977733  0.012380709  0.0006477656  1.277881e-14


###########################
# THE VECM MODEL
###########################
```

```
# Comment: below 'ect1' and 'ect2' denotes the cointegrating relationships
# The term 'X.dl1' refers to the first-difference of the lag of X.
# The variables starting with D are dummies, and variables beginning with
# 'sd' are seasonal dummies

Response oilprice.d :

Call:
lm(formula = oilprice.d ~ ect1 + ect2 + constant + D2011_02 +
    D2020_04 + D2020_05 + D2020_06 + D2020_07 + D2020_12 + D2021_02 +
    D2021_03 + D2021_04 + D2021_06 + sd1 + sd2 + sd3 + sd4 +
    sd5 + sd6 + sd7 + sd8 + sd9 + sd10 + sd11 + oilprice.dl1 +
    cost.dl1 + rigs.dl1 + oil.dl1 - 1, data = data.mat)

Residuals:
     Min       1Q   Median       3Q      Max
-0.43605 -0.03985  0.00088  0.05177  0.21114

Coefficients:
               Estimate Std. Error t value Pr(>|t|)
ect1           0.053320   0.025896   2.059  0.04169 *
ect2           4.406041   2.085358   2.113  0.03672 *
constant     -23.976944  11.345050  -2.113  0.03667 *
D2011_02       0.002116   0.096744   0.022  0.98259
D2020_04      -0.359714   0.115072  -3.126  0.00223 **
D2020_05       0.760768   0.130569   5.827 5.01e-08 ***
D2020_06       0.185352   0.181250   1.023  0.30858
D2020_07      -0.153199   0.121875  -1.257  0.21123
D2020_12       0.130501   0.098388   1.326  0.18727
D2021_02       0.059241   0.098958   0.599  0.55055
D2021_03       0.081952   0.131957   0.621  0.53576
D2021_04      -0.156689   0.136717  -1.146  0.25408
D2021_06       0.140898   0.098488   1.431  0.15519
sd1            0.027934   0.038506   0.725  0.46961
sd2            0.026669   0.042383   0.629  0.53042
sd3           -0.008709   0.039331  -0.221  0.82515
sd4            0.015484   0.040878   0.379  0.70553
sd5           -0.013156   0.039998  -0.329  0.74280
sd6           -0.029915   0.040842  -0.732  0.46533
sd7            0.008806   0.040007   0.220  0.82617
sd8           -0.039387   0.039490  -0.997  0.32061
sd9            0.019328   0.038713   0.499  0.61853
sd10           0.011662   0.038301   0.304  0.76130
sd11          -0.017042   0.038222  -0.446  0.65651
oilprice.dl1   0.274193   0.107419   2.553  0.01197 *
cost.dl1      -0.911256   2.545492  -0.358  0.72099
rigs.dl1      -0.163043   0.152578  -1.069  0.28744
oil.dl1        0.670907   0.575489   1.166  0.24605
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.09048 on 118 degrees of freedom
Multiple R-squared:  0.4875,    Adjusted R-squared:  0.3658
F-statistic: 4.008 on 28 and 118 DF,  p-value: 5.986e-08


Response cost.d :

Call:
lm(formula = cost.d ~ ect1 + ect2 + constant + D2011_02 + D2020_04 +
    D2020_05 + D2020_06 + D2020_07 + D2020_12 + D2021_02 + D2021_03 +
    D2021_04 + D2021_06 + sd1 + sd2 + sd3 + sd4 + sd5 + sd6 +
    sd7 + sd8 + sd9 + sd10 + sd11 + oilprice.dl1 + cost.dl1 +
    rigs.dl1 + oil.dl1 - 1, data = data.mat)

Residuals:
       Min         1Q     Median         3Q        Max
-0.0078250 -0.0014402 -0.0000076  0.0012390  0.0122329
```

```
Coefficients:
               Estimate Std. Error t value Pr(>|t|)
ect1         -3.485e-03  8.631e-04  -4.038 9.61e-05 ***
ect2         -3.404e-01  6.951e-02  -4.897 3.12e-06 ***
constant      1.850e+00  3.781e-01   4.892 3.18e-06 ***
D2011_02     -5.973e-04  3.224e-03  -0.185 0.853367
D2020_04     -1.371e-03  3.835e-03  -0.358 0.721320
D2020_05     -5.230e-03  4.352e-03  -1.202 0.231831
D2020_06      1.513e-02  6.041e-03   2.504 0.013634 *
D2020_07      1.656e-03  4.062e-03   0.408 0.684327
D2020_12      3.188e-06  3.279e-03   0.001 0.999226
D2021_02     -1.073e-03  3.298e-03  -0.325 0.745599
D2021_03      7.634e-03  4.398e-03   1.736 0.085238 .
D2021_04     -1.026e-02  4.557e-03  -2.253 0.026133 *
D2021_06      3.046e-03  3.283e-03   0.928 0.355375
sd1           3.014e-03  1.283e-03   2.349 0.020501 *
sd2           1.572e-03  1.413e-03   1.113 0.268122
sd3           2.835e-04  1.311e-03   0.216 0.829178
sd4           2.294e-03  1.362e-03   1.683 0.094924 .
sd5           1.920e-03  1.333e-03   1.440 0.152432
sd6           2.446e-03  1.361e-03   1.797 0.074947 .
sd7           6.894e-04  1.333e-03   0.517 0.606139
sd8           4.067e-04  1.316e-03   0.309 0.757861
sd9          -1.117e-04  1.290e-03  -0.087 0.931133
sd10         -1.378e-04  1.277e-03  -0.108 0.914204
sd11         -3.790e-05  1.274e-03  -0.030 0.976316
oilprice.dl1 -1.301e-02  3.580e-03  -3.635 0.000414 ***
cost.dl1      1.582e-01  8.484e-02   1.865 0.064673 .
rigs.dl1      4.437e-03  5.085e-03   0.872 0.384757
oil.dl1       3.096e-02  1.918e-02   1.614 0.109190
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.003016 on 118 degrees of freedom
Multiple R-squared:  0.5156,    Adjusted R-squared:  0.4007
F-statistic: 4.486 on 28 and 118 DF,  p-value: 4.309e-09


Response rigs.d :

Call:
lm(formula = rigs.d ~ ect1 + ect2 + constant + D2011_02 + D2020_04 +
    D2020_05 + D2020_06 + D2020_07 + D2020_12 + D2021_02 + D2021_03 +
    D2021_04 + D2021_06 + sd1 + sd2 + sd3 + sd4 + sd5 + sd6 +
    sd7 + sd8 + sd9 + sd10 + sd11 + oilprice.dl1 + cost.dl1 +
    rigs.dl1 + oil.dl1 - 1, data = data.mat)

Residuals:
      Min        1Q    Median        3Q       Max
-0.086281 -0.011834 -0.000172  0.011046  0.057840

