Speculation in the Oil Market

Luciana Juvenal (Federal Reserve Bank of St. Louis and International Monetary Fund)

Ivan Petrella (Birkbeck College, University of London)

Disclaimer

The views expressed are those of the individual authors and do not necessarily reflect official positions of the Federal Reserve Bank of St. Louis, the Federal Reserve System, the Board of Governors, or the International Monetary Fund.

Big picture

⟶ Increase in oil prices since 2004 coincided with:

- Large flow of investment into commodity markets
  - Financial institutions invested billions of dollars in futures markets
  - Assets allocated to commodity index trading increased from 13$ billion to 260$ billion between 2004 and 2007 (Masters, 2008)
- Increased price comovement between different commodities

Big picture

⟶ Increase in oil prices since 2004 coincided with:

- Large flow of investment into commodity markets
  - Financial institutions invested billions of dollars in futures markets
  - Assets allocated to commodity index trading increased from 13$ billion to 260$ billion between 2004 and 2007 (Masters, 2008)
- Increased price comovement between different commodities

⟶ What are the drivers of oil price fluctuations?

- Supply and demand factors
- Role of speculation

In business....

"The increase in (oil) prices has not been driven by supply and demand."

Lord Browne, Group Chief Executive of British Petroleum (2006)

In the news....

The New York Times Reprints

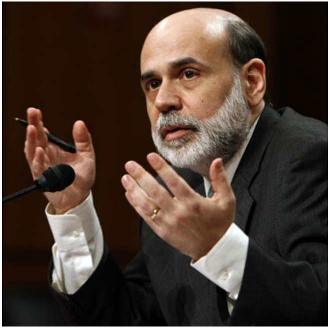
May 24, 2011

U.S. Suit Sees Manipulation of Oil Trades

By GRAHAM BOWLEY

After oil prices surged past $100 a barrel in 2008, suspicions that traders had manipulated the market led to Congressional hearings and regulatory investigations. But they produced no solid cases in the record run-up in gasoline prices.

In policy....



"[...] the sharp increases and extreme volatility of oil prices have led observers to suggest that some part of the rise in prices reflects a speculative component arising from the activities of traders in the oil markets."

Ben Bernanke, Federal Reserve Board (2004)

Our paper

- We examine the effects of oil shocks in a FAVAR with sign restrictions
  - Shocks: supply, global demand, oil inventory demand, speculation
  - FAVAR: small-scale VAR is not informationally sufficient to identify the shocks.

Our paper

- We examine the effects of oil shocks in a FAVAR with sign restrictions
  - Shocks: supply, global demand, oil inventory demand, speculation
  - FAVAR: small-scale VAR is not informationally sufficient to identify the shocks.
- We analyze the transmission of oil shocks to a large number of variables
  - Conditional correlations between oil prices and the price of other commodities.

What we find...

- Global demand shocks account for the largest share of oil price fluctuations
- Speculative shocks also played a role in the oil price increase between 2004-2008 and its subsequent collapse
- The comovement between oil prices and the prices of other commodities are mainly explained by global demand followed by speculative shocks

Speculation

- *Speculative demand shock* in Kilian and Murphy (2011) is a shock to the demand of inventories arising from:
  - Expected shortfalls of future oil supply relative to future oil demand
  - Increased uncertainty in the oil market
  - Speculative activity
- Producers (partially) accommodate the higher demand

$\longrightarrow$ We name this shock *oil inventory demand shock*

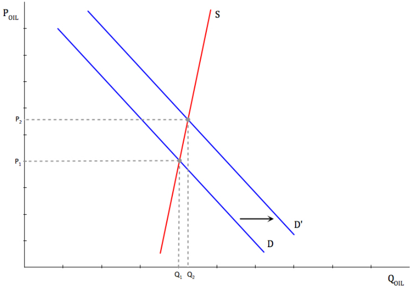
Speculation

- *Speculative demand shock* in Kilian and Murphy (2011) is a shock to the demand of inventories arising from:
  - Expected shortfalls of future oil supply relative to future oil demand
  - Increased uncertainty in the oil market
  - Speculative activity
- Producers (partially) accommodate the higher demand

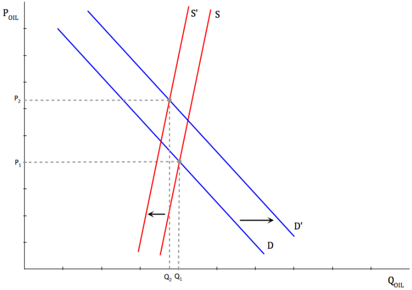
→ We name this shock *oil inventory demand shock*

