Oil Shocks: A Textual Analysis Approach

Deepa D. Datta   Daniel A. Dias

Federal Reserve Board

November 2020

The views expressed are those of the authors and do not necessarily reflect the views of anyone else in the Federal Reserve System.

Overview

We use textual analysis of a key oil industry news source to measure global supply and demand developments.

1. Develop a systematic and automated process to *read the news*, analogous to the informal approach we employ as Federal Reserve Board oil analysts.
    - Use Energy Intelligence Group's flagship publications "Oil Daily" and "International Oil Daily."
2. Provide a quantitative narrative of market conditions in real-time, as the news develops.
    - Earlier data availability than for oil production or IP.
    - Avoids reliance on alternative markets, e.g. metals, equities.
3. Use the publication archive to relate text-based indexes to supply, demand, and prices.
    - Identify supply- and demand-driven oil price dynamics using a structural VAR model.

Beyond article counts and sentiment analysis

- Use **phrase counts** to develop a quantitative narrative of the relative importance of supply and demand.
  - Baker, et al. (2016) construct EPU index by counting the number of articles that contain terms related to EPU in 10 international newspapers.
  - MPU, GPR, TPU indexes use similar methods (Husted, et al. (2019); Caldara and Iacoviello (2018); Caldara, et al. (2019)).

- Focus on **direction of supply and demand** to go beyond simple sentiment analysis for oil markets.
  - Loughran, et al. (2019) – short-horizon price movements related to sentiment in DJ Energy Service.
  - Cakir Melek, et al. (2019) – sentiment in Thomson-Reuters oil articles helps forecast oil prices.
  - Brandt and Gao (2019) – sentiment on macroeconomic and geopolitical news in RavenPack affects oil prices.

Full replicability, Straightforward updating

- Using natural language processing, we extract signals and construct the quantitative narrative from news articles, while **preserving replicability and straightforward updating**.
- Wu and Cavallo (working paper, 2012) similarly combine narrative and quantitative approaches to construct measures of oil price shocks.
  - Narrative approach involves human auditing of Oil Daily, Oil & Gas Journal, and Monthly Energy Chronology.
  - Attribute daily changes in oil prices to 22 types of oil-related events, e.g. weather changes, oil field discoveries, political and military actions, and changes in actual or expected inventories.
  - Aggregate select event types to generate exogenous oil shocks series, and show substantial effects.

Index Construction

Directional phrase counts

1. Construct vocabulary lists for supply, demand, increase, and decrease
2. Count the number of times a "supply" word is found in proximity to an "increase" word.
3. Repeat to obtain counts for "supply decrease," "demand increase," and "demand decrease."

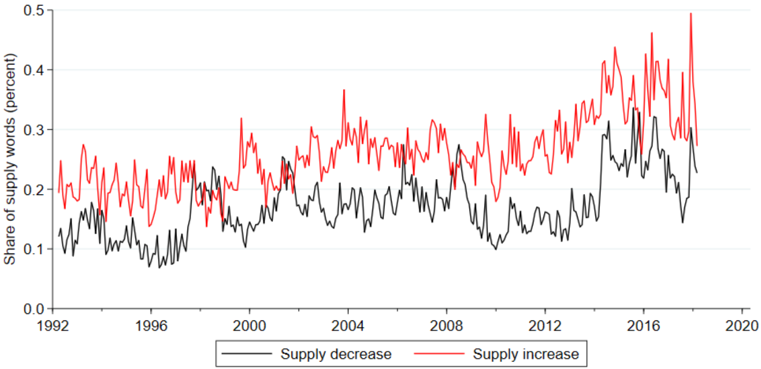
Index Construction

$$Index_t = \frac{\frac{PhraseCount_t}{WordCount_t} - Mean(\frac{PhraseCount_{1995-2004}}{WordCount_{1995-2004}})}{StDev(\frac{PhraseCount_{1995-2004}}{WordCount_{1995-2004}})}$$

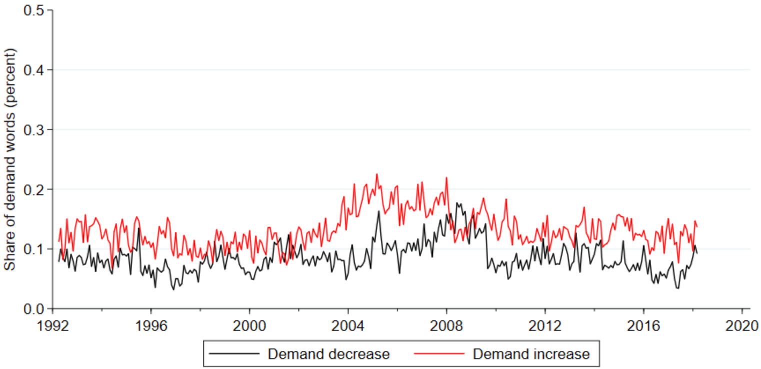
Obtain:

- 4 directional indexes
  1. SI_t - Supply Increase
  2. SD_t - Supply Decrease
  3. DI_t - Demand Increase
  4. DD_t - Demand Decrease
- 2 net indexes
  1. $Net\text{-}S_t$ - Net Supply
  2. $Net\text{-}D_t$ - Net Demand

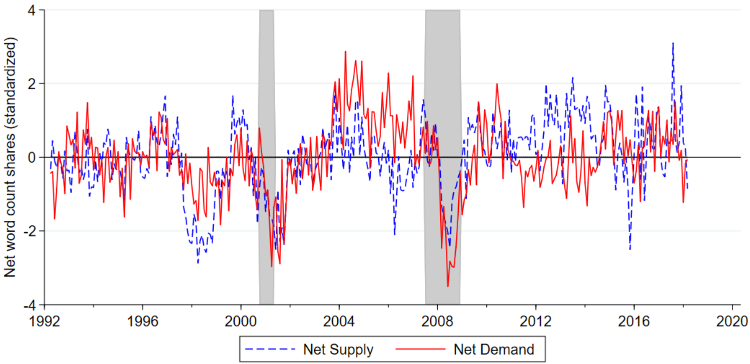
Directional Supply - Phrase Count Shares



Directional Demand - Phrase Count Shares



Net Supply and Net Demand Indexes



Oil Supply, Aggregate Demand, and Oil Prices

Contemporaneous regression:

$$y_t - y_{t-1} = \beta_0 + \beta_1 SI_t + \beta_2 SD_t + \beta_3 DI_t + \beta_4 DD_t + X_{t-1} + \epsilon_t$$

$$y_t - y_{t-1} = \beta_0 + \beta_1 NetS_t + \beta_2 NetD_t + X_{t-1} + \epsilon_t$$

Also consider *future* supply, demand, and oil prices:

$$y_{t+h} - y_t = \beta_0 + \beta_1 SI_t + \beta_2 SD_t + \beta_3 DI_t + \beta_4 DD_t + X_t + \epsilon_t$$

$$y_{t+h} - y_t = \beta_0 + \beta_1 NetS_t + \beta_2 NetD_t + X_t + \epsilon_t$$

X_{t-1} - Lagged values of changes in oil supply, aggregate demand, oil spot and futures prices
X_t - Current and lagged values of changes in oil supply, aggregate demand, oil spot and futures prices

Oil production, real economic activity, and prices

Table: Contemporaneous Movements

	Oil production		REA		WTI spot		WTI 12-month	
Supply increase	3.39 (0.86)***		0.64 (0.50)		-14.26 (6.55)**		-9.28 (5.42)*	
Supply decrease	-3.11 (0.73)***		-1.54 (0.44)***		5.72 (5.75)		3.25 (4.28)	
Demand increase	-0.82 (0.69)		1.68 (0.51)***		29.00 (7.85)***		23.23 (6.35)***	
Demand decrease	0.11 (0.54)		-1.40 (0.42)***		-15.59 (5.44)***		-12.85 (4.04)***	
Net supply		2.53 (0.66)***		0.43 (0.36)		-9.96 (4.98)**		-6.45 (3.93)
Net demand		-0.99 (0.57)*		0.76 (0.37)**		20.83 (5.98)***		16.85 (4.72)***
R^2	.265	.257	.429	.392	.333	.330	.340	.337
Indexes R^2 share	.148	.119	.249	.133	.104	.093	.184	.167

Note: *, **, and *** denote statistical significance at 10%, 5%, and 1%. Heteroskedasticity and autocorrelation corrected standard errors in parentheses. 296 observations.

$$y_t - y_{t-1} = \beta_0 + \beta_1 SI_t + \beta_2 SD_t + \beta_3 DI_t + \beta_4 DD_t + X_{t-1} + \epsilon_t$$

Oil production, real economic activity, and prices

Table: 12-Months Ahead

	Oil production		REA		WTI spot		WTI 12-month	
Supply increase	0.34		-1.22		-17.35		-13.03	
	(0.26)		(0.35)***		(5.32)***		(3.69)***	
Supply decrease	-0.55		0.30		14.59		10.21	
	(0.30)*		(0.36)		(4.07)***		(2.95)***	
Demand increase	0.10		1.26		11.88		10.39	
	(0.24)		(0.34)***		(3.76)***		(2.96)***	
Demand decrease	-0.08		-0.88		-6.85		-5.19	
	(0.19)		(0.43)**		(3.35)**		(2.54)**	
Net supply		0.26		-0.99		-13.38		-9.67
		(0.21)		(0.29)***		(4.15)***		(2.80)***
Net demand		-0.10		0.64		10.63		8.74
		(0.16)		(0.30)**		(2.73)***		(2.09)***
R^2	.315	.289	.368	.331	.362	.359	.274	.269
Indexes R^2 share	.143	.076	.207	.159	.422	.392	.458	.432