Coefficients:
               Estimate Std. Error t value Pr(>|t|)
ect1          0.0067378  0.0069177   0.974  0.33205
ect2         -0.8670470  0.5570736  -1.556  0.12228
constant      4.7213072  3.0306675   1.558  0.12195
D2011_02     -0.0195846  0.0258436  -0.758  0.45008
D2020_04     -0.2286275  0.0307399  -7.437 1.80e-11 ***
D2020_05     -0.1128728  0.0348797  -3.236  0.00157 **
D2020_06      0.1435665  0.0484184   2.965  0.00366 **
D2020_07      0.0425918  0.0325572   1.308  0.19334
D2020_12     -0.0518423  0.0262829  -1.972  0.05090 .
D2021_02     -0.0050052  0.0264353  -0.189  0.85015
D2021_03      0.0053978  0.0352505   0.153  0.87856
D2021_04     -0.0636530  0.0365221  -1.743  0.08396 .
D2021_06     -0.0469090  0.0263097  -1.783  0.07716 .
sd1          -0.0177226  0.0102863  -1.723  0.08752 .
sd2          -0.0227320  0.0113221  -2.008  0.04695 *
sd3          -0.0075718  0.0105066  -0.721  0.47254
```

```
sd4           -0.0025595  0.0109201  -0.234  0.81509
sd5           -0.0205342  0.0106849  -1.922  0.05704 .
sd6            0.0070061  0.0109103   0.642  0.52202
sd7           -0.0007843  0.0106873  -0.073  0.94162
sd8           -0.0044582  0.0105492  -0.423  0.67335
sd9           -0.0195112  0.0103417  -1.887  0.06167 .
sd10          -0.0081336  0.0102316  -0.795  0.42824
sd11          -0.0125133  0.0102105  -1.226  0.22281
oilprice.dl1   0.1372046  0.0286956   4.781 5.07e-06 ***
cost.dl1       1.0681550  0.6799916   1.571  0.11890
rigs.dl1       0.8613029  0.0407591  21.132  < 2e-16 ***
oil.dl1        0.3309851  0.1537337   2.153  0.03336 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.02417 on 118 degrees of freedom
Multiple R-squared:  0.915, Adjusted R-squared:  0.8948
F-statistic: 45.37 on 28 and 118 DF,  p-value: < 2.2e-16


Response oil.d :

Call:
lm(formula = oil.d ~ ect1 + ect2 + constant + D2011_02 + D2020_04 +
    D2020_05 + D2020_06 + D2020_07 + D2020_12 + D2021_02 + D2021_03 +
    D2021_04 + D2021_06 + sd1 + sd2 + sd3 + sd4 + sd5 + sd6 +
    sd7 + sd8 + sd9 + sd10 + sd11 + oilprice.dl1 + cost.dl1 +
    rigs.dl1 + oil.dl1 - 1, data = data.mat)

Residuals:
      Min        1Q    Median        3Q       Max
-0.028129 -0.008090  0.000000  0.006654  0.047152

Coefficients:
               Estimate Std. Error t value Pr(>|t|)
ect1          -0.0160738  0.0036869  -4.360 2.80e-05 ***
ect2          -0.0733277  0.2969028  -0.247 0.805356
constant       0.4070573  1.6152511   0.252 0.801473
D2011_02      -0.0486711  0.0137739  -3.534 0.000586 ***
D2020_04      -0.0708750  0.0163834  -4.326 3.19e-05 ***
D2020_05      -0.1874052  0.0185898 -10.081  < 2e-16 ***
D2020_06       0.1019895  0.0258055   3.952 0.000132 ***
D2020_07       0.0534755  0.0173520   3.082 0.002561 **
D2020_12      -0.0076051  0.0140080  -0.543 0.588214
D2021_02      -0.1624927  0.0140892 -11.533  < 2e-16 ***
D2021_03       0.1744309  0.0187874   9.284 9.89e-16 ***
D2021_04      -0.0017776  0.0194651  -0.091 0.927390
D2021_06       0.0008844  0.0140222   0.063 0.949815
sd1           -0.0093236  0.0054823  -1.701 0.091639 .
sd2            0.0093180  0.0060343   1.544 0.125227
sd3            0.0148446  0.0055997   2.651 0.009128 **
sd4            0.0024768  0.0058200   0.426 0.671202
sd5            0.0100303  0.0056947   1.761 0.080773 .
sd6            0.0048972  0.0058148   0.842 0.401384
sd7            0.0117005  0.0056960   2.054 0.042169 *
sd8            0.0152681  0.0056224   2.716 0.007610 **
sd9            0.0140792  0.0055118   2.554 0.011911 *
sd10           0.0157590  0.0054531   2.890 0.004587 **
sd11           0.0102593  0.0054419   1.885 0.061855 .
oilprice.dl1   0.0454821  0.0152938   2.974 0.003566 **
cost.dl1       0.0978957  0.3624143   0.270 0.787539
rigs.dl1       0.0599975  0.0217233   2.762 0.006667 **
oil.dl1        0.0770421  0.0819352   0.940 0.348995
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.01288 on 118 degrees of freedom
Multiple R-squared:  0.9109,    Adjusted R-squared:  0.8898
F-statistic: 43.09 on 28 and 118 DF,  p-value: < 2.2e-16
```

```
Response oilprice.d :

Call:
lm(formula = oilprice.d ~ constant + D2011_02 + D2020_04 + D2020_05 +
    D2020_06 + D2020_07 + D2020_12 + D2021_02 + D2021_03 + D2021_04 +
    D2021_06 + sd1 + sd2 + sd3 + sd4 + sd5 + sd6 + sd7 + sd8 +
    sd9 + sd10 + sd11 + oilprice.dl1 + cost.dl1 + rigs.dl1 +
    oil.dl1 + oilprice.l1 + cost.l1 + rigs.l1 + oil.l1 + trend.l1 -
    1, data = data.mat)

Residuals:
     Min       1Q   Median       3Q      Max
-0.43566 -0.03739  0.00080  0.05033  0.18226

Coefficients:
               Estimate Std. Error t value Pr(>|t|)
constant     -20.878269  12.790049  -1.632  0.10533
D2011_02      -0.028858   0.097421  -0.296  0.76760
D2020_04      -0.384450   0.119769  -3.210  0.00172 **
D2020_05       0.680450   0.142289   4.782 5.18e-06 ***
D2020_06       0.097283   0.188082   0.517  0.60598
D2020_07      -0.179140   0.122366  -1.464  0.14593
D2020_12       0.127726   0.098261   1.300  0.19625
D2021_02       0.057001   0.098482   0.579  0.56386
D2021_03       0.059273   0.133612   0.444  0.65815
D2021_04      -0.138245   0.136406  -1.013  0.31296
D2021_06       0.131467   0.099413   1.322  0.18865
sd1            0.027117   0.038336   0.707  0.48078
sd2            0.027522   0.042166   0.653  0.51525
sd3           -0.012859   0.039334  -0.327  0.74432
sd4            0.014538   0.040985   0.355  0.72344
sd5           -0.014098   0.040207  -0.351  0.72650
sd6           -0.030698   0.041469  -0.740  0.46065
sd7            0.008023   0.040420   0.198  0.84301
sd8           -0.039262   0.039755  -0.988  0.32542
sd9            0.018958   0.038714   0.490  0.62529
sd10           0.011155   0.038142   0.292  0.77046
sd11          -0.015956   0.038042  -0.419  0.67569
oilprice.dl1   0.245702   0.123444   1.990  0.04892 *
cost.dl1      -0.789770   2.618857  -0.302  0.76352
rigs.dl1      -0.272236   0.189059  -1.440  0.15260
oil.dl1        0.529419   0.580289   0.912  0.36350
oilprice.l1    0.035708   0.069807   0.512  0.60996
cost.l1        3.935281   2.334743   1.686  0.09460 .
rigs.l1        0.012818   0.040468   0.317  0.75201
oil.l1        -0.288047   0.138721  -2.076  0.04008 *
trend.l1       0.008143   0.004435   1.836  0.06893 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.08999 on 115 degrees of freedom
Multiple R-squared:  0.5059,    Adjusted R-squared:  0.3727
F-statistic: 3.798 on 31 and 115 DF,  p-value: 9.853e-08


