Stock-Oil Comovement: Fundamentals or Financialization? *

Alessandro Melone† Otto Randl‡ Leopold Sögner§ Josef Zechner¶

First draft: August 2020
This version: December 2021

Abstract

We investigate the sources of time-variation in the stock-oil correlation over the period 1983–2019. We first derive a novel oil futures return news decomposition following Campbell and Shiller (1988) and Campbell (1991). Then, for both stocks and oil, we split unexpected returns into cash flow news (which can be related to asset-specific fundamentals) and discount rate news (which can be driven by shocks to investors holding both assets) using a vector autoregressive (VAR) model. We find that about 79% of the time-varying correlation is related to the comovement of cash flow news between the two assets. This result is robust to different specifications of the VAR model used to decompose returns. We provide supportive evidence that underlying changes in the structure of the real economy, such as the increased oil production in the U.S., are key drivers for the changing stock-oil comovement beyond the financialization of commodities.

JEL classification: G12, Q43, E44.
Keywords: Asset comovements, oil futures, VAR return decomposition, real economy.

*We thank Suleyman Basak, John Campbell, Itay Goldstein, Semyon Malamud, Anna Pavlova, Alberto Rossi, Eduardo Schwartz, Veronika Selezneva (discussant), Ivan Shaliastovich, Valeri Sokolovski, and Michael Weber, and participants at the VGSF Conference 2018, the 33rd Australasian Finance & Banking Conference, the 2020 Commodity and Energy Market (CEMA) Annual Conference, the 2021 German Finance Association (DGF) Annual Meeting, the 2021 Financial Management Association (FMA) Annual Meeting, the 2021 Workshop of the Austrian Working Group on Banking and Finance (AWG) for insightful discussions and helpful comments. Alessandro Melone acknowledges financial support from the Austrian Science Fund (FWF).

†Vienna Graduate School of Finance (VGSF), alessandro.melone@vgsf.ac.at.
‡WU Vienna University of Economics and Business, otto.randl@wu.ac.at.
§Institute for Advances Studies (IHS) and VGSF, soegner@ihs.ac.at.
¶WU Vienna University of Economics and Business and VGSF, josef.zechner@wu.ac.at.

"[...] *The tendency of stocks to fall along with oil prices is surprising. The usual presumption is that a decline in oil prices is good news for the economy, at least for net oil importers like the United States and China*[...]"

Ben S. Bernanke (2016)

1 Introduction

Although oil continues to be the most important source of energy for the global economy, the degree to which stock market returns and returns from investing in crude oil comove varies greatly over time. Figure 1 shows the time-varying correlation between the U.S. stock market and crude oil excess returns for the period March 1983 to December 2019.[1] In particular, the stock-oil comovement appears strikingly different before and after 2008. Before 2008, the unconditional correlation between the two assets is about −8%. Following the statement by Bernanke (2016), a negative correlation is what one would expect to find.[2] For the non-oil producing sectors of an economy, a drop in oil prices should be good news. However, after 2008, the unconditional stock-oil correlation is +55%, contributing to a slightly positive correlation over the entire sample of about +14%. This fact is not only puzzling from an economic perspective (for example, see Draghi, 2017), but it also naturally poses concerns about the diversification benefit traditionally pursued by investors when investing in commodities such as oil (e.g., Gorton and Rouwenhorst, 2006). The sources of this change over time of the correlation between stocks and oil are still a matter of debate.

[1] For stocks, we use Center for Research in Securities Prices (CRSP) value-weighted excess returns; oil excess returns come from holding a long position in WTI futures. We compute the conditional correlation by a DCC-GARCH(1,1) specification as outlined in Section 2.

[2] The puzzling positive correlation between stocks and oil in the last decade has been at the center of a lively policy debate. See, for example, IMF Blog, Obstfeld et al. (2016) and Vox EU/CEPR, Obstfeld et al. (2016). See also Mohaddes and Pesaran (2017) for a recent review on the relationship between oil markets and the economy.

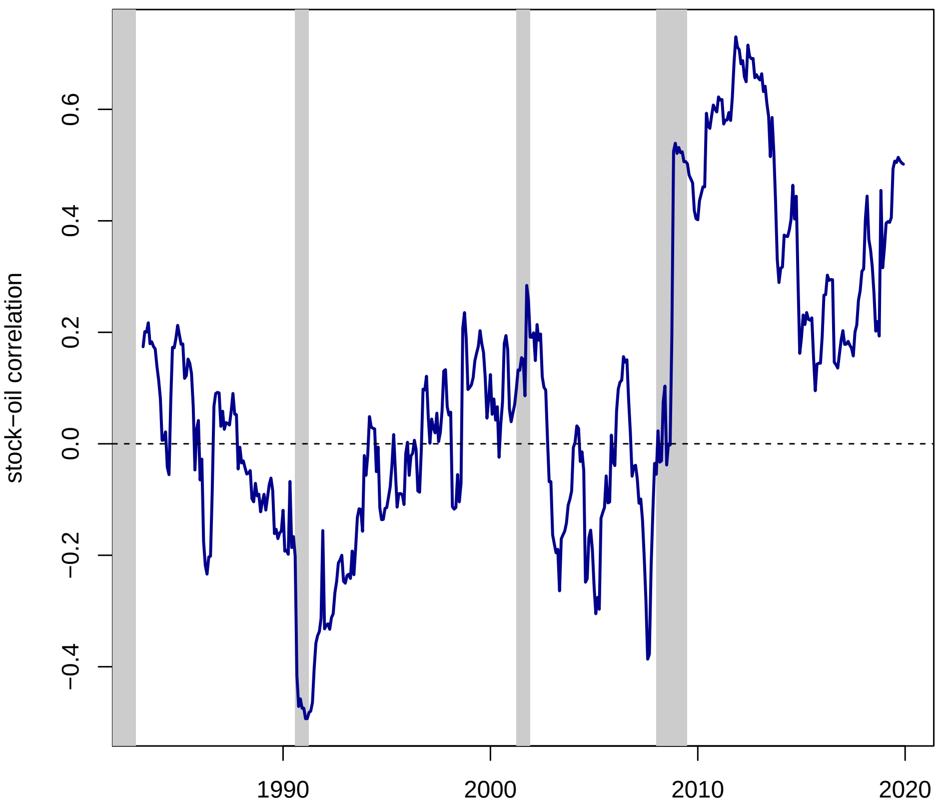

Figure 1: Stock-Oil Comovement

This figure shows time-varying correlation between equity and oil excess returns computed by a DCC-GARCH(1,1) specification. For U.S. equity, we use the CRSP value weighted index, and for oil WTI futures contracts. Shaded area are NBER recessions. The sample period is March 1983 to December 2019.

The main objective of our work is to shed light on the dynamics of stock-oil comovements. Specifically, we address what we label as "Bernanke's puzzle": Can we explain the time-varying correlation between stocks and oil? We are particularly interested in understanding why this correlation remains positive and high in recent years. We answer this question by exploring two main channels.

On one side, returns of two assets may correlate because their cash flows may be correlated. Thus, news about the cash flows generated by these two assets may be correlated, and this correlation potentially exhibits time variation. When pursuing this possible channel, one needs to recognize that oil does not produce cash flows such as dividends, as is the case for stocks. However, physically holding a commodity gives its owner a benefit, the convenience yield, e.g., from avoiding the risk of manufacturing disruptions in the presence of supply shortages. Several papers have shown that a convenience yield can arise endogenously as a result of the interaction between supply, demand, and storage decisions (e.g., Brennan, 1958; Deaton and Laroque, 1992; Routledge et al., 2000). One can compute the convenience yield from observed spot and futures prices. Estimates of the convenience yield show considerable time variation, with a tendency to be high in periods of low inventory levels.