- We disentangle the *speculative shock* following Hamilton (2009a, 2009b)
  - Speculative activity affects the incentive faced by producers
  - As producers/retailers expect higher prices, they hold back oil from the market (Hotelling's principle) and accumulate inventories

Oil Inventory Demand Shock



Speculative Shock



Literature

- Oil prices have historically been driven by demand factors
  - Kilian (2009)
  - Hicks and Kilian (2012), Kilian and Murphy (2011a, 2011b), Kilian and Park (2009), Baumeister and Peersman(2011, 2012), Peersman and Van Robays (2009, 2011)
- Role of speculation
  - Kilian and Murphy (2011), Lombardi and Van Robays (2011), Hamilton (2009a, 2009b)
- Finance literature
  - Tang and Xiong (2011), Singleton (2011), Hamilton and Wu (2011)
  - Haigh et al. (2007), Büyüksahin et al. (2008), Büyüksahin and Harris (2011)

Outline

1. Econometric Method: FAVAR
2. Data and Specification
3. Comparison VAR and FAVAR
4. Speculation
5. Extended Model Results
6. Speculation and Oil Prices in the Past Decade

FAVAR (Bernanke et al., 2005)

- Low-dimension VAR: All information to identify shocks is contained in the dataset
- Issue: Additional information of other economic series excluded from the VAR may be relevant to the dynamic relations implied in the VAR
- FAVAR allows us to identify oil shocks from a large dataset
- By-product of the FAVAR is to obtain IRFs for any variable included in the dataset

FAVAR

Each variable in our dataset, x_{it}, is expressed as the sum of a common component and an idiosyncratic component that are mutually orthogonal and unobservable:

$$x_{it} = \boldsymbol{\lambda}_i \mathbf{f}_t + \xi_{it},$$

where $\mathbf{f}_t$ represents r unobserved factors $(N \gg r)$, $\boldsymbol{\lambda}_i$ is the r-dimensional vector of factor loadings, and ξ_{it} are idiosyncratic components of x_{it} uncorrelated with $\mathbf{f}_t$.

The idiosyncratic components are weakly correlated across the cross-sectional dimension. We can consider them as shocks that affect a single variable or a small group of variables.

The common component is responsible for the main bulk of the co-movements between macroeconomic variables.

FAVAR

Let $\mathbf{y}_t$ denote the M-dimensional vector of variables describing the dynamics of the oil market: growth rate of oil production, oil inventories and real oil prices

The dynamics in the oil market can be well represented by the following FAVAR:

$$\begin{bmatrix} \mathbf{y}_t \\ \mathbf{f}_t \end{bmatrix} = \Phi(L) \begin{bmatrix} \mathbf{y}_{t-1} \\ \mathbf{f}_{t-1} \end{bmatrix} + \mathbf{u}_t,$$

where $\Phi(L)$ is the lag polynomial in the lag operator L, and $\mathbf{u}_t$ is the error term with mean zero and variance-covariance matrix Σ.

Estimation: Two-step procedure

1. The unobserved factors and loadings are estimated using the principal components method described by Stock and Watson (2002)
2. We use the estimated factors along with the oil variables to estimate our VAR model.

- The unobserved factors are estimated and then included as regressors in the FAVAR model $\Longrightarrow$ the two-step approach might suffer from the "generated regressor" problem.
- We adopt a non-overlapping block bootstrap technique.

Estimation: Non-overlapping block bootstrap

- We partition the $T \times N$ matrix of data $\mathbf{X} = [x_{it}]$ into S sub-matrices $\mathbf{X}_s$ (blocks), $s = 1, ..., S$, of dimension $\tau \times N$, where τ is an integer part of T/S.
- An integer h_s between 1 and S is drawn randomly with reintroduction S times to obtain the sequence $h_1, ..., h_s$.
- We then generate an artificial sample $\mathbf{X}^* = \left[\mathbf{X}'_{h_1}, ..., \mathbf{X}'_{h_s}\right]'$ of dimension $\tau S \times N$ and the corresponding impulse responses are estimated.

Identification of the shocks

- To give a structural interpretation to the shocks we follow the approach based on sign restrictions proposed by Uhlig (2005).
- We identify the shocks imposing economically meaningful sign restrictions on the impulse responses of a subset of variables
- Let $\mathbf{Q}$ denote an orthonormal matrix such that $\mathbf{Q}'\mathbf{Q} = \mathbf{I}$. The structural shocks can be recovered as $\boldsymbol{\eta}_t = \mathbf{Q}\mathbf{u}_t$.
- The orthonormal matrices $\mathbf{Q}$ are found from the eigenvalue decomposition of a random $(q \times q)$ matrix drawn from a normal distribution with unitary variance (see Rubio-Ramirez et al., 2010).