Response cost.d :

Call:
lm(formula = cost.d ~ constant + D2011_02 + D2020_04 + D2020_05 +
    D2020_06 + D2020_07 + D2020_12 + D2021_02 + D2021_03 + D2021_04 +
    D2021_06 + sd1 + sd2 + sd3 + sd4 + sd5 + sd6 + sd7 + sd8 +
    sd9 + sd10 + sd11 + oilprice.dl1 + cost.dl1 + rigs.dl1 +
    oil.dl1 + oilprice.l1 + cost.l1 + rigs.l1 + oil.l1 + trend.l1 -
    1, data = data.mat)

Residuals:
      Min        1Q    Median        3Q       Max
```

```
-0.008106 -0.001563  0.000000  0.001150  0.012663

Coefficients:
               Estimate Std. Error t value Pr(>|t|)
constant      2.084e+00  4.309e-01   4.835 4.16e-06 ***
D2011_02     -3.654e-04  3.282e-03  -0.111  0.91156
D2020_04     -2.399e-03  4.035e-03  -0.595  0.55334
D2020_05     -5.706e-03  4.794e-03  -1.190  0.23638
D2020_06      1.545e-02  6.337e-03   2.438  0.01628 *
D2020_07      1.756e-03  4.123e-03   0.426  0.67101
D2020_12     -2.903e-04  3.311e-03  -0.088  0.93027
D2021_02     -9.304e-04  3.318e-03  -0.280  0.77968
D2021_03      8.668e-03  4.502e-03   1.925  0.05665 .
D2021_04     -1.049e-02  4.596e-03  -2.281  0.02436 *
D2021_06      2.846e-03  3.349e-03   0.850  0.39724
sd1           2.944e-03  1.292e-03   2.280  0.02448 *
sd2           1.588e-03  1.421e-03   1.118  0.26606
sd3           4.569e-04  1.325e-03   0.345  0.73090
sd4           2.508e-03  1.381e-03   1.816  0.07192 .
sd5           2.144e-03  1.355e-03   1.582  0.11632
sd6           2.786e-03  1.397e-03   1.994  0.04848 *
sd7           9.886e-04  1.362e-03   0.726  0.46934
sd8           6.512e-04  1.339e-03   0.486  0.62776
sd9           4.324e-05  1.304e-03   0.033  0.97361
sd10         -7.115e-05  1.285e-03  -0.055  0.95594
sd11         -3.885e-05  1.282e-03  -0.030  0.97587
oilprice.dl1 -1.037e-02  4.159e-03  -2.494  0.01404 *
cost.dl1      1.849e-01  8.823e-02   2.096  0.03830 *
rigs.dl1      7.992e-03  6.370e-03   1.255  0.21216
oil.dl1       3.421e-02  1.955e-02   1.750  0.08279 .
oilprice.l1  -6.240e-03  2.352e-03  -2.653  0.00911 **
cost.l1      -3.827e-01  7.866e-02  -4.865 3.66e-06 ***
rigs.l1      -2.595e-03  1.363e-03  -1.903  0.05955 .
oil.l1        2.337e-02  4.674e-03   5.000 2.07e-06 ***
trend.l1     -7.445e-04  1.494e-04  -4.982 2.24e-06 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.003032 on 115 degrees of freedom
Multiple R-squared:  0.5228,    Adjusted R-squared:  0.3942
F-statistic: 4.065 on 31 and 115 DF,  p-value: 2.097e-08


Response rigs.d :

Call:
lm(formula = rigs.d ~ constant + D2011_02 + D2020_04 + D2020_05 +
    D2020_06 + D2020_07 + D2020_12 + D2021_02 + D2021_03 + D2021_04 +
    D2021_06 + sd1 + sd2 + sd3 + sd4 + sd5 + sd6 + sd7 + sd8 +
    sd9 + sd10 + sd11 + oilprice.dl1 + cost.dl1 + rigs.dl1 +
    oil.dl1 + oilprice.l1 + cost.l1 + rigs.l1 + oil.l1 + trend.l1 -
    1, data = data.mat)

Residuals:
     Min       1Q   Median       3Q      Max
-0.08898 -0.01203  0.00000  0.01044  0.04983

Coefficients:
               Estimate Std. Error t value Pr(>|t|)
constant      1.5934522  3.3188881   0.480  0.63206
D2011_02     -0.0293562  0.0252797  -1.161  0.24794
D2020_04     -0.2170456  0.0310789  -6.984 1.98e-10 ***
D2020_05     -0.1209920  0.0369225  -3.277  0.00139 **
D2020_06      0.1211548  0.0488053   2.482  0.01449 *
D2020_07      0.0359121  0.0317527   1.131  0.26041
D2020_12     -0.0476943  0.0254977  -1.871  0.06395 .
D2021_02     -0.0077181  0.0255552  -0.302  0.76319
D2021_03     -0.0155689  0.0346709  -0.449  0.65424
D2021_04     -0.0565148  0.0353960  -1.597  0.11309
```

```
D2021_06      -0.0455640  0.0257967  -1.766  0.08000 .
sd1           -0.0167619  0.0099478  -1.685  0.09470 .
sd2           -0.0228205  0.0109418  -2.086  0.03922 *
sd3           -0.0111567  0.0102067  -1.093  0.27665
sd4           -0.0061713  0.0106351  -0.580  0.56286
sd5           -0.0242899  0.0104334  -2.328  0.02165 *
sd6            0.0014068  0.0107609   0.131  0.89621
sd7           -0.0057208  0.0104885  -0.545  0.58651
sd8           -0.0083416  0.0103160  -0.809  0.42041
sd9           -0.0220609  0.0100458  -2.196  0.03010 *
sd10          -0.0092986  0.0098975  -0.939  0.34945
sd11          -0.0122855  0.0098716  -1.245  0.21583
oilprice.dl1   0.0894456  0.0320326   2.792  0.00613 **
cost.dl1       0.6653005  0.6795669   0.979  0.32963
rigs.dl1       0.7830993  0.0490588  15.962  < 2e-16 ***
oil.dl1        0.2512629  0.1505791   1.669  0.09791 .
oilprice.l1    0.0473114  0.0181143   2.612  0.01021 *
cost.l1       -0.2825076  0.6058422  -0.466  0.64188
rigs.l1       -0.0600216  0.0105010  -5.716 8.74e-08 ***
oil.l1         0.0332740  0.0359966   0.924  0.35723
trend.l1      -0.0008365  0.0011508  -0.727  0.46876
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.02335 on 115 degrees of freedom
Multiple R-squared:  0.9227,    Adjusted R-squared:  0.9019
F-statistic: 44.28 on 31 and 115 DF,  p-value: < 2.2e-16