Note: *, **, and *** denote statistical significance at 10%, 5%, and 1%. Heteroskedasticity and autocorrelation corrected standard errors in parentheses. 296 observations.

$$y_{t+12} - y_t = \beta_0 + \beta_1 SI_t + \beta_2 SD_t + \beta_3 DI_t + \beta_4 DD_t + X_t + \epsilon_t$$

Estimating a structural VAR of the Oil Market

Key advantages:

1. Higher-frequency and more promptly available data
2. Current and prospective information on market conditions

Assume that in the short-run, the supply and demand sides of the oil market only interact with each other via prices.

$$\begin{cases} SI_t = \alpha_P^{SI} \Delta ln(P_t) + \epsilon_t^{S^+} \\ SD_t = \alpha_P^{SD} \Delta ln(P_t) + \epsilon_t^{S^-} \\ DI_t = \alpha_P^{DI} \Delta ln(P_t) + \epsilon_t^{D^+} \\ DD_t = \alpha_P^{DD} \Delta ln(P_t) + \epsilon_t^{D^-} \\ \Delta ln(P_t) = \gamma_{SI} SI_t + \gamma_{SD} SD_t + \gamma_{DI} DI_t + \gamma_{DD} DD_t + \epsilon_t^{NFP} \end{cases}$$

Weekly SVAR Model

Table: Short-Run Dynamics

	SI_t (1)	SD_t (2)	DI_t (3)	DD_t (4)	$\Delta ln(Price_t)$ (5)
Panel B: Weekly model					
SI_t	-	-	-	-	-3.807
					(0.58)***
SD_t	-	-	-	-	2.870
					(0.599)***
DI_t	-	-	-	-	6.030
					(0.733)***
DD_t	-	-	-	-	-4.159
					(0.386)***
$\Delta ln(Price_t)$	0.052	-0.023	-0.115	0.080	-
	(0.01)***	(0.01)**	(0.023)***	(0.012)***	
Sample			4/06/1994 - 10/10/2018		
Number observations			1280		

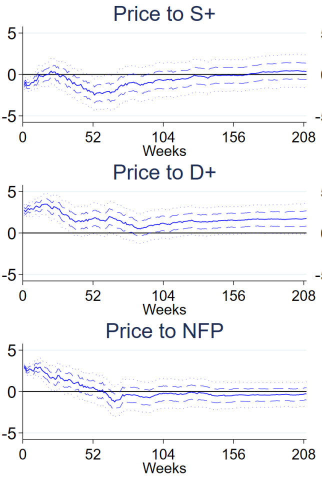
Weekly SVAR Model

Table: Forecast Error Variance Decomposition

Weekly model							
	S^+	S^-	S	D^+	D^-	D	NFP
$t = 1$	7.3	4.2	11.5	32.7	16.5	49.2	39.3
$t = \infty$	12.9	11.0	23.8	25.9	17.8	43.6	32.6

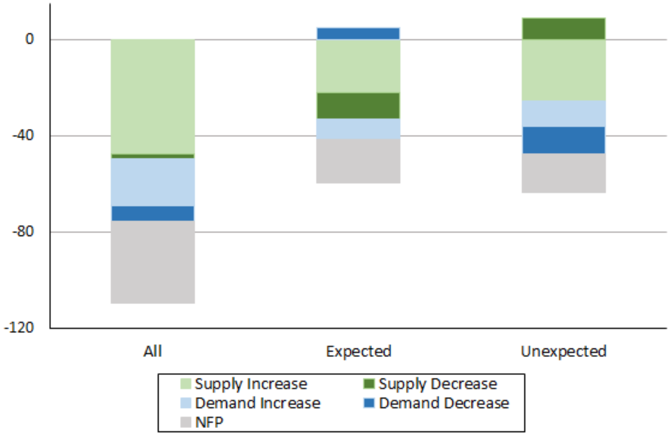
Weekly SVAR Model - Impulse Responses

Figure: Oil Price Response to Structural Shocks



The Oil Price Collapse of 2014

Figure: Oil Price Change Decomposition of the Oil Price Collapse of 2014



Note: Each bar corresponds to the sum of all the structural shocks causing oil prices to change between end May 2014 and end February 2016.

Conclusions

- Using textual analysis, we construct indexes containing information about supply and demand developments in the oil market.
  - Our indexes correlate well with existing measures of oil supply, demand, and prices.
  - The indexes contain substantial information about current and future oil price movements.

- Used the new indexes to estimate a structural VAR model of the oil market.
  - Results are in line with economic theory and are of plausible magnitudes.
  - Historical decomposition of well-known episodes in the oil market provide further evidence that our indexes contain substantial information about prospective oil price movements.