Estimation: Impulse responses

- The corresponding structural impulse response function for the oil variables can be recovered as

$$\mathbf{y}_t = [\mathbf{I}_3; \mathbf{0}_{3\times r}] \left[\mathbf{I}_{3+r} - \widehat{\Phi}(L)L\right]^{-1} \mathbf{Q}' \boldsymbol{\eta}_t,$$

where the moving average representation of the *i*-th variable in the dataset can be written as

$$x_{it} = [\mathbf{0}_{1\times 3}; \widehat{\boldsymbol{\lambda}}_i] \left[\mathbf{I}_{3+r} - \widehat{\Phi}(L)L\right]^{-1} \mathbf{Q}' \boldsymbol{\eta}_t.$$

Data

- Quarterly data 1972-2009
- 151 series
  - Macroeconomic and financial variables of the G7 countries: output, prices, interest rates, labor market indicators, trade, stock market indices, exchange rates (IFS, OECD)
  - Oil market variables: prices, production (DOE), and inventories (EIA)
  - Commodity prices (IFS)
  - Aggregate economic activity: aggregate industrial production (IFS) and Kilian (2009) measure of real global economic activity
- All data are transformed to reach stationarity

Test of sufficient information

- Does the large dataset contain valuable information wrt the small-scale VAR to identify the shocks?
- Test for information sufficiency using Forni and Gambetti (2011)
- Set maximum number of factors r=6 and compute the 6 principal components
- Test whether the principal components Granger-cause the variables of the VAR
- If the null of no Granger-causality is not rejected for any of the successive combinations of principal components, the variables of the VAR are informationally sufficient

Test of sufficient information

4 variable-VAR with aggregate industrial production

	VAR	VAR+1F	VAR+2F	VAR+3F	VAR+4F
1F	0.0233	—	—	—	—
2F	0.0100	0.1133	—	—	—
3F	0.0033	0.2333	0.4433	—	—
4F	0.0067	0.0167	0.0233	0.0000	—
5F	0.0200	0.0200	0.0033	0.0000	0.1433
6F	0.0300	0.0167	0.0000	0.0000	0.1200

Bootstrapped *p*-values of the Granger causality test for the VAR and VAR augmented with factors. Based on the Gelper and Croux (2007) multivariate extension of the out-of-sample Granger-causality test.

Test of sufficient information

4 variable-VAR with Kilian measure of real global economic activity

	VAR	VAR+1F	VAR+2F	VAR+3F	VAR+4F
1F	0.0400	—	—	—	—
2F	0.0033	0.0133	—	—	—
3F	0.0033	0.0033	0.0733	—	—
4F	0.0000	0.0000	0.0000	0.0067	—
5F	0.0000	0.0000	0.0200	0.0133	0.1333
6F	0.0033	0.0300	0.1600	0.0100	0.3500

Bootstrapped *p*-values of the Granger causality test for the VAR and VAR augmented with factors. Based on the Gelper and Croux (2007) multivariate extension of the out-of-sample Granger-causality test.

Evaluating latent and observed factors

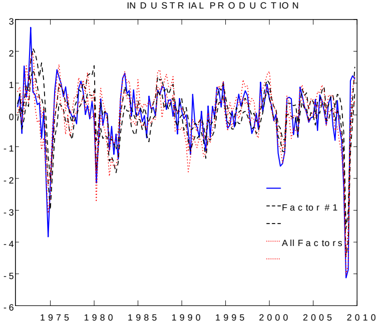
- Do observable economic variables span the same information as the unobserved factors?
- Based on Bai and Ng (2006) test for this hypothesis: No