Alternatively, stock and oil returns may correlate if the discount rates applied to the cash flows generated by the two assets are correlated. This view is particularly plausible since investors frequently hold both stocks as well as oil-related financial products (e.g., oil futures or commodity-ETFs) in their portfolios. Cheng and Xiong (2014) review recent studies and argue that the financialization of commodity markets starting around 2004 has impacted prices substantially through risk sharing and information discovery by financial investors. Singleton (2014) argues that informational frictions and the associated speculative activity can lead prices to drift away from fundamentals. The author finds that increases in flows into commodity index funds over the preceding three months predict higher subsequent crude oil futures prices. Hamilton and Wu (2014) provide further evidence that index-fund investing has increased in importance relative to commercial hedging of oil producers and/or consumers in determining crude oil futures risk premia. Thus, even if cash flow fundamentals of the two assets are independent of one another, we would observe comovement between stocks and oil, especially in light of the increasing financialization of commodities (e.g., Basak and Pavlova, 2016 or Goldstein and Yang, 2019). According to this view, a surge in speculation might make commodity markets more volatile and more highly correlated with other assets. We refer to this channel as the discount rate channel of return correlation.[3]

To investigate the importance of the two channels, we decompose equity and oil return innovations into cash flow and discount rate news. While our decomposition is novel in the context of the stock-oil correlation, its foundation is the present value relation for stocks in Campbell and Shiller (1988).[4] We estimate VAR models to quantify the news components and calculate the time-varying correlations both between equity and oil cash flow news and between equity and oil discount rate news. We find that the correlation between cash flow news plays a dominant role in explaining the overall dynamics between equity and oil return comovement, accounting for about 79% of the total correlation. The correlation of cash flow news shifts upwards with increased U.S. oil production, consistent with oil on average turning from from a cost factor into a revenue factor for the U.S. economy.

Our paper is complementary to recent work that connect oil, the stock market, and the real economy. Ready (2018b) finds that demand shocks correlate positively with stock market returns, while supply shocks correlate negatively. Using a different classification of

[3] Brogaard et al. (2019) find a strong feedback effect between the financialization of commodity markets and the real economy.

[4] Szymanowska et al. (2014) builds on Campbell and Shiller (1988) to explore the relation between the commodity basis and expected returns.

shocks, Hitzemann (2016) provides theoretical support for the findings of Kilian and Park (2009) that there is a positive relationship between oil price changes and stock market returns in response to aggregate growth shocks, but not in connection with oil-specific productivity shocks. Gilje et al. (2016) find that technology innovations related to the shale oil revolution are key drivers of the aggregate stock market. Ready (2018a) finds that changes in long-run oil supply uncertainty relate to changes in the market price of oil risk, and rationalizes this fact in a production-based long-run risk model. David (2019) derives a production equilibrium model of drilling, exploration, and storage to understand the different impact of oil firm investments on the oil futures basis and oil futures risk premia.

This paper also speaks to the literature that studies the relationship between commodity futures and the stock market. Boons and Szymanowska (2014) find that commodity risk is priced in the cross-section of U.S. stock returns with opposite signs before and after the financialization of commodity markets. Increased commodity index investing potentially has adverse effects on firm performance. Brogaard et al. (2019) provide evidence for declining price informativeness of index commodities that leads to lower sensitivity of commodity using firms' investment to commodity futures prices. Fernandez-Perez et al. (2017) show that backwardated and contangoed portfolios predict the aggregate stock market as they capture long-run changes in investment opportunities. Christoffersen and Pan (2018) find that oil price volatility significantly predicts stock returns both in the time-series and in the cross-section, after the financialization of commodities.

The remainder of the paper is organized as follows. In Section 2, we describe our methodology and derive a decomposition of oil returns into their cash flow news and discount rate news components. We present our empirical analysis in Section 3. Finally, Section 4 concludes.

2 Methodology

To formally explore the time-varying relationship between stock and oil excess returns, we proceed in the following steps. First, we use a Dynamic Conditional Correlation Generalized Autoregressive Conditional Heteroskedasticity (DCC-GARCH) model to estimate the stock-oil correlation over time. Second, we decompose unexpected returns into their cash flow news and discount rate news components. Finally, we estimate the time-varying correlation of these components using DCC-GARCH and explore how much of the total correlation can be explained by cash flow news correlation and discount rate news correlation.

2.1 Estimation of time-varying correlations

Reliable estimates of the joint dynamics of equity and oil have to take into account that all components of the covariance between equity and oil excess returns are time-varying: both assets' standard deviation and their correlation. While simple methods such as rolling historical correlations and exponential smoothing are widely used, these methods are not necessarily consistent with the stochastic properties of the individual time series. For example, rolling window estimation gives equal weight to all observations from the periods used and zero weight to all other observations. Thus, results depend on the somewhat arbitrary choice of the interval length used for estimation. An intuitive alternative that is the DCC-GARCH (1,1) method suggested by Engle (2002). This method decomposes the covariance into its components, all of which are dynamic. Estimation follows a two-step procedure.

First, for each of the two assets its volatility is estimated using a univariate GARCH(1,1) specification. Second, estimation of the time-varying correlation uses the standardized errors of the first step, estimates the parameters of a generalized autoregressive proxy process, and rescales the result to ensure the correlation is in the interval [-1,+1]. An appealing property of the estimation procedure is that multivariate and univariate measures of variability are consistent with each other. The time-varying covariance of equity and oil excess returns is the product of the conditional standard deviations and the conditional correlation; all three components are mean reverting processes.

2.2 Oil return decomposition

In order to better understand the time variation in correlations, we split returns into components that we can analyze separately. For both stocks and oil, unexpected returns must be associated with changes in expectations of future cash flows or discount rates. As stated for equity returns by Campbell and Vuolteenaho (2004), the component of unexpected returns that is generated by cash flow news can be considered permanent, while returns generated by news on discount rates are transitory. This is because a positive shock on the discount rate leads to capital losses today, but these will be offset by improved investment opportunities. Thus it is important for investors to understand if commonalities in the cash flows or discount rates drive time-variation in the correlation.

We derive a novel decomposition for oil return news into discount rate news and cash flow news. Our decomposition builds on the seminal work by Campbell and Shiller (1988) and Campbell (1991). The Campbell-Shiller log-linearization allows to decompose the variation in stock log-returns[5] r^{eq}_{t+1} into revisions in expectations of dividend growth Δd_{t+1} and future returns:

$$r^{eq}_{t+1} - E_t(r^{eq}_{t+1}) = \underbrace{(E_{t+1} - E_t)\left[\sum_{j=0}^{\infty} \varrho^j \Delta d_{t+1+j}\right]}_{N^{eq}_{CF,t+1}} - \underbrace{(E_{t+1} - E_t)\left[\sum_{j=1}^{\infty} \varrho^j r^{eq}_{t+1+j}\right]}_{N^{eq}_{DR,t+1}}, \quad (1)$$

where $\varrho = \frac{1}{1+\exp(\overline{d-p})} < 1$ is a log-linearization parameter that depends on the average dividend-to-price ratio $\overline{d-p}$.[6]

The cost-of-carry model for the oil futures price (e.g., Fama and French, 1988; Szymanowska et al., 2014) is an identity relating futures to spot prices, the risk-free rate, storage costs, and the convenience yield. As in Szymanowska et al. (2014), for ease of exposition, we assume that the risk-free rate and storage costs are constant. Our specification then is:

$$F_{t,t+1} = S_t\delta - C_{t+1}, \quad (2)$$

where $F_{t,t+1}$ is the futures price at time t with delivery at time $t+1$, S_t is the oil spot price at time t, δ is a constant containing the risk-free interest rate and the rate for storage costs from t to $t+1$, and C_{t+1} is the income from convenience yield. We model the convenience yield as a cash payment to the commodity owner that occurs at time $t+1$ but it is known at time t.[7] It can be interpreted as the implicit benefits that accrue only to holders of physical

[5] Let P_t denote the asset price in period t, $p_t := \ln P_t$, D_t the corresponding dividend payment, and $d_t := \ln D_t$ (where ln denotes the natural logarithm). Then, $r^{eq}_{t+1} := \ln\left(P^{eq}_{t+1} + D_{t+1}\right) - \ln P^{eq}_t$. In addition, $\Delta d_{t+1} := d_{t+1} - d_t$. $E_t(\cdot)$ abbreviates the conditional expectation,

[6] A similar return decomposition has been proposed for bonds by Campbell and Ammer (1993). Recently, Balduzzi and Chiang (2020) and Chiang and Mo (2019) propose a currency returns decomposition.