Response oil.d :

Call:
lm(formula = oil.d ~ constant + D2011_02 + D2020_04 + D2020_05 +
    D2020_06 + D2020_07 + D2020_12 + D2021_02 + D2021_03 + D2021_04 +
    D2021_06 + sd1 + sd2 + sd3 + sd4 + sd5 + sd6 + sd7 + sd8 +
    sd9 + sd10 + sd11 + oilprice.dl1 + cost.dl1 + rigs.dl1 +
    oil.dl1 + oilprice.l1 + cost.l1 + rigs.l1 + oil.l1 + trend.l1 -
    1, data = data.mat)

Residuals:
      Min        1Q    Median        3Q       Max
-0.024421 -0.007929  0.000000  0.005418  0.042466

Coefficients:
               Estimate Std. Error t value Pr(>|t|)
constant      -0.5823479  1.8207628  -0.320 0.749672
D2011_02      -0.0521117  0.0138686  -3.758 0.000271 ***
D2020_04      -0.0673255  0.0170501  -3.949 0.000136 ***
D2020_05      -0.1907334  0.0202559  -9.416 5.94e-16 ***
D2020_06       0.0939519  0.0267749   3.509 0.000643 ***
D2020_07       0.0510826  0.0174197   2.932 0.004059 **
D2020_12      -0.0062816  0.0139882  -0.449 0.654233
D2021_02      -0.1633933  0.0140197 -11.655  < 2e-16 ***
D2021_03       0.1674250  0.0190207   8.802 1.59e-14 ***
D2021_04       0.0006998  0.0194185   0.036 0.971317
D2021_06       0.0012402  0.0141522   0.088 0.930319
sd1           -0.0090185  0.0054574  -1.653 0.101156
sd2            0.0092966  0.0060027   1.549 0.124194
sd3            0.0136444  0.0055995   2.437 0.016354 *
sd4            0.0012954  0.0058345   0.222 0.824688
sd5            0.0088022  0.0057238   1.538 0.126840
sd6            0.0030716  0.0059035   0.520 0.603851
sd7            0.0100902  0.0057541   1.754 0.082166 .
sd8            0.0140077  0.0056594   2.475 0.014777 *
sd9            0.0132478  0.0055112   2.404 0.017823 *
sd10           0.0153762  0.0054298   2.832 0.005466 **
sd11           0.0103427  0.0054156   1.910 0.058651 .
oilprice.dl1   0.0297231  0.0175733   1.691 0.093473 .
cost.dl1      -0.0319228  0.3728146  -0.086 0.931912
```

```
rigs.dl1       0.0336548  0.0269140   1.250 0.213670
oil.dl1        0.0499284  0.0826086   0.604 0.546770
oilprice.l1  -0.0030453  0.0099376  -0.306 0.759824
cost.l1        0.1125057  0.3323688   0.338 0.735606
rigs.l1        0.0292600  0.0057609   5.079 1.48e-06 ***
oil.l1        -0.0321923  0.0197479  -1.630 0.105804
trend.l1       0.0005529  0.0006313   0.876 0.382954
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.01281 on 115 degrees of freedom
Multiple R-squared:  0.9141,    Adjusted R-squared:  0.891
F-statistic:  39.5 on 31 and 115 DF,  p-value: < 2.2e-16



###########################
# BOOTSTRAPPED STANDARD ERRORS
###########################

# Comment: below 'ect1' and 'ect2' denotes the cointegrating relationships

OOTSTRAP OF LINEAR MODEL  (method = residuals)

Original Model Fit
------------------
Call:
lm(formula = oilprice.d ~ ect1 + ect2 + constant + oilprice.dl1 +
    cost.dl1 + rigs.dl1 + oil.dl1 - 1 + sd1 + sd2 + sd3 + sd4 +
    sd5 + sd6 + sd7 + sd8 + sd9 + sd10 + sd11 + D2011_02 + D2020_04 +
    D2020_05 + D2020_06 + D2020_07 + D2020_12 + D2021_02 + D2021_03 +
    D2021_04 + D2021_06, data = OLSVecmData)

Coefficients:
        ect1          ect2      constant  oilprice.dl1      cost.dl1
    0.053320      4.406041    -23.976944      0.274193     -0.911256
         sd2           sd3           sd4           sd5           sd6
    0.026669     -0.008709      0.015484     -0.013156     -0.029915
        sd10          sd11      D2011_02      D2020_04      D2020_05
    0.011662     -0.017042      0.002116     -0.359714      0.760768
    D2021_02      D2021_03      D2021_04      D2021_06
    0.059241      0.081952     -0.156689      0.140898

    rigs.dl1       oil.dl1           sd1
   -0.163043      0.670907      0.027934
         sd7           sd8           sd9
    0.008806     -0.039387      0.019328
    D2020_06      D2020_07      D2020_12
    0.185352     -0.153199      0.130501

Bootstrap SD's:
        ect1          ect2      constant  oilprice.dl1      cost.dl1
  0.02330706    1.85344055   10.08340417    0.09671740    2.28814979
         sd2           sd3           sd4           sd5           sd6
  0.03778301    0.03487192    0.03654656    0.03501348    0.03662265
        sd10          sd11      D2011_02      D2020_04      D2020_05
  0.03418084    0.03408228    0.08638777    0.10229120    0.11835289
    D2021_02      D2021_03      D2021_04      D2021_06
  0.08877278    0.12325111    0.12434183    0.08685555

    rigs.dl1       oil.dl1           sd1
  0.13828255    0.52724336    0.03391253
         sd7           sd8           sd9
  0.03578275    0.03524540    0.03479240
    D2020_06      D2020_07      D2020_12
  0.16410042    0.11076542    0.08647370

BOOTSTRAP OF LINEAR MODEL  (method = residuals)
```

```
Original Model Fit
------------------
Call:
lm(formula = cost.d ~ ect1 + ect2 + constant + oilprice.dl1 +
    cost.dl1 + rigs.dl1 + oil.dl1 - 1 + sd1 + sd2 + sd3 + sd4 +
    sd5 + sd6 + sd7 + sd8 + sd9 + sd10 + sd11 + D2011_02 + D2020_04 +
    D2020_05 + D2020_06 + D2020_07 + D2020_12 + D2021_02 + D2021_03 +
    D2021_04 + D2021_06, data = OLSVecmData)