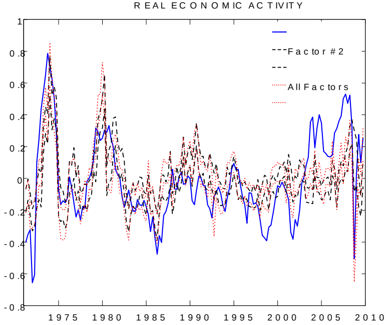
Go to Test

Go to IRF

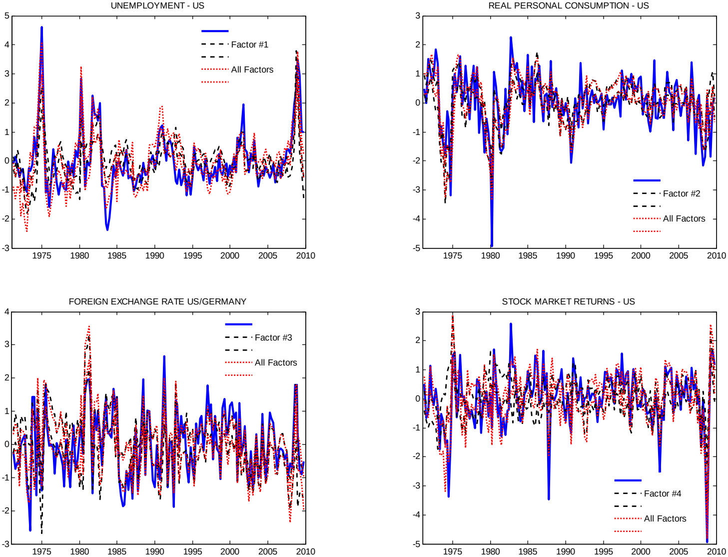
What are the factors?

Factors and proxies of aggregate economic activity





What are the factors?



Shocks

- Oil supply shock: Shock to the flow of crude oil production
- Global demand shock: Shock to the demand of oil driven by the global business cycle
- Oil inventory demand shock: Shock to the demand of oil inventories arising from expected shortage (or uncertainty of future supply)

- Speculative Shocks: *Shock driven by higher expected future prices (where these might be influenced by traders activity in the oil futures market).*

Identification of the shocks

Shock	Oil prod.	Oil inv.	Real oil P	Real activity
Oil supply	−	−	+	−
Oil inventory demand	+	+	+	−
Global Demand	+		+	+

- Inspired by Kilian and Murphy (2011)
- Additional restriction: upper bound for the response of impact elasticity of oil production wrt real price of oil is 0.0257 (Kilian and Murphy, 2011)
- We estimate a VAR and FAVAR

Generalizing Kilian and Murphy (2011)

- The first factor is a proxy for global economic activity
- VAR à la Kilian and Murphy (2011) can be considered as a special case of our model
  - VAR augmented with one factor

Results

Variance Decomposition of the Real Oil Price

Horizon		Oil supply	Oil inventory demand	Global demand
1	VAR (KM)	0.0865	0.2850	0.5415
	VAR (AIP)	0.1171	0.2937	0.4758
	FAVAR	0.0609	0.1254	0.3769
2	VAR (KM)	0.0732	0.1997	0.6259
	VAR (AIP)	0.1162	0.2379	0.5212
	FAVAR	0.0443	0.0712	0.4242
3	VAR (KM)	0.0351	0.1623	0.6920
	VAR (AIP)	0.0784	0.2528	0.5439
	FAVAR	0.0297	0.0469	0.4461
4	VAR (KM)	0.0280	0.1361	0.7128
	VAR (AIP)	0.0655	0.2805	0.5327
	FAVAR	0.0272	0.0384	0.4449
8	VAR (KM)	0.0306	0.0687	0.7766
	VAR (AIP)	0.0868	0.1846	0.5993
	FAVAR	0.0573	0.0467	0.3834
12	VAR (KM)	0.0307	0.0837	0.7613
	VAR (AIP)	0.0879	0.2019	0.5814
	FAVAR	0.0951	0.0696	0.3372

Speculation

- We identify a speculative shock using sign restrictions inspired by Hamilton (2009a, 2009b)

Speculation

- We identify a speculative shock using sign restrictions inspired by Hamilton (2009a, 2009b)
- Speculators can affect the incentives faced by producers
- As producers/retailers expect higher prices, they hold back oil from the market and accumulate inventories
- Hotelling (1931) principle: it would pay oil producers/retailers to forego current production in order to sell the oil at higher future price

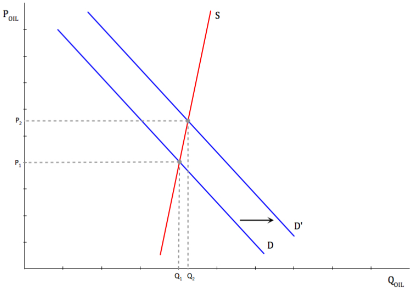
Speculation

Sign Restrictions

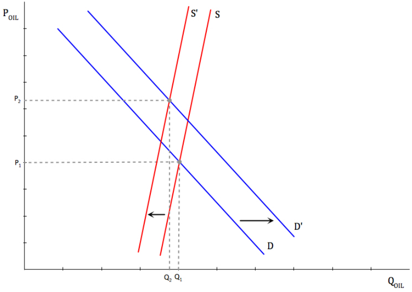
Shock	Oil prod.	Oil inv.	Real oil P	Real activity
Oil supply	−	−	+	−
Oil Inventory demand	+	+	+	−
Global Demand	+		+	+
Speculation	−	+	+	