[7] This assumption is made also in Szymanowska et al. (2014).

inventories but not to owners of futures contracts. For example, in manufacturing holding inventory can avoid costly time delays and short-term changes in output. As inventories can never be negative, one cannot borrow from future oil supply to fulfill current demand. In an alternative interpretation, the cash payment from the convenience yield stems from the value of the option to sell oil out of storage. To equalize returns from owning oil physically or via owning long positions in oil futures contracts, the latter therefore trade at a discount to the (interest- and storage-cost-adjusted) spot price. At maturity, that discount disappears such that $F_{t+1,t+1} = S_{t+1}$. Hence, while the cash flows from holding commodities are difficult to quantify directly, they can be quantified with the convenience yield implied in the futures curve. In the literature, the convenience yield has been either modeled as an exogenous "dividend" process (e.g., Gibson and Schwartz, 1990; Cassacus and Collin-Dufresne, 2005) or modeled as the endogenous result of the interaction among supply, demand, and storage decisions (e.g., Routledge et al., 2000; Gorton et al., 2013).[8]

From Eq. (2), we compute the excess return on the futures contract that matures at time $t+1$:

$$R_{t+1}^{oil} = \frac{F_{t+1,t+1}}{F_{t,t+1}} = \frac{S_{t+1}}{S_t\delta - C_{t+1}}.$$

Taking logs, we have:

$$r_{t+1}^{oil} = \frac{\ln S_{t+1}}{\ln(S_t\delta - C_{t+1})} = s_{t+1} - \ln(S_t\delta - C_{t+1}),$$

where lowercase letters denote the natural logarithm of the corresponding uppercase letters. To make sure we take the log of a positive number, we make the plausible assumption that

[8] In addition to the convenience yield, the risk premium determines futures prices as the futures price at time t is the expected future spot price under the risk neutral measure, i.e. $F_{t,t+1} = E_t^Q(S_{t+1})$.

the current spot price of oil, adjusted by interest rates and storage cost, exceeds the cash income from the convenience yield ($C_{t+1} < \delta S_{t+1}$). Taking S_t out of log we get:

$$r_{t+1}^{oil} = s_{t+1} - s_t - \ln(\delta - \exp(c_{t+1} - s_t)). \tag{3}$$

The last term in Eq. (3) is non-linear. Following Campbell and Shiller (1988), we log-linearize this term by a first-order Taylor approximation around the long-run average convenience yield-price ratio $\overline{c} - \overline{s} \equiv \overline{cs}$. Thus, we obtain:

$$\ln(\delta - \exp(c_{t+1} - s_t)) \approx k + (1 - \rho)(c_{t+1} - s_t),$$

with $\rho = \frac{\delta}{\delta - \exp(\overline{cs})}$ and $k = \ln \delta - \ln \rho + (1 - \rho)\overline{cs}$.

We can now re-write Eq. (3) as an approximately linear relationship:

$$r_{t+1}^{oil} \approx -k + s_{t+1} - \rho s_t - (1 - \rho)c_{t+1}. \tag{4}$$

From Eq. (4), we obtain the identity for the oil spot price:

$$s_t = \tilde{k} + \theta s_{t+1} - \theta r_{t+1}^{oil} + (1 - \theta)c_{t+1}, \tag{5}$$

where $\theta = 1/\rho$ and $\tilde{k} = -k/\rho$; subtracting from both sides c_t, we have:

$$s_t - c_t = \tilde{k} + \theta(s_{t+1} - c_{t+1}) - \theta r_{t+1}^{oil} + \Delta c_{t+1}, \tag{6}$$

with $\Delta c_{t+1} = c_{t+1} - c_t$.

Iterating Eq. (6) forward and imposing the transversality condition $\lim_{k\to\infty}\theta^k(s_{t+k}-c_{t+k})=0$ (no rational bubbles), we get:

$$s_t - c_t = \tilde{k} + \sum_{j=1}^{\infty}\theta^{j-1}\Delta c_{t+j} - \sum_{j=1}^{\infty}\theta^j r^{oil}_{t+j}. \tag{7}$$

For $0<\theta<1$, we need $\rho>1$. This is the case when the average income from the convenience yield is positive ($\overline{cs}>0$). Eq. (7) says that the log oil price to convenience yield ratio is high when future yields are expected to increase fast, or when returns from holding long positions in oil futures are expected to be low. Intuitively, if the oil price is high today, there must either be high utility from owning oil now or low returns from holding oil in the future.[9]

Finally, rearranging Eq. (7) to express r^{oil}_{t+1}, and taking conditional expectations, we obtain:[10]

$$r^{oil}_{t+1} - E_t(r^{oil}_{t+1}) = \underbrace{(E_{t+1}-E_t)\left[\sum_{j=0}^{\infty}\theta^{j-1}\Delta c_{t+1+j}\right]}_{N^{oil}_{CF,t+1}} - \underbrace{(E_{t+1}-E_t)\left[\sum_{j=1}^{\infty}\theta^{j} r^{oil}_{t+1+j}\right]}_{N^{oil}_{DR,t+1}}. \tag{8}$$

Excess returns from investing in commodity futures can therefore be decomposed into revisions of expectations about future cash flows (i.e., convenience yields) and revisions of expectations about future discount rates (i.e, returns from holding commodity futures). We call the first and the second term on the right-hand side of Eq. (8) oil cash flow news and oil

[9] Szymanowska et al. (2014) log-linearize oil futures return around the mean basis to show that the basis contains information about risk premia.

[10] Recall that $E_t(c_{t+1}) = c_{t+1}$. Hence, $(E_{t+1}-E_t)\,\theta^{-1}\Delta c_{t+1} = 0$. Equation (8) follows the notation in Szymanowska et al. (2014) and includes the "$j = 0$-term" for cash flow news.

discount rate news, respectively. As in the corresponding decomposition for equity returns, the sign is positive on cash flow news and negative on discount rate news: An increase in expected cash flows is good news for the holder of an asset, while an increase in the discount rate reduces its value.

2.3 VAR

To implement the decomposition of returns into their cash flow news and discount rate news components, we use VAR models. This approach has been introduced by Campbell (1991) for equity returns and has been used widely thereafter. We follow the description in Campbell and Vuolteenaho (2004) and Campbell et al. (2018) and apply their VAR methodology to estimate the news components for stock returns. In addition, we apply a variant of this VAR methodology for estimation of cash flow and discount rate news of oil returns.

The VAR first estimates the conditionally expected oil excess return $E_t(r_{t+1}^{oil})$. Using Eq. (8), we can then split the unexpected component of excess returns $r_{t+1}^{oil} - E_t(r_{t+1}^{oil})$ into its components discount rate news $N_{DR,t+1}^{oil}$ and cash flow news $N_{CF,t+1}^{oil}$.[11] We specify a stationary first-order VAR model as

$$\mathbf{z}_{t+1} = \mathbf{a} + \boldsymbol{\Gamma}\mathbf{z}_t + \mathbf{u}_{t+1}\,, \tag{9}$$

where $\mathbf{z}_{t+1}$ is a state vector with r_{t+1}^{oil} as its first element and variables that are useful for oil return prediction as its further elements. $\mathbf{a}$ is a vector of constants, $\boldsymbol{\Gamma}$ a matrix of constant parameters, and $\mathbf{u}_{t+1}$ a vector of i.i.d. shocks. The assumption that $\mathbf{z}_{t+1}$ follows a first-order

[11] Because of the approximate identity linking returns, the convenience yield, and the oil price this approach is almost equivalent to forecasting cash flows explicitly using the same forecasting variables.