Coefficients:
        ect1          ect2      constant  oilprice.dl1      cost.dl1
  -3.485e-03    -3.404e-01     1.850e+00    -1.301e-02     1.582e-01
         sd2           sd3           sd4           sd5           sd6
   1.572e-03     2.835e-04     2.294e-03     1.920e-03     2.446e-03
        sd10          sd11      D2011_02      D2020_04      D2020_05
  -1.378e-04    -3.790e-05    -5.973e-04    -1.371e-03    -5.230e-03
    D2021_02      D2021_03      D2021_04      D2021_06
  -1.073e-03     7.634e-03    -1.026e-02     3.046e-03

    rigs.dl1       oil.dl1           sd1
    4.437e-03     3.096e-02     3.014e-03
          sd7           sd8           sd9
    6.894e-04     4.067e-04    -1.117e-04
     D2020_06      D2020_07      D2020_12
    1.513e-02     1.656e-03     3.188e-06

Bootstrap SD's:
        ect1          ect2      constant  oilprice.dl1      cost.dl1
0.0007749857  0.0615213721  0.3346975065  0.0031864244  0.0754844429
         sd2           sd3           sd4           sd5           sd6
0.0012491280  0.0011784897  0.0012388887  0.0012130685  0.0012243585
        sd10          sd11      D2011_02      D2020_04      D2020_05
0.0011554824  0.0011356505  0.0028697292  0.0033510873  0.0038999831
    D2021_02      D2021_03      D2021_04      D2021_06
0.0030109808  0.0038928498  0.0040661882  0.0029582294

     rigs.dl1       oil.dl1           sd1
  0.0045811746  0.0170399314  0.0011391127
           sd7           sd8           sd9
  0.0011887416  0.0011842977  0.0011522069
      D2020_06      D2020_07      D2020_12
  0.0054773515  0.0036248978  0.0028990696

BOOTSTRAP OF LINEAR MODEL  (method = residuals)

Original Model Fit
------------------
Call:
lm(formula = rigs.d ~ ect1 + ect2 + constant + oilprice.dl1 +
    cost.dl1 + rigs.dl1 + oil.dl1 - 1 + sd1 + sd2 + sd3 + sd4 +
    sd5 + sd6 + sd7 + sd8 + sd9 + sd10 + sd11 + D2011_02 + D2020_04 +
    D2020_05 + D2020_06 + D2020_07 + D2020_12 + D2021_02 + D2021_03 +
    D2021_04 + D2021_06, data = OLSVecmData)

Coefficients:
        ect1          ect2      constant  oilprice.dl1      cost.dl1
   0.0067378    -0.8670470     4.7213072     0.1372046     1.0681550
         sd2           sd3           sd4           sd5           sd6
  -0.0227320    -0.0075718    -0.0025595    -0.0205342     0.0070061
        sd10          sd11      D2011_02      D2020_04      D2020_05
  -0.0081336    -0.0125133    -0.0195846    -0.2286275    -0.1128728
    D2021_02      D2021_03      D2021_04      D2021_06
  -0.0050052     0.0053978    -0.0636530    -0.0469090

     rigs.dl1       oil.dl1           sd1
    0.8613029     0.3309851    -0.0177226
          sd7           sd8           sd9
   -0.0007843    -0.0044582    -0.0195112
    D2020_06      D2020_07      D2020_12
```

```
      0.1435665     0.0425918    -0.0518423

Bootstrap SD's:
        ect1          ect2      constant  oilprice.dl1      cost.dl1
 0.006201117   0.497795129   2.708178287   0.025460138   0.604440964
         sd2           sd3           sd4           sd5           sd6
 0.010163415   0.009433324   0.009951486   0.009555364   0.009821618
        sd10          sd11      D2011_02      D2020_04      D2020_05
 0.009089393   0.009186831   0.022581038   0.027923055   0.030895788
    D2021_02      D2021_03      D2021_04      D2021_06
 0.023897138   0.031807934   0.033497375   0.023674525

      rigs.dl1       oil.dl1           sd1
   0.036102838   0.138643464   0.009331100
           sd7           sd8           sd9
   0.009619435   0.009497837   0.009297496
      D2020_06      D2020_07      D2020_12
   0.043239704   0.029188031   0.023451276

BOOTSTRAP OF LINEAR MODEL  (method = residuals)

Original Model Fit
------------------
Call:
lm(formula = oil.d ~ ect1 + ect2 + constant + oilprice.dl1 +
    cost.dl1 + rigs.dl1 + oil.dl1 - 1 + sd1 + sd2 + sd3 + sd4 +
    sd5 + sd6 + sd7 + sd8 + sd9 + sd10 + sd11 + D2011_02 + D2020_04 +
    D2020_05 + D2020_06 + D2020_07 + D2020_12 + D2021_02 + D2021_03 +
    D2021_04 + D2021_06, data = OLSVecmData)

Coefficients:
        ect1          ect2      constant  oilprice.dl1      cost.dl1
  -0.0160738    -0.0733277     0.4070573     0.0454821     0.0978957
         sd2           sd3           sd4           sd5           sd6
   0.0093180     0.0148446     0.0024768     0.0100303     0.0048972
        sd10          sd11      D2011_02      D2020_04      D2020_05
   0.0157590     0.0102593    -0.0486711    -0.0708750    -0.1874052
    D2021_02      D2021_03      D2021_04      D2021_06
  -0.1624927     0.1744309    -0.0017776     0.0008844

     rigs.dl1       oil.dl1           sd1
    0.0599975     0.0770421    -0.0093236
          sd7           sd8           sd9
    0.0117005     0.0152681     0.0140792
     D2020_06      D2020_07      D2020_12
    0.1019895     0.0534755    -0.0076051

Bootstrap SD's:
        ect1          ect2      constant  oilprice.dl1      cost.dl1
 0.003254010   0.264335503   1.438088125   0.013737407   0.326664301
         sd2           sd3           sd4           sd5           sd6
 0.005457923   0.005051122   0.005154447   0.005192784   0.005198355
        sd10          sd11      D2011_02      D2020_04      D2020_05
 0.004959553   0.004827921   0.012390825   0.014674603   0.017014460
    D2021_02      D2021_03      D2021_04      D2021_06
 0.012611701   0.016716820   0.017572841   0.012860749

      rigs.dl1       oil.dl1           sd1
   0.019457615   0.073212890   0.005008806
           sd7           sd8           sd9
   0.005196691   0.005042842   0.005012952
      D2020_06      D2020_07      D2020_12
   0.023280811   0.015409781   0.012579474


###########################
# AUTOCORRELATION, HETEROSCEDASTICITY AND NORMALITY
###########################
```

```
# Autocorrelation:

Portmanteau Test (asymptotic)
data:  Residuals of VAR object VAR
Chi-squared = 392.72, df = 356, p-value = 0.08749