Go to Model

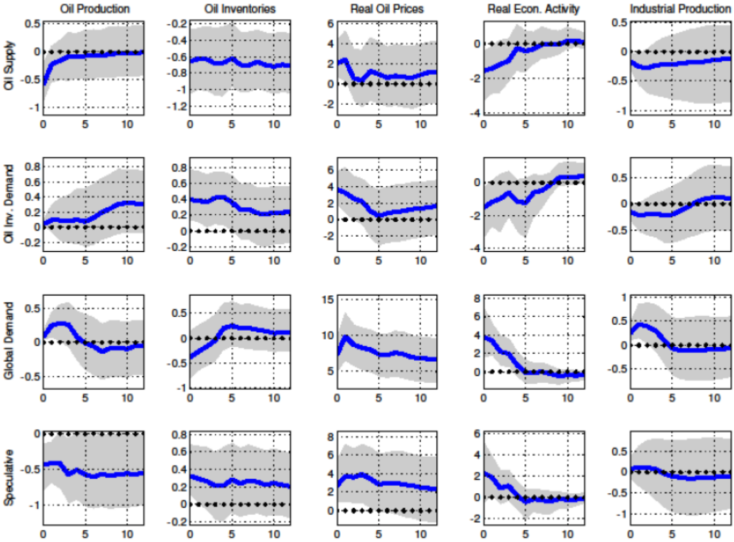
Oil Inventory Demand Shock



Speculative Shock



IRF FAVAR 4 Shocks



Variance Decomposition: Real oil price

Horizon	Oil supply	Oil inventory demand	Global demand	Speculative
1	0.0638	0.1315	0.3924	0.0900
2	0.0459	0.0742	0.4378	0.0984
3	0.0289	0.0475	0.4596	0.1095
4	0.0253	0.0388	0.4555	0.1269
8	0.0484	0.0464	0.4078	0.1043
12	0.0842	0.0677	0.3595	0.0924

Comovement of commodity prices

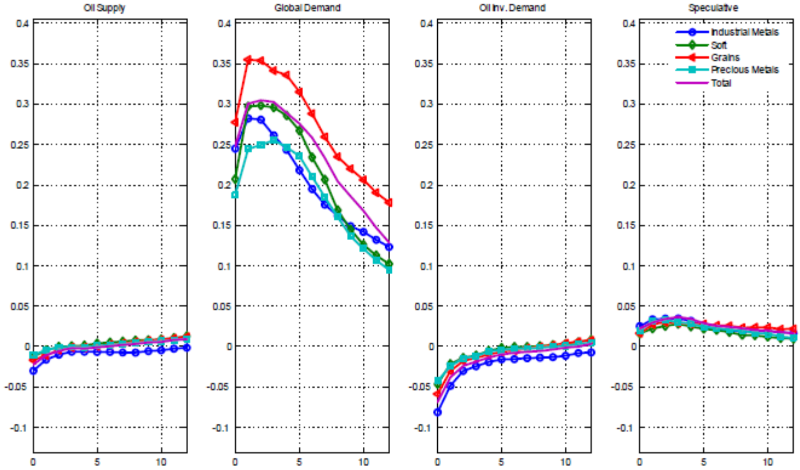
Following Den Haan and Sterk (2011), the correlation (COR) between the Kth-period-ahead forecast error of two variables, v_t and z_t, is

$$COR(v_t, z_t; K, s) = \frac{\sum_{k=1}^{K} v_k^{imp,s} z_k^{imp,s}}{SD(v_t; K) SD(z_t; K)}.$$

where $v_k^{imp,s}$ and $z_k^{imp,s}$ are the Kth-period responses of v and z to a 1-standard deviation innovation of the sth structural shock, and SD denotes the total standard deviation of the Kth-period-ahead forecast error given by

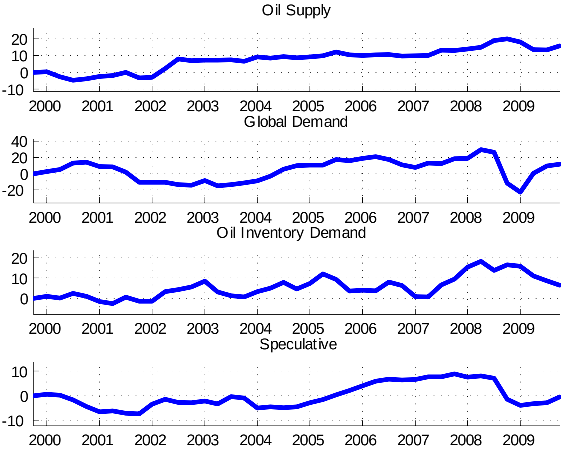
$$SD(b_t; K) = \left[\sum_{k=1}^{K} COV\left(b_t, b_t; K, s\right) \right]^{1/2} \quad \text{for } b_t = v_t, z_t,$$

Conditional correlations with oil prices



Historical Decomposition: Real oil price

2000 - 2010



Go to Subsample

What explains the 2003-08 oil price shock?