VAR is not restrictive, because a higher-order VAR can be rewritten as a first-order VAR with an expanded state vector.[12] Provided that the data generating process is described by Eq. (9), the unexpected oil excess return is given by $r^{oil}_{t+1} - E_t(r^{oil}_{t+1}) = \mathbf{e}_1' \mathbf{u}_{t+1}$, where $\mathbf{e}_1'$ is a row vector with its first element equal to one and all other elements equal to zero. From Eq. (8), we can express discount rate news and cash flow news as

$$\begin{aligned} N^{oil}_{DR,t+1} &= \mathbf{e}_1' \boldsymbol{\lambda} \mathbf{u}_{t+1} \\ N^{oil}_{CF,t+1} &= \left(\mathbf{e}_1' + \mathbf{e}_1' \lambda\right) \mathbf{u}_{t+1} \ , \end{aligned} \tag{10}$$

where $\boldsymbol{\lambda} \equiv \boldsymbol{\theta} \boldsymbol{\Gamma} \left(\mathbf{I} - \boldsymbol{\theta} \boldsymbol{\Gamma}\right)^{-1}$.

The formula for discount rate news can be seen from $N^{oil}_{DR,t+1} = (E_{t+1} - E_t) \left[\sum_{j=1}^{\infty} \theta^j r^{oil}_{t+1+j}\right] = \mathbf{e}_1' \left[\sum_{j=1}^{\infty} \boldsymbol{\theta}^j \boldsymbol{\Gamma}^j \mathbf{u}_{t+1}\right]$ and the fact that $N^{oil}_{CF,t+1} = \left(r^{oil}_{t+1} - E_t(r^{oil}_{t+1})\right) + N^{oil}_{DR,t+1}$. The interpretation of Eq. (10) mirrors the description in Campbell and Vuolteenaho (2004) for the equity decomposition and VAR estimation. $\mathbf{e}_1' \boldsymbol{\lambda}$ captures the long-run importance of VAR shocks to discount rate expectations. Variables with large absolute values of the coefficient in the returns prediction equation (top row of $\boldsymbol{\Gamma}$) are important for discount rate news. Further, the term $\left(\mathbf{I} - \boldsymbol{\theta} \boldsymbol{\Gamma}\right)^{-1}$ shows that more persistent variables receive more weight.

2.4 Dissecting the time-varying correlation

The decomposition of both stock market excess returns and oil futures excess returns into its cash flow news and discount rate news components allows to further investigate the time-varying correlation of equity and oil returns. To do so, we use again the DCC-GARCH(1,1) to model the time-varying correlation between equity cash flow news and oil cash flow news

[12]While our main analysis uses a VAR(1) specification, we report results from the estimation of a VAR(2) specification in Section 3.5.

on the one hand, and equity discount rate news and oil discount rate news on the other hand. We then use OLS regressions to see how much variation in the stock-oil correlation can be explained by the variation in the correlations of the cash flow news components alone. Similarly, we check the explanatory power of the discount rate news components. In principle, the correlation could be driven by a combination of common cash flow news and common discount rate news. This is because supply shocks might have opposing effects on prices of stocks and oil, but positive demand shocks are likely to boost both oil prices and corporate profitability. The correlation of discount rate shocks will depend on the extent to which oil is seen as an investment asset (financialization) and the degree of integration of these markets.

3 Empirical Analysis

3.1 Data for returns and yields

We use monthly log excess returns from CRSP for U.S. equity. To compute oil excess returns, we follow standard procedures, as outlined in Szymanowska et al. (2014) and Bakshi et al. (2019). We use end-of-month prices for all traded WTI futures contracts (ticker CL) downloaded from Bloomberg for the period March 1983 to December 2019. To avoid using thinly traded contracts, we exclude prices of futures contracts observed a month prior to and during the delivery month. In each month, we then select the futures contract with the second shortest time to maturity $t+2$ and calculate its log excess return over the next month as $r_{t+1}^{oil} = \ln F_{t+1,t+2} - \ln F_{t,t+2}$.

As WTI futures are available with monthly maturity, this calculation corresponds to rolling into the next futures contract on a monthly basis.[13]

Table 1 reports summary statistics for stocks and oil excess returns, while Figure 2 shows monthly log index levels and excess returns. For the U.S. equity market, we use the CRSP value-weighted index, and for the oil price level the WTI futures contract with the shortest time to maturity. Over the period from March 1983 to December 2019, investors have earned an annualized 7.8 percent excess return from taking long positions in WTI oil futures contracts. This is lower than the 8.3 percent excess return that investors could earn in the U.S. stock market. While over the period analyzed equity has an higher Sharpe Ratio than oil, oil has nevertheless also turned out to be an attractive investment.

Table 1: Descriptive Statistics CRSP and WTI

This table reports summary statistics for stocks (r_t^{eq}) and oil (r_t^{oil}) monthly excess returns. Numbers are given in percent. The sample period is March 1983 to December 2019.

Statistic	Min	Max	Mean	St. Dev.	Skew.	Kurt.
r_t^{eq}	−23.24	12.47	0.69	4.33	−0.92	2.72
r_t^{oil}	−32.37	46.01	0.65	9.40	0.31	2.24

3.2 VAR state variables

For estimation of the equity VAR, we closely follow Campbell and Vuolteenaho (2004) and Campbell et al. (2018). All data are measured at a monthly frequency.

[13] The strategy implied by this return calculation makes sure that positions are closed before a contract's first notice day, thus ensuring that the investor never faces physical delivery.