Portmanteau Test (adjusted)
data:  Residuals of VAR object VAR
Chi-squared = 431.76, df = 356, p-value = 0.003636
(Comment: The low p-value is likely caused by the high number of dummy variables)

Breusch-Godfrey LM test
data:  Residuals of VAR object VAR
Chi-squared = 94.455, df = 80, p-value = 0.1287

Edgerton-Shukur F test
data:  Residuals of VAR object VAR
F statistic = 0.91043, df1 = 80, df2 = 365, p-value = 0.6893

# Heteroscedasticity:
ARCH (multivariate)
data:  Residuals of VAR object VAR
Chi-squared = 1230.1, df = 1200, p-value = 0.2667

# Normality

JB-Test (multivariate)
data:  Residuals of VAR object VAR
Chi-squared = 366.51, df = 8, p-value < 2.2e-16

Skewness only (multivariate)
data:  Residuals of VAR object VAR
Chi-squared = 59.177, df = 4, p-value = 4.32e-12

Kurtosis only (multivariate)
data:  Residuals of VAR object VAR
Chi-squared = 307.33, df = 4, p-value < 2.2e-16


###########################
# UNIT ROOT TESTS ON MODEL RESIDUALS
###########################

Phillips-Perron Unit Root Test
data:  res_oil
Dickey-Fuller Z(alpha) = -154.09, Truncation lag parameter = 4, p-value = 0.01
alternative hypothesis: stationary

Phillips-Perron Unit Root Test
data:  res_rigs
Dickey-Fuller Z(alpha) = -153.65, Truncation lag parameter = 4, p-value = 0.01
alternative hypothesis: stationary

Phillips-Perron Unit Root Test
data:  res_cost
Dickey-Fuller Z(alpha) = -144.95, Truncation lag parameter = 4, p-value = 0.01
alternative hypothesis: stationary

Phillips-Perron Unit Root Test
data:  res_oilprice
Dickey-Fuller Z(alpha) = -134.85, Truncation lag parameter = 4, p-value = 0.01
alternative hypothesis: stationary


###########################
# CORRELATION MATRIX FOR RESIDUALS FROM VECM
###########################

            oilprice.d      cost.d      rigs.d       oil.d
```

```
oilprice.d  1.00000000 -0.08727788  0.08959654  0.05712338
cost.d     -0.08727788  1.00000000 -0.09109386 -0.09569478
rigs.d      0.08959654 -0.09109386  1.00000000  0.09002453
oil.d       0.05712338 -0.09569478  0.09002453  1.00000000


###########################
# CHOLESKY DECOMPOSITION FOR ORTHOGONAL IMPULSE RESPONSES
###########################

              oilprice.d       cost.d      rigs.d      oil.d
oilprice.d  0.0906860367  0.000000000 0.000000000 0.00000000
cost.d     -0.0002623479  0.002994422 0.000000000 0.00000000
rigs.d      0.0021732060 -0.002027589 0.024072676 0.00000000
oil.d       0.0007385619 -0.001177293 0.001006951 0.01281483


###########################
# TESTS FOR WEAK EXOGENEITY
###########################

# TESTS FOR ZERO COEFFICIENT ON COEFFICIENT IN ALPHA MATRIX
# (see the Johansen procedure above for the alpha matrix)

# H0: OILPRICE IS WEAKLY EXOGENOUS

Estimation and testing under linear restrictions on alpha/beta

The VECM has been estimated subject to:
beta=H*phi and/or alpha=A*psi

     [,1] [,2] [,3]
[1,]    0    0    0
[2,]    1    0    0
[3,]    0    1    0
[4,]    0    0    1

Eigenvalues of restricted VAR (lambda):
[1] 0.4525 0.1863 0.1122 0.0000 0.0000

The value of the likelihood ratio test statistic:
4.31 distributed as chi square with 2 df.
The p-value of the test statistic is: 0.12

Eigenvectors, normalised to first column
of the restricted VAR:

                     [,1]     [,2]
RK.oilprice.l1   1.0000   1.0000
RK.cost.l1     -16.8492 113.3666
RK.rigs.l1      -2.8552   1.0951
RK.oil.l1        2.7754  -6.4979
RK.trend.l1     -0.0573   0.2164

Weights W of the restricted VAR:

        [,1]    [,2]
[1,]  0.0000  0.0000
[2,]  0.0003 -0.0024
[3,]  0.0138 -0.0093
[4,] -0.0128 -0.0037

# H0: COST IS WEAKLY EXOGENOUS

Estimation and testing under linear restrictions on alpha/beta

The VECM has been estimated subject to:
beta=H*phi and/or alpha=A*psi

```

```
     [,1] [,2] [,3]
[1,]    1    0    0
[2,]    0    0    0
[3,]    0    1    0
[4,]    0    0    1

Eigenvalues of restricted VAR (lambda):
[1] 0.4512 0.1365 0.0536 0.0000 0.0000

The value of the likelihood ratio test statistic:
13.32 distributed as chi square with 2 df.
The p-value of the test statistic is: 0

Eigenvectors, normalised to first column
of the restricted VAR:

                  [,1]     [,2]
RK.oilprice.l1  1.0000   1.0000
RK.cost.l1     -5.2874 -10.9624
RK.rigs.l1     -2.4136  -0.7721
RK.oil.l1       1.8253   0.3625
RK.trend.l1    -0.0314  -0.0192

Weights W of the restricted VAR:

        [,1]   [,2]
[1,]  0.0060 0.0128
[2,]  0.0000 0.0000
[3,]  0.0160 0.0287
[4,] -0.0148 0.0096

# H0: RIGS IS WEAKLY EXOGENOUS

Estimation and testing under linear restrictions on alpha/beta

The VECM has been estimated subject to:
beta=H*phi and/or alpha=A*psi

     [,1] [,2] [,3]
[1,]    1    0    0
[2,]    0    1    0
[3,]    0    0    0
[4,]    0    0    1

Eigenvalues of restricted VAR (lambda):
[1] 0.3725 0.2033 0.0311 0.0000 0.0000

The value of the likelihood ratio test statistic:
21.12 distributed as chi square with 2 df.
The p-value of the test statistic is: 0

Eigenvectors, normalised to first column
of the restricted VAR:

                   [,1]    [,2]
RK.oilprice.l1   1.0000  1.0000
RK.cost.l1     -14.0846 69.5831
RK.rigs.l1      -6.5881  0.4746
RK.oil.l1        6.4303 -4.3264
RK.trend.l1     -0.1013  0.1362

Weights W of the restricted VAR:

        [,1]    [,2]
[1,]  0.0001  0.0753
[2,]  0.0001 -0.0053
[3,]  0.0000  0.0000
[4,] -0.0046  0.0008
```

```
# H0: OIL IS WEAKLY EXOGENOUS

stimation and testing under linear restrictions on alpha/beta

The VECM has been estimated subject to:
beta=H*phi and/or alpha=A*psi

     [,1] [,2] [,3]
[1,]    1    0    0
[2,]    0    1    0
[3,]    0    0    1
[4,]    0    0    0