- No evidence that supply shocks were behind the oil price increase
- Strong evidence that a booming world economy was the cause of the price increase
- Evidence that speculation also played a role

Conclusion

- Small-scale VAR of the oil market is not informationally sufficient $\Longrightarrow$ FAVAR
- We analyze the role of speculation compared with supply and demand factors as drivers of oil prices
- We find that global demand shocks are the main drivers of oil price fluctuations
- Speculative shocks are the second most important driver
- Positive comovement between oil prices and the price of other commodities
  - Driven mainly by global demand shocks
  - Speculative shocks also played a role

Evaluating latent and observed factors

Bai and Ng (2006)

Before proceeding to describe our identification method it is interesting to consider to what extent some observable economic variables span the same information of the unobserved factors. Bai and Ng (2006) propose a test of this hypothesis based on the t-statistic

$$\tau_t(j) = \frac{\widehat{x}_{jt} - x_{jt}}{\sqrt{\widehat{var}\left(\widehat{x}_{jt}\right)}}$$

where $\widehat{x}_{jt}(= \widehat{\delta}_j' \widehat{\mathbf{f}}_t)$ is the least square projection of the variable x_{jt} on the estimated latent factors and the associate variance is constructed as detailed in Bai and Ng (2006).

Can any variable be considered as an observable factor? NO!

	$A(j)$	$M(j)$	$NS\ (j)$	$R^2\ (j)$
Oil production	0.793	38.776	6.112	0.140
Real oil prices	0.767	25.572	2.081	0.324
Oil inventories	0.916	83.424	28.093	0.034
Aggregate industrial production	0.567	9.495	0.289	0.775
Kilian measure of economic activity	0.709	15.752	1.101	0.475

Go back

Orthogonality test

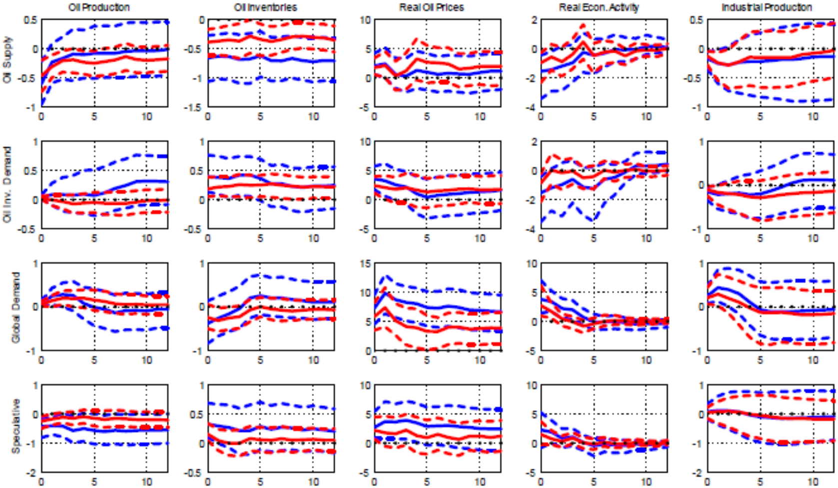
- If the structural shocks of the VAR are orthogonal to any available information at time t, some shocks could still be identified in the VAR.
- Otherwise, the identified shocks cannot be considered structural (Forni and Gambetti, 2011)
- We test wether each possible set of shocks is Granger-caused by the lagged factors.