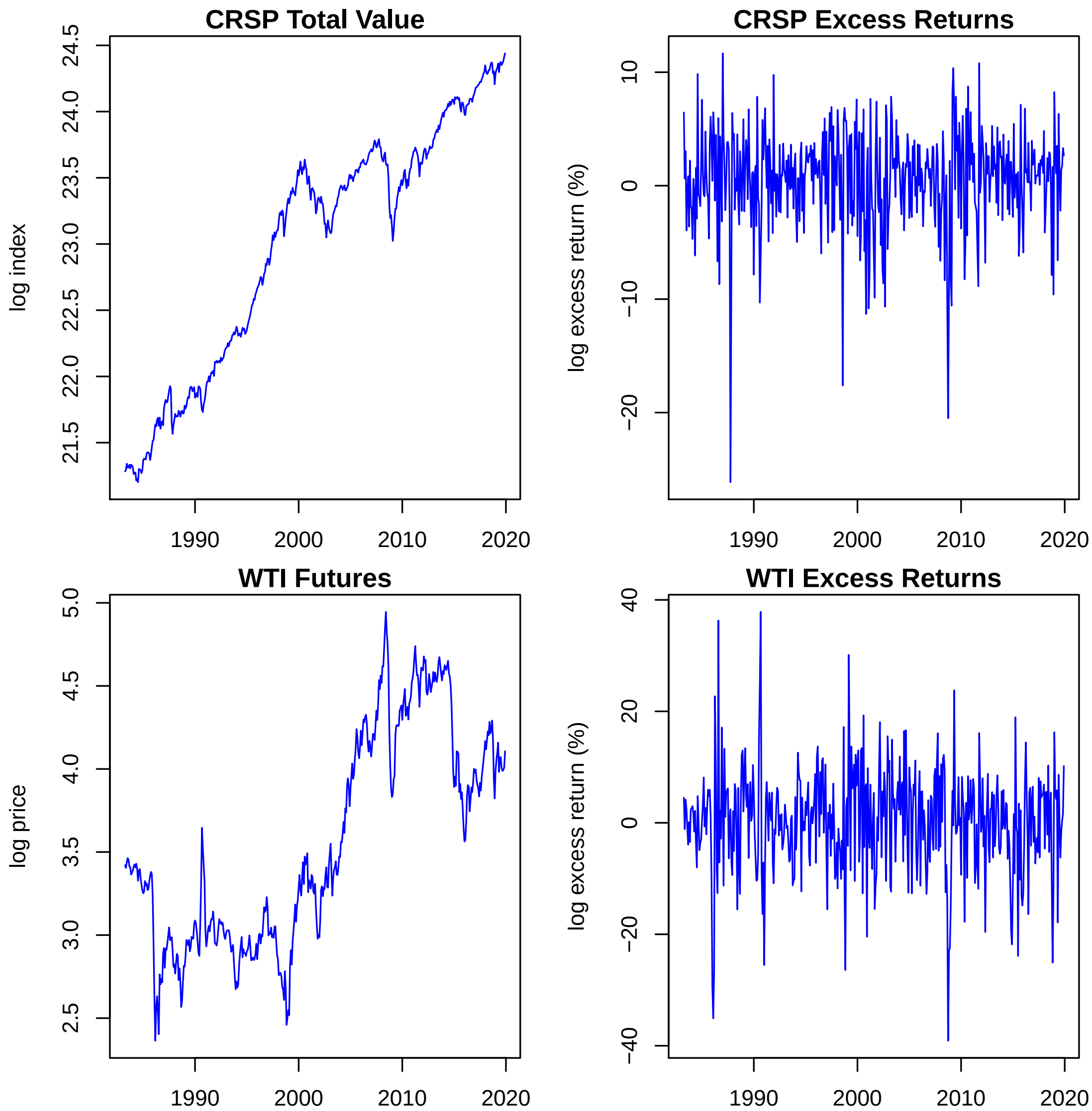

Figure 2: Stocks and Oil Dynamics. This figure shows monthly log prices and log excess returns for stocks and oil. For U.S. equity, we use the CRSP value weighted index, and for oil WTI futures contracts. The sample period is March 1983 to December 2019.

Return. The return to be predicted in the equity VAR is r^{eq}, the excess log return on the CRSP value-weighted stock market index. The risk-free rate is constructed from Treasury bills with approximately three months maturity, obtained from CRSP.

Term spread. To capture the impact of the business cycle on equity returns, we include the term yield spread *termsp*, measured as the yield difference between ten-year constant-maturity taxable bonds and short-term taxable notes.

Price-earnings ratio. *cape* is the log ratio of the Shiller (2000) S&P 500's price to the S&P 500's ten-year moving average of earnings. Earnings are averaged over multiple years to avoid temporary spikes in the price-earnings ratio caused by cyclical declines in earnings. The rationale for including *cape* as a potential return predictor is that with constant earnings growth high price-earnings ratios imply low long-run expected returns.

Small stock value spread. Next, we include the small stock value spread *valuesp* in the VAR. High returns to small growth stocks should forecast low returns on the aggregate stock market if their intertemporal hedging value is the reason for their low average returns, as their returns must be negatively correlated with investment opportunities. *valuesp* is based on two of the six elementary portfolios constructed by Davis et al. (2000). At the end of June of every year, we construct the small-stock value spread as the difference between the log book-to-market ratios of small value and small growth stocks, where book equity and market equity are measured at the end of December of the previous calendar year. For months from July to May, the small-stock value spread is constructed by adding the cumulative log return from June on the small low-book-to-market portfolio to, and subtracting the cumulative log return on the small high-book-to-market portfolio from, the end-of-June small stock value spread.

Default spread. Finally, we add the default spread as suggested in Campbell et al. (2018). $defsp$ is the difference between the log yield on Moody's BAA bonds and the log yield on Moody's AAA bonds, and should provide information on future corporate profits. In particular, shocks to the default spread reflect news about aggregate default probabilities and should therefore in turn reflect news about the market's future cash flows.

In Table 2, we provide descriptive statistics (Panel A) and correlation matrices (Panel B) for the equity VAR state variables. The sample period is from March 1983 to December 2019. Table 3 reports the OLS parameter estimates for the first-order VAR model on equity excess returns. The R^2 of 1.5 percent is in line with typical values for monthly equity return prediction. The lagged excess return of the market predicts with a positive sign and is barely significant. The most important variable is *cape*: High valuation levels predict low returns.

For the oil VAR, we select state variables that plausibly predict returns from investing in oil commodity futures.

Return. The return to be predicted in the VAR is r^{oil}, the excess log return on the WTI futures contract. The contract used for return calculation is the contract with the second shortest time to maturity. A strategy that purchases this contract at time t and sells it at $t + 1$ (with a time-to-maturity one month shorter) requires no cash investment and gives directly the excess return.

Table 2: Equity VAR State Variables

This table reports the descriptive statistics for the equity VAR state variables over the whole sample from March 1983 to December 2019, in percentage points. r_t^{eq} is the excess log return on the CRSP value-weight index. $termsp_t$ is the term yield spread measured as the yield difference between ten-year constant-maturity taxable bonds and short-term taxable notes. $cape$ is the log ratio of the S&P 500's price to the S&P 500's ten-year moving average of earnings. $valuesp_t$ is the small-stock value-spread, the difference in the log book-to-market ratios of small value and small growth stocks. The small-value and small-growth portfolios are two of the six elementary portfolios constructed by Davis et al. (2000). $defsp_t$ is the difference between the log yield on Moody's BAA bonds and the log yield on Moody's AAA bonds.

Panel A: Descriptive Statistics.

Statistic	Mean	St. Dev.	Min	Median	Max
$termsp_t$	1.78	1.09	−0.61	1.77	3.91
$cape_t$	24.58	6.99	11.72	24.62	44.20
$valuesp_t$	1.59	0.19	1.15	1.57	2.15
$defsp_t$	0.91	0.35	0.51	0.85	3.17

Panel B: Correlation Matrix.

	r^{eq}	$termsp$	$cape$	$valuesp$	$defsp$
r_t^{eq}	1	−0.030	0.037	0.100	−0.098
$termsp_t$	−0.030	1	−0.414	0.319	0.248
$cape_t$	0.037	−0.414	1	0.159	−0.531
$valuesp_t$	0.100	0.319	0.159	1	−0.188
$defsp_t$	−0.098	0.248	−0.531	−0.188	1

Table 3: VAR Parameter Estimates for Stocks

This table reports the OLS parameter estimates for a first-order VAR model including a constant, r_t^{eq}, $termsp_t$, $cape_t$, $valuesp_t$, and $defsp_t$. Values in parenthesis are OLS standard errors. ***, **, and * indicates respectively 1%, 5%, and 10% level of significance. The sample period is March 1983 to December 2019.

	r_{t+1}^{eq}	$termsp_{t+1}$	$cape_{t+1}$	$valuesp_{t+1}$	$defsp_{t+1}$
r_t^{eq}	0.060	0.005	0.096***	−0.000	−0.008***
	(0.050)	(0.003)	(0.008)	(0.001)	(0.001)
$termsp_t$	−0.423*	0.964***	−0.056	0.012***	−0.001
	(0.243)	(0.016)	(0.041)	(0.004)	(0.005)
$cape_t$	−0.098**	−0.000	0.983***	0.001*	0.000
	(0.039)	(0.002)	(0.007)	(0.001)	(0.001)
$valuesp_t$	2.062	−0.046	0.090	0.889***	−0.033
	(1.277)	(0.082)	(0.217)	(0.023)	(0.025)
$defsp_t$	−0.555	0.076*	−0.058	0.005	0.952***
	(0.697)	(0.045)	(0.118)	(0.013)	(0.014)
Constant	0.921	0.069	0.419	0.120***	0.100**
	(2.216)	(0.142)	(0.376)	(0.040)	(0.043)
R^2	0.024	0.934	0.989	0.830	0.939

Basis. We include the *basis* as an important commodity return predictor. It is defined as the log ratio of the spot price to the futures price. The futures price is the market price at time t of the contract used for calculation of the return r^{oil}. We follow common practice to use a short-maturity futures contract as a proxy for the spot price.[14] Specifically, we use the outstanding contract that has the closest but shorter maturity than the contract used for return calculation. The economic rationale for including the *basis* is its relation to commodity returns. Gorton et al. (2013) provide both theoretical and empirical evidence that the basis predicts future returns of commodities as it is related to inventory levels and the risk premium.