Eigenvalues of restricted VAR (lambda):
[1] 0.2371 0.1904 0.0278 0.0000 0.0000

The value of the likelihood ratio test statistic:
52 distributed as chi square with 2 df.
The p-value of the test statistic is: 0

Eigenvectors, normalised to first column
of the restricted VAR:

                      [,1]    [,2]
RK.oilprice.l1    1.0000  1.0000
RK.cost.l1     -125.7488 39.2094
RK.rigs.l1       -4.1415 -0.2760
RK.oil.l1         8.6363 -2.2847
RK.trend.l1      -0.2614  0.0740

Weights W of the restricted VAR:

        [,1]    [,2]
[1,] -0.0056  0.1137
[2,]  0.0010 -0.0062
[3,]  0.0127  0.0320
[4,]  0.0000  0.0000


###########################
# THE UNRESTRICTED VECM
###########################

# Comment: The term 'X.dl1' refers to the first-difference of the lag of X.
# The variables starting with D are dummies, variables beginning with
# 'sd' are seasonal dummies


> summary(VECM_unrestr)
Response oilprice.d :

Call:
lm(formula = oilprice.d ~ constant + D2011_02 + D2020_04 + D2020_05 +
    D2020_06 + D2020_07 + D2020_12 + D2021_02 + D2021_03 + D2021_04 +
    D2021_06 + sd1 + sd2 + sd3 + sd4 + sd5 + sd6 + sd7 + sd8 +
    sd9 + sd10 + sd11 + oilprice.dl1 + cost.dl1 + rigs.dl1 +
    oil.dl1 + oilprice.l1 + cost.l1 + rigs.l1 + oil.l1 + trend.l1 -
    1, data = data.mat)

Residuals:
     Min       1Q   Median       3Q      Max
-0.43566 -0.03739  0.00080  0.05033  0.18226

Coefficients:
               Estimate Std. Error t value Pr(>|t|)
constant     -20.878269  12.790049  -1.632  0.10533
D2011_02      -0.028858   0.097421  -0.296  0.76760
D2020_04      -0.384450   0.119769  -3.210  0.00172 **
D2020_05       0.680450   0.142289   4.782 5.18e-06 ***
```

```
D2020_06       0.097283   0.188082   0.517  0.60598
D2020_07      -0.179140   0.122366  -1.464  0.14593
D2020_12       0.127726   0.098261   1.300  0.19625
D2021_02       0.057001   0.098482   0.579  0.56386
D2021_03       0.059273   0.133612   0.444  0.65815
D2021_04      -0.138245   0.136406  -1.013  0.31296
D2021_06       0.131467   0.099413   1.322  0.18865
sd1            0.027117   0.038336   0.707  0.48078
sd2            0.027522   0.042166   0.653  0.51525
sd3           -0.012859   0.039334  -0.327  0.74432
sd4            0.014538   0.040985   0.355  0.72344
sd5           -0.014098   0.040207  -0.351  0.72650
sd6           -0.030698   0.041469  -0.740  0.46065
sd7            0.008023   0.040420   0.198  0.84301
sd8           -0.039262   0.039755  -0.988  0.32542
sd9            0.018958   0.038714   0.490  0.62529
sd10           0.011155   0.038142   0.292  0.77046
sd11          -0.015956   0.038042  -0.419  0.67569
oilprice.dl1   0.245702   0.123444   1.990  0.04892 *
cost.dl1      -0.789770   2.618857  -0.302  0.76352
rigs.dl1      -0.272236   0.189059  -1.440  0.15260
oil.dl1        0.529419   0.580289   0.912  0.36350
oilprice.l1    0.035708   0.069807   0.512  0.60996
cost.l1        3.935281   2.334743   1.686  0.09460 .
rigs.l1        0.012818   0.040468   0.317  0.75201
oil.l1        -0.288047   0.138721  -2.076  0.04008 *
trend.l1       0.008143   0.004435   1.836  0.06893 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.08999 on 115 degrees of freedom
Multiple R-squared:  0.5059,    Adjusted R-squared:  0.3727
F-statistic: 3.798 on 31 and 115 DF,  p-value: 9.853e-08


Response cost.d :

Call:
lm(formula = cost.d ~ constant + D2011_02 + D2020_04 + D2020_05 +
    D2020_06 + D2020_07 + D2020_12 + D2021_02 + D2021_03 + D2021_04 +
    D2021_06 + sd1 + sd2 + sd3 + sd4 + sd5 + sd6 + sd7 + sd8 +
    sd9 + sd10 + sd11 + oilprice.dl1 + cost.dl1 + rigs.dl1 +
    oil.dl1 + oilprice.l1 + cost.l1 + rigs.l1 + oil.l1 + trend.l1 -
    1, data = data.mat)

Residuals:
      Min        1Q    Median        3Q       Max
-0.008106 -0.001563  0.000000  0.001150  0.012663

Coefficients:
               Estimate Std. Error t value Pr(>|t|)
constant      2.084e+00  4.309e-01   4.835 4.16e-06 ***
D2011_02     -3.654e-04  3.282e-03  -0.111  0.91156
D2020_04     -2.399e-03  4.035e-03  -0.595  0.55334
D2020_05     -5.706e-03  4.794e-03  -1.190  0.23638
D2020_06      1.545e-02  6.337e-03   2.438  0.01628 *
D2020_07      1.756e-03  4.123e-03   0.426  0.67101
D2020_12     -2.903e-04  3.311e-03  -0.088  0.93027
D2021_02     -9.304e-04  3.318e-03  -0.280  0.77968
D2021_03      8.668e-03  4.502e-03   1.925  0.05665 .
D2021_04     -1.049e-02  4.596e-03  -2.281  0.02436 *
D2021_06      2.846e-03  3.349e-03   0.850  0.39724
sd1           2.944e-03  1.292e-03   2.280  0.02448 *
sd2           1.588e-03  1.421e-03   1.118  0.26606
sd3           4.569e-04  1.325e-03   0.345  0.73090
sd4           2.508e-03  1.381e-03   1.816  0.07192 .
sd5           2.144e-03  1.355e-03   1.582  0.11632
sd6           2.786e-03  1.397e-03   1.994  0.04848 *
sd7           9.886e-04  1.362e-03   0.726  0.46934
```

```
sd8             6.512e-04  1.339e-03   0.486  0.62776
sd9             4.324e-05  1.304e-03   0.033  0.97361
sd10           -7.115e-05  1.285e-03  -0.055  0.95594
sd11           -3.885e-05  1.282e-03  -0.030  0.97587
oilprice.dl1 -1.037e-02  4.159e-03  -2.494  0.01404 *
cost.dl1        1.849e-01  8.823e-02   2.096  0.03830 *
rigs.dl1        7.992e-03  6.370e-03   1.255  0.21216
oil.dl1         3.421e-02  1.955e-02   1.750  0.08279 .
oilprice.l1   -6.240e-03  2.352e-03  -2.653  0.00911 **
cost.l1        -3.827e-01  7.866e-02  -4.865 3.66e-06 ***
rigs.l1        -2.595e-03  1.363e-03  -1.903  0.05955 .
oil.l1          2.337e-02  4.674e-03   5.000 2.07e-06 ***
trend.l1       -7.445e-04  1.494e-04  -4.982 2.24e-06 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.003032 on 115 degrees of freedom
Multiple R-squared:  0.5228,    Adjusted R-squared:  0.3942
F-statistic: 4.065 on 31 and 115 DF,  p-value: 2.097e-08