Orthogonality test

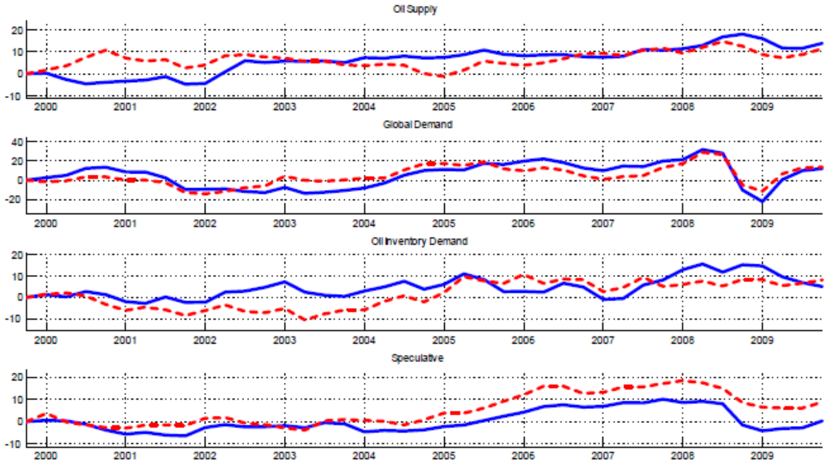
# of factors	Oil supply	Oil inventory demand	Global demand
1	0.0000	0.0180	0.0020
2	0.3470	0.1710	0.4440
3	0.3590	0.3870	0.2240
4	0.0010	0.8600	0.5860

Notes: Size of the rejection set (at the 10% level) of the F-test of orthogonality for each of the shocks identified from the VAR with sign restrictions.

Subsample analysis

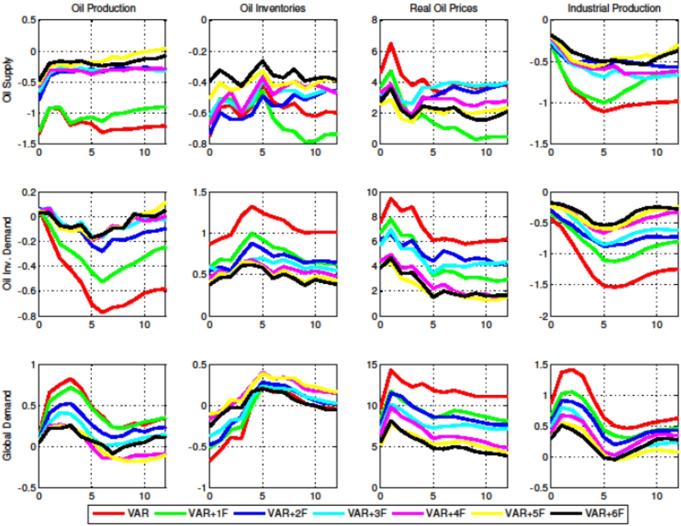


Subsample analysis: Historical decomposition



Go back

Robustness: number of factors



Go back

A simplified model of the oil market

The demand for oil (follows Hamilton, 2009)

- The demand for oil originates from the demand of gasoline retailers.
- Oil (X_t) is used as an intermediate input for the production of gasoline, whose real price is G_t
- $F(X_t, I_t)$ is the production function and depends on the current level of inventories, I_t.
- The first order conditions of the retailers' problem are:

$$F'_X(X_t, I_t) = \frac{P_t}{G_t}$$

$$P_t + \mathbb{C}'(I_{t+1}) = \frac{G_{t+1}F'_I(X_{t+1}, I_{t+1}) + P_{t+1}}{1 + r_t}$$

- Current oil production is either consumed for the production of gasoline or stored as inventories (for future production of gasoline): $\Delta I_{t+1} = Q_t - X_t$.
- To close the model we assume the following demand for gasoline

$$F\left(X_t, I_t\right) = \frac{\exp\left(\gamma_t\right)}{G_t^{\beta}}$$

- The inverse demand function for crude oil is given by

$$P_t = \left[\frac{\exp\left(\gamma_t\right)}{F\left(X_t, I_t\right)}\right]^{\frac{1}{\beta}} F'_X\left(X_t, I_t\right)$$

- The inverse demand function of inventories is

$$P_t = \frac{G_{t+1} F_I' \left(X_{t+1}, I_{t+1} \right) + P_{t+1}}{1 + r_t} - \mathbb{C}' \left(I_{t+1} \right)$$

- The total demand function for oil is

$$Q_t = D_{Inv} \left(P_t, P_{t+1}, G_{t+1}, X_{t+1}, r_t \right) - I_t + D_{Cons} \left(P_t, \gamma_t, I_t \right).$$

- This shows that a shift in the future oil price manifests itself into a shift in demand, specifically into an increase in the demand for inventories (Hamilton, 2009a, and Kilian and Murphy, 2011a).