Hedging pressure. Keynes' theory of normal backwardation gives the futures basis the interpretation of a risk premium. The theory states that commodity producers hedge future spot price risk by taking short positions in the futures market. As hedgers have to allow speculators to earn a premium, futures prices are set in equilibrium at a discount relative to expected future spot prices. This is consistent with a downward sloping futures curve. However, as expected future spot prices are not directly observable, the basis basis will not fully reveal the magnitude of this risk premium. Where therefore use data on the positions of hedgers in oil futures as an additional state variable. We construct the variable *hedging.pressure* as the number of short commercial positions minus the number of long commercial positions, divided by the total number of outstanding contracts. We download these time series from Bloomberg; original data source it the Commodity Futures Trading Commission (CFTC) in the Commitments of Traders Report. Variants of the hedging

[14] As for most commodities, cash prices for commodities delivered physically are difficult to obtain and noisy. In contrast to prices observed on futures markets, where specifications are standardized, cash prices depend on the quality of the product delivered, the delivery location, and even the customer-client specific relationship.

pressure variable have been used in various papers to better understand commodity futures prices (see, for example, Bakshi et al., 2019; Cortazar et al., 2018; Chari and Christiano, 2017; Szymanowska et al., 2014).[15]

Commodity momentum factor. Another variable that may capture commodity speculators' effects on return expectations is the commodity momentum factor $momC$. Following Bakshi et al. (2019), this variable is constructed as the return on a portfolio that is long in the five commodities with the highest returns over the previous six months and short in the ones with the lowest returns over the previous six months. Kang et al. (2020) find that speculators are momentum traders, as they invest more aggressively in commodities that increase in price while reducing their positions in commodities that decrease in price.

Term spread. As in the VAR for equity returns, we include the term yield spread $termsp$ in the VAR for oil returns. This variable is intended to capture the impact of the business cycle on oil returns.

Table 4 shows descriptive statistics (Panel A) and correlation matrices (Panel B) for the oil VAR state variables over the whole sample from March 1983 to December 2019, in percentage points. Table 5 shows the OLS parameter estimates for the first-order VAR model specified in Eq. (9). The R^2 of 5.1 percent is higher than the R^2 from the equity VAR. The most important variable for return prediction is the lagged oil excess return.

[15] In a recent contribution, Bianchi (2021) finds that hedging pressure is a key driver of the time-variation in commodity risk premium.

Table 4: Oil VAR State Variables

This table reports the descriptive statistics for the oil VAR state variables over the whole sample from March 1983 to December 2019, in percentage points. r_t^{oil} is the excess log return on a long position in a WTI futures contract. $basis$ is the difference between the log spot price and the log WTI futures price. $hedging.pressure$ is the difference between hedgers and speculators CFTC positions divided by open interest. $defsp$ is the difference between the log yield on Moody's BAA bonds and the log yield on Moody's AAA bonds. $momC$ is constructed following Bakshi et al. (2019).

Panel A: Descriptive Statistics.

Statistic	Mean	St. Dev.	Min	Median	Max
$basis_t$	−0.03	1.88	−10	−0.1	6
$hedging.pressure_t$	8.67	15.64	−28.28	5.46	48.41
$termsp_t$	1.78	1.09	−0.61	1.77	3.91
$momC_t$	0.49	9.07	−34.50	−0.002	37.08

Panel B: Correlation Matrix.

	r_t^{oil}	$basis_t$	$hedging.pressure_t$	$termsp_t$	$momC_t$
r_t^{oil}	1	0.361	0.084	−0.018	−0.020
$basis_t$	0.361	1	−0.026	−0.030	0.056
$hedging.pressure_t$	0.084	−0.026	1	0.044	−0.028
$termsp_t$	−0.018	−0.030	0.044	1	−0.046
$momC_t$	−0.020	0.056	−0.028	−0.046	1

Table 5: VAR Parameter Estimates for Oil

This table reports the OLS parameter estimates for a first-order VAR model including a constant, r_t^{oil}, $basis_t$, $hedging.pressure_t$, and $termsp_t$, and $momC_t$. Values in parenthesis are OLS standard errors. ***, **, and * indicates respectively 1%, 5%, and 10% level of significance. The sample period is March 1983 to December 2019.

	r_{t+1}^{oil}	$basis_{t+1}$	$hedging.pressure_{t+1}$	$termsp_{t+1}$	$momC_{t+1}$
r_t^{oil}	0.198***	0.019***	0.131***	0.003	0.010
	(0.052)	(0.007)	(0.035)	(0.002)	(0.050)
$basis_t$	0.082	0.720***	−0.389**	−0.009	0.259
	(0.265)	(0.034)	(0.180)	(0.008)	(0.256)
$hedging.pressure_t$	−0.045	−0.001	0.909***	−0.001	−0.035
	(0.030)	(0.004)	(0.020)	(0.001)	(0.029)
$termsp_t$	−0.429	−0.031	−0.031	0.966***	0.103
	(0.429)	(0.056)	(0.292)	(0.013)	(0.414)
$momC_t$	−0.057	0.012*	−0.004	−0.003	−0.066
	(0.052)	(0.007)	(0.035)	(0.002)	(0.050)
Constant	1.410	0.048	0.899	0.067**	0.616
	(0.945)	(0.123)	(0.642)	(0.028)	(0.912)
R^2	0.051	0.585	0.836	0.934	0.011

3.3 The dynamics of cash flow and discount rate news correlations

Figure 3 shows the estimates for the cash flow and discount rate news dynamics for U.S. equities and crude oil. Both time series plots show considerable variation, with large (small) movements following on average large (small) movements.[16] From this chart it is difficult to see the degree to which these fluctuations covary over time, and the dynamics of these fluctuations require the use of an appropriate methodology to take into account volatility clustering. Thus, to investigate the relation between return correlation and the correlation of the cash flow news components more formally, we consider (adjusted) conditional correlation (DCC) estimates of the equity and oil excess returns, $\widehat{\rho}_t^r$, of the equity and oil cash-flow news, $\widehat{\rho}_t^{CF}$, and of the equity and oil discount-rate news, $\widehat{\rho}_t^{DR}$. We regress $\widehat{\rho}_t^r$ on either $\widehat{\rho}_t^{CF}$, $\widehat{\rho}_t^{DR}$ or both. We are primarily interested in the R^2 from these regressions, in order to quantify to what degree the fluctuations of the equity-oil correlation can be explained by its cash flow news components. Table 6 provides the results. In column (1) we observe that the cash flow news correlation explains more than three quarters of the stock-oil correlation. Then, in column (2) we observe that $\widehat{\rho}_t^{DR}$ has low explanatory power. This finding is confirmed in column (3) where we use $\widehat{\rho}_t^{CF}$ and $\widehat{\rho}_t^{DR}$ as regressors. The results are not sensitive to estimation specifics. In the appendix we provide tables with alternative estimation of standard errors (Table A.1) and adding control variables (Table A.2).

[16] Oil cash flow news can be driven by both demand and supply shocks which are difficult to disentangle empirically. In a recent paper, Känzig (2021) provides a clean measure of oil supply shocks based on OPEC announcements. We compare our measure of cash flow news to the shock series provided via Diego Känzig's website and find a correlation coefficient of 0.48.