Response rigs.d :

Call:
lm(formula = rigs.d ~ constant + D2011_02 + D2020_04 + D2020_05 +
    D2020_06 + D2020_07 + D2020_12 + D2021_02 + D2021_03 + D2021_04 +
    D2021_06 + sd1 + sd2 + sd3 + sd4 + sd5 + sd6 + sd7 + sd8 +
    sd9 + sd10 + sd11 + oilprice.dl1 + cost.dl1 + rigs.dl1 +
    oil.dl1 + oilprice.l1 + cost.l1 + rigs.l1 + oil.l1 + trend.l1 -
    1, data = data.mat)

Residuals:
     Min       1Q   Median       3Q      Max
-0.08898 -0.01203  0.00000  0.01044  0.04983

Coefficients:
               Estimate Std. Error t value Pr(>|t|)
constant      1.5934522  3.3188881   0.480  0.63206
D2011_02     -0.0293562  0.0252797  -1.161  0.24794
D2020_04     -0.2170456  0.0310789  -6.984 1.98e-10 ***
D2020_05     -0.1209920  0.0369225  -3.277  0.00139 **
D2020_06      0.1211548  0.0488053   2.482  0.01449 *
D2020_07      0.0359121  0.0317527   1.131  0.26041
D2020_12     -0.0476943  0.0254977  -1.871  0.06395 .
D2021_02     -0.0077181  0.0255552  -0.302  0.76319
D2021_03     -0.0155689  0.0346709  -0.449  0.65424
D2021_04     -0.0565148  0.0353960  -1.597  0.11309
D2021_06     -0.0455640  0.0257967  -1.766  0.08000 .
sd1          -0.0167619  0.0099478  -1.685  0.09470 .
sd2          -0.0228205  0.0109418  -2.086  0.03922 *
sd3          -0.0111567  0.0102067  -1.093  0.27665
sd4          -0.0061713  0.0106351  -0.580  0.56286
sd5          -0.0242899  0.0104334  -2.328  0.02165 *
sd6           0.0014068  0.0107609   0.131  0.89621
sd7          -0.0057208  0.0104885  -0.545  0.58651
sd8          -0.0083416  0.0103160  -0.809  0.42041
sd9          -0.0220609  0.0100458  -2.196  0.03010 *
sd10         -0.0092986  0.0098975  -0.939  0.34945
sd11         -0.0122855  0.0098716  -1.245  0.21583
oilprice.dl1  0.0894456  0.0320326   2.792  0.00613 **
cost.dl1      0.6653005  0.6795669   0.979  0.32963
rigs.dl1      0.7830993  0.0490588  15.962  < 2e-16 ***
oil.dl1       0.2512629  0.1505791   1.669  0.09791 .
oilprice.l1   0.0473114  0.0181143   2.612  0.01021 *
cost.l1      -0.2825076  0.6058422  -0.466  0.64188
rigs.l1      -0.0600216  0.0105010  -5.716 8.74e-08 ***
oil.l1        0.0332740  0.0359966   0.924  0.35723
trend.l1     -0.0008365  0.0011508  -0.727  0.46876
---
```

```
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.02335 on 115 degrees of freedom
Multiple R-squared:  0.9227,    Adjusted R-squared:  0.9019
F-statistic: 44.28 on 31 and 115 DF,  p-value: < 2.2e-16


Response oil.d :

Call:
lm(formula = oil.d ~ constant + D2011_02 + D2020_04 + D2020_05 +
    D2020_06 + D2020_07 + D2020_12 + D2021_02 + D2021_03 + D2021_04 +
    D2021_06 + sd1 + sd2 + sd3 + sd4 + sd5 + sd6 + sd7 + sd8 +
    sd9 + sd10 + sd11 + oilprice.dl1 + cost.dl1 + rigs.dl1 +
    oil.dl1 + oilprice.l1 + cost.l1 + rigs.l1 + oil.l1 + trend.l1 -
    1, data = data.mat)

Residuals:
      Min        1Q    Median        3Q       Max
-0.024421 -0.007929  0.000000  0.005418  0.042466

Coefficients:
               Estimate Std. Error t value Pr(>|t|)
constant     -0.5823479  1.8207628  -0.320 0.749672
D2011_02     -0.0521117  0.0138686  -3.758 0.000271 ***
D2020_04     -0.0673255  0.0170501  -3.949 0.000136 ***
D2020_05     -0.1907334  0.0202559  -9.416 5.94e-16 ***
D2020_06      0.0939519  0.0267749   3.509 0.000643 ***
D2020_07      0.0510826  0.0174197   2.932 0.004059 **
D2020_12     -0.0062816  0.0139882  -0.449 0.654233
D2021_02     -0.1633933  0.0140197 -11.655  < 2e-16 ***
D2021_03      0.1674250  0.0190207   8.802 1.59e-14 ***
D2021_04      0.0006998  0.0194185   0.036 0.971317
D2021_06      0.0012402  0.0141522   0.088 0.930319
sd1          -0.0090185  0.0054574  -1.653 0.101156
sd2           0.0092966  0.0060027   1.549 0.124194
sd3           0.0136444  0.0055995   2.437 0.016354 *
sd4           0.0012954  0.0058345   0.222 0.824688
sd5           0.0088022  0.0057238   1.538 0.126840
sd6           0.0030716  0.0059035   0.520 0.603851
sd7           0.0100902  0.0057541   1.754 0.082166 .
sd8           0.0140077  0.0056594   2.475 0.014777 *
sd9           0.0132478  0.0055112   2.404 0.017823 *
sd10          0.0153762  0.0054298   2.832 0.005466 **
sd11          0.0103427  0.0054156   1.910 0.058651 .
oilprice.dl1  0.0297231  0.0175733   1.691 0.093473 .
cost.dl1     -0.0319228  0.3728146  -0.086 0.931912
rigs.dl1      0.0336548  0.0269140   1.250 0.213670
oil.dl1       0.0499284  0.0826086   0.604 0.546770
oilprice.l1  -0.0030453  0.0099376  -0.306 0.759824
cost.l1       0.1125057  0.3323688   0.338 0.735606
rigs.l1       0.0292600  0.0057609   5.079 1.48e-06 ***
oil.l1       -0.0321923  0.0197479  -1.630 0.105804
trend.l1      0.0005529  0.0006313   0.876 0.382954
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.01281 on 115 degrees of freedom
Multiple R-squared:  0.9141,    Adjusted R-squared:  0.891
F-statistic:  39.5 on 31 and 115 DF,  p-value: < 2.2e-16
```