Modeling Oil Extraction

- Q_t is the production or extraction of oil in period t
- $\mathbb{Q}_t$ is the cumulative extraction at the end of period t, so that: $\mathbb{Q}_t = \sum_{\tau=0}^{t} Q_\tau$
- Let $\Re_t$ be the amount of proven reserves so that the total amount of the resources exploitable at time t is $R_t = \Re_t - \mathbb{Q}_t$.
- Optimal extraction profile, $\{Q_\tau, R_\tau\}_{\tau=t}^{T}$, is obtained by maximizing the discounted stream of profits over the life of the field:

$$\Pi_t = \sum_{\tau=t}^{T} \frac{1}{\prod_{s=t}^{\tau}(1+r_s)} \left[P_\tau Q_\tau - C\left(Q_\tau, \mathbb{Q}_\tau\right)\right]$$

given the resource constraint

$$R_t = R_{t-1} - Q_t + e_t$$

where $e_t = \Re_t - \Re_{t-1}$

- Optimality condition for the extraction rate is

$$P_t - C'_Q\left(Q_t, \mathbb{Q}_t\right) - C'_{\mathbb{Q}}\left(Q_t, \mathbb{Q}_t\right) = \frac{P_{t+1} - C'_Q\left(Q_{t+1}, \mathbb{Q}_{t+1}\right)}{1 + r_t}$$

- If $C'_{\mathbb{Q}} = 0$, then the relation above is the Hotelling Principle: The price of the resource net of marginal extraction cost is expected to rise with the discount rate, r.
- If the firm were to face an increase in price ($P_{t+1} > P_t$), it would respond by decreasing the amount of current production, until the condition given in the equation was restored

Go back

Effect of US macro news on oil prices

We estimate:

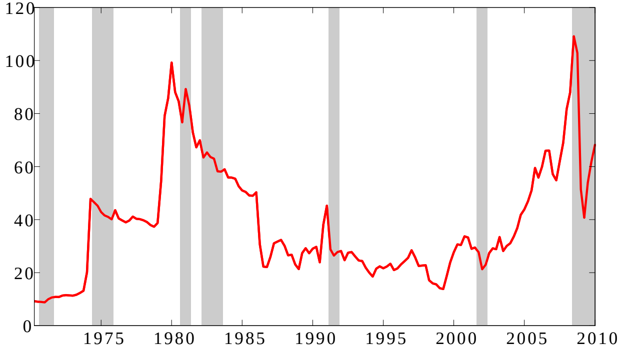
$$R_{t+1} = \alpha + \beta_i S_{it} + \epsilon_{t+1}$$

R_{t+1} : returns
S_{it} : news surprise

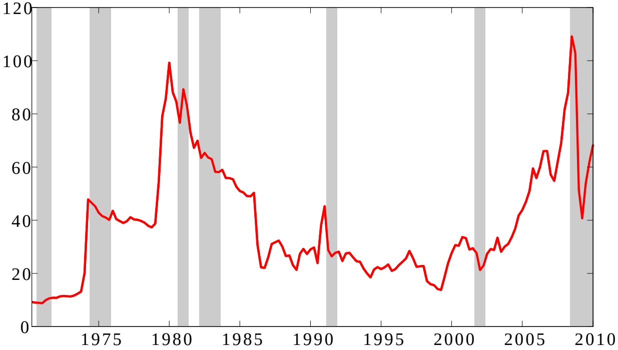
Table: Intra-Daily Crude Oil Prices

	Beta	p-Value	R2
GDP advance	0.203	0.018	0.167
Nonfarm payroll employment	0.246	0.000	0.121
Retail Sales	0.114	0.003	0.089
Industrial Production	0.072	0.003	0.085
Capacity Utilization	0.062	0.026	0.050
Factory orders	0.102	0.049	0.039
Trade Balance	0.056	0.014	0.062
NAPM Index	0.149	0.023	0.052
Index of leading indicators	0.167	0.015	0.059
Target federal funds rate	-0.150	0.008	0.035
Initial Unemployment Claims	-0.031	0.036	0.010

Real Oil Prices

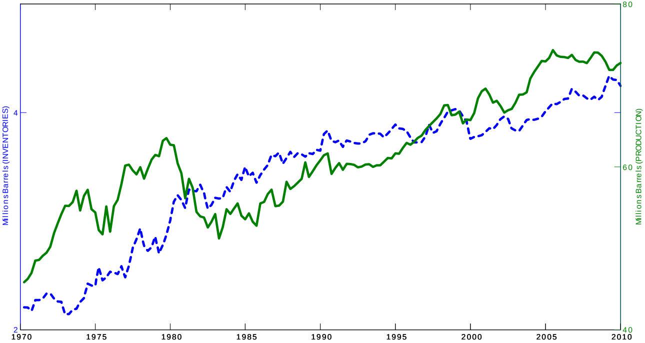


Real Oil Prices



Nine out of ten of the U.S. recessions since World War II were preceded by a spike up in oil prices (Hamilton 2005).

Oil Production and Inventories



Share of Non-Commercial Spread Positions in Reportable Open Interest: Futures Market