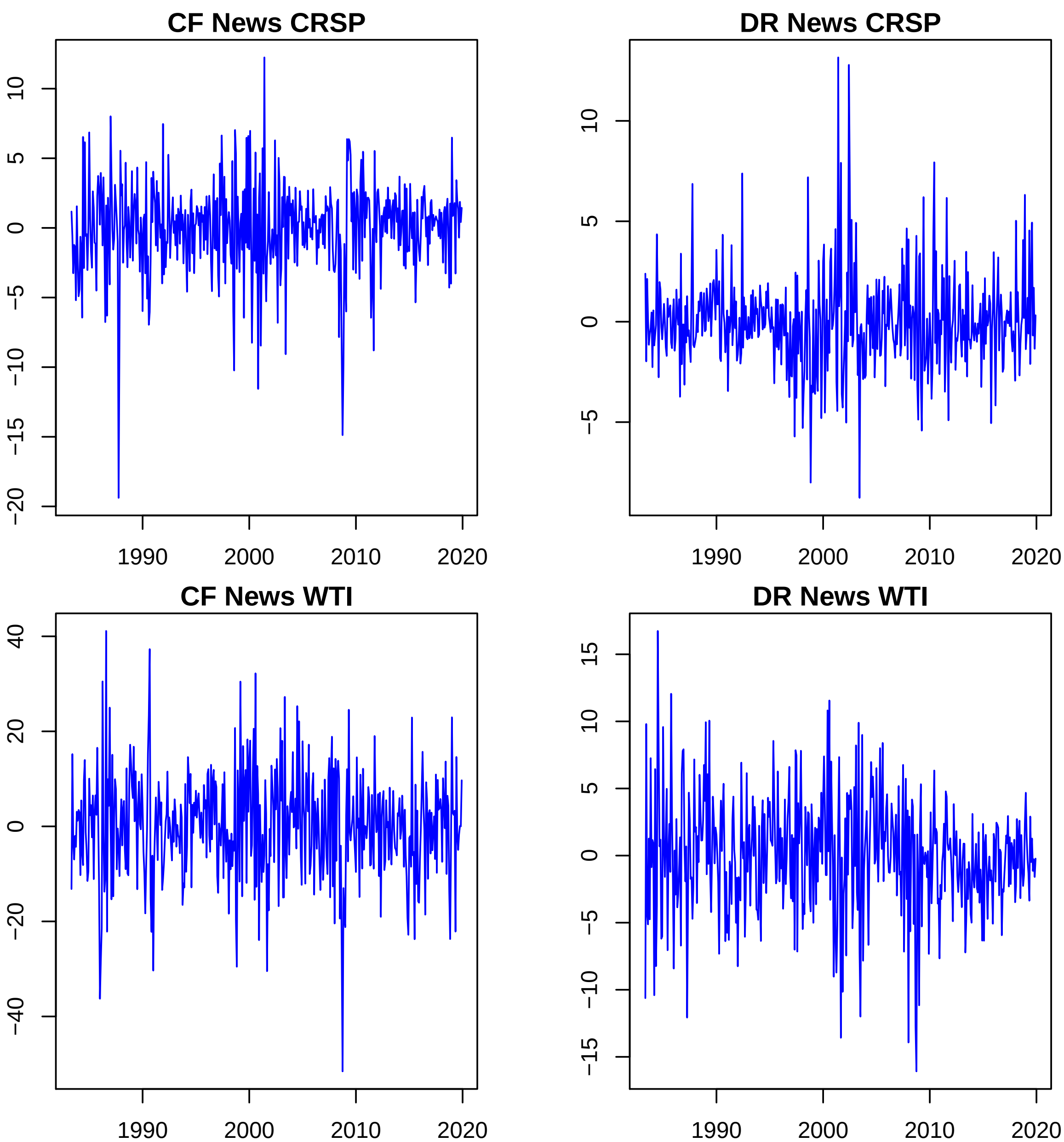

Figure 3: Cash Flow vs Discount Rate News for Stocks and Oil. This figure shows cash flow and discount rate news dynamics for equity and oil excess returns. For U.S. equity, we use the CRSP value weighted index, and for oil WTI futures contracts. The sample period is March 1983 to December 2019.

Table 6: Stock-Oil and Stock-Oil News Correlation

This table reports the estimated coefficients for the regression $\widehat{\rho}_t^r = \gamma_0 + \gamma_1 \widehat{\rho}_t^{news} + \varepsilon_t$. $\widehat{\rho}_t^r$ is the (adjusted) conditional correlation (DCC) estimate of equity and oil excess returns. For U.S. equity, we use the CRSP value weighted index, and for oil WTI futures contracts. $\widehat{\rho}_t^{news}$ is either the (adjusted) conditional correlation (DCC) estimate of cash flow news for equity and oil (column 1) or the (adjusted) conditional correlation (DCC) estimate of discount rate news for equity and oil (column 2) or both (column 3). Values in parenthesis are heteroskedasticity and autocorrelation consistent (HAC) standard errors using Newey and West (1987) with automatic bandwidth selection procedure as described in Newey and West (1994). ***, **, and * indicates respectively 1%, 5%, and 10% level of significance. The sample period is March 1983 to December 2019.

	Stock-Oil Correlation		
	(1)	(2)	(3)
Cash Flow News Stock-Oil Correlation	2.865***		2.864***
	(0.289)		(0.286)
Discount Rate News Stock-Oil Correlation		0.102	0.039
		(0.635)	(0.191)
Constant	−0.196***	0.120	−0.196***
	(0.038)	(0.096)	(0.038)
Observations	440	440	440
Adjusted R^2	0.788	−0.001	0.787

3.4 A possible explanation

While the stock-oil correlation fluctuates considerably over time, there is a striking jump around 2008, as can be seen from Figure 1. This correlation was negative most of the time before that jump, but has stayed in positive territory since then. Given our evidence that the stock-oil correlation is primarily driven by cash flow news, a positive correlation essentially means that positive cash flow shocks on oil returns tend to be good news for the stock market. This is at first surprising since oil is an input to most other companies. For an economy that imports considerable amounts of its oil consumption, spillover effects from oil producing firms to the rest of the economy are likely dominated by the negative effects from price increases. The key observation is that the U.S. has over time turned from a net importer of crude oil to a net exporter. While the U.S. has only recently become a net exporter of oil, financial markets tend to anticipate developments. This is naturally the case for oil production, where new technologies (shale oil production) and investment patterns made this change foreseeable. Figure 4 shows that the turning point in U.S. oil production occurred at around the same time as the regime shift in the stock-oil correlation pattern. In Appendix B, Figure B.1 provides further evidence on the changing pattern in U.S. oil exports and imports. We test the empirical relationship between the stock-oil correlation and U.S. oil production in Table 7. An increase in U.S. oil production is positively and significantly related to both the stock-oil correlation and the stock-oil cash flow news correlation, but not with the comovement of discount rate news.[17]

[17] Results remain qualitatively the same if we use other proxies for increased U.S. oil production or if we add speculative pressure as a control variable to the regression.

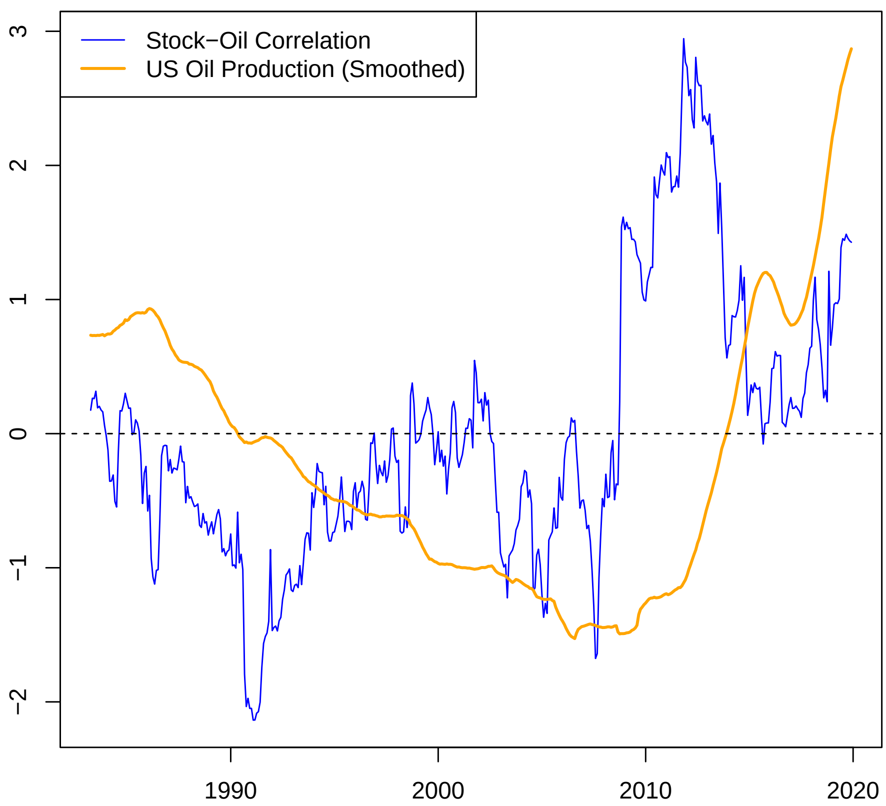

Figure 4: Stock-Oil Correlation and U.S. Oil Production. This figure shows the (standardized) time-varying correlation between equity and oil excess returns and (yearly smoothed) U.S. oil production for the period is 1983 to 2019. For U.S. equity, we use the CRSP value weighted index, and for oil WTI futures contracts.

Table 7: Stock-Oil and Stock-Oil News Correlation: Oil Production

This table reports the estimated coefficients for the regression $\widehat{\rho}_t = \pi_0 + \pi_1 oil\ prod_t + \beta' \mathbf{x}_t + \varepsilon_t$, where $\widehat{\rho}_t$ is the (adjusted) conditional correlation estimate of equity and oil excess returns $\widehat{\rho}_t^r$ (column 1), the (adjusted) conditional correlation (DCC) estimate of cash flow news for equity and oil $\widehat{\rho}_t^{CF}$ (column 2), or the (adjusted) conditional correlation (DCC) estimate of discount rate news for equity and oil $\widehat{\rho}_t^{DR}$ (column 3); $oil\ prod_t$ is the log change of crude oil barrels produced in the US (millions/day); $\mathbf{x}_t$ comprises the variables used in VAR specifications for both equity and oil. For U.S. equity, we use the CRSP value weighted index, and for oil WTI futures contracts. Values in parenthesis are heteroskedasticity and autocorrelation consistent (HAC) standard errors using Newey and West (1987) with automatic bandwidth selection procedure as described in Newey and West (1994). ***, **, and * indicates respectively 1%, 5%, and 10% level of significance. The sample period is March 1983 to December 2019.

	$\widehat{\rho}_t$		
	(1)	(2)	(3)
$oil\ prod_t$	1.511**	0.375**	−0.015
	(0.619)	(0.173)	(0.122)
Constant	−0.621	−0.092	−0.241**
	(0.417)	(0.125)	(0.114)
Observations	441	440	440
Adjusted R^2	0.338	0.231	0.384

Our analysis shows a changing pattern of the cash flow news correlation with the transition of the U.S. from a net importer to a net exporter of crude oil. While we do not observe multiple changing trends in U.S. oil production in our sample period, we resort to other countries for additional plausibility. Among traditional oil exporting countries, we analyze Canada and Russia. Since significant production from oil sands started in the 1980's, total Canadian oil production has constantly increased. Thus, cash flow news can be expected to be positively correlated, with increasing magnitude over time. The regressions shown in Table 8, Panel A, confirm that the stock-oil correlation is mainly driven by cash flow news correlation for Canada and Russia.

The cash flow news correlation explains more than 45% of the overall stock-oil correlation in the case of Canada. While the R^2 is lower for Russia, the relationship is again highly significant. Other extremes are Hong Kong and Japan, which we use to illustrate oil-importing economies. Positive oil cash flow news that are demand shocks can be positive for the economy and thus equity markets, while cash flow news on the oil price that are driven by a negative supply shock are clearly a cost to the economy. Thus, we expect that the cash flow news correlation does not dominate the total stock-oil correlation. Indeed Table 8, Panel B, shows that the pattern observed for Canada and Russia is nonexistent for Japan and Hong Kong, both oil importers.

Table 8: International Stock-Oil Correlation and CF News

This table reports the estimated coefficients for the regression $\widehat{\rho}_t^r = \gamma_0 + \gamma_1 \widehat{\rho}_t^{CF} + \varepsilon_t$, where $\widehat{\rho}_t^r$ is the (adjusted) conditional correlation estimate of equity and oil excess returns and $\widehat{\rho}_t^{CF}$ is the (adjusted) conditional correlation (DCC) estimate of cash flow news for equity and oil. Panel A reports results for oil exporting countries Canada (CAN, column 1) and Russia (RUS, column 2). Panel B reports results for oil importing countries Japan (JAP, column 1) and Hong Kong (HK, column 2). News are obtained using in the VAR return decomposition specification the first four macro-factors constructed in Ludvigson and Ng (2009) as state variables. Constant term estimates are not tabulated. Values in parenthesis are heteroskedasticity and autocorrelation consistent (HAC) standard errors using Newey and West (1987) with automatic bandwidth selection procedure as described in Newey and West (1994). ***, **, and * indicates respectively 1%, 5%, and 10% level of significance. The sample period is March 1983 to December 2019.

Panel A: Oil Exporting Countries.

	CAN	RUS
	(1)	(2)
CF News CAN Stock-Oil Correlation	1.053***	
	(0.278)	
CF News RUS Stock-Oil Correlation		1.898***
		(0.268)
Observations	408	290
Adjusted R^2	0.460	0.116

Panel B: Oil Importing Countries.

	JAP	HK
	(1)	(2)
CF News JAP Stock-Oil Correlation	−0.004	
	(0.124)	
CF News HK Stock-Oil Correlation		0.556
		(0.368)
Observations	408	408
Adjusted R^2	0.000	0.010

3.5 Robustness

Principal components of macro variables. Campbell and Vuolteenaho (2004) and Campbell et al. (2018) directly estimate discount rate news and back out cash flow news as the residual. As pointed out by Chen and Zhao (2009), this approach may suffer from misspecification, as discount rate news are difficult to measure. Consequently cash flow news inherit the potential misspecification error of discount rate news. A remedy suggested by Chen and Zhao (2009) to deal with model misspecification is to use few principal components extracted from a large set of variables that can predict asset returns.

To address this issue, we use the first four macro-factors extracted from a large cross-section of macro variables; factors are constructed as in Ludvigson and Ng (2009).[18] We include four factors in the VAR specification to keep the parsimonious specification with five state variables proposed by Campbell et al. (2018). The portion of variation explained by using the four macro-factors is practically identical to the variation explained by our previous specification. We report the parameter estimates for this VAR return decomposition specification in appendix C, Table C.1. We then compute conditional correlations between the cash flow and discount rate news of U.S. equity and oil excess returns, respectively, using a DCC-GARCH(1,1) specification. Table C.2 in the appendix reports the estimated coefficients for the regression $\widehat{\rho}_t = \gamma_0 + \gamma_1 \widehat{\rho}_t^{news} + \varepsilon_t$, where $\widehat{\rho}_t^{news}$ is either $\widehat{\rho}_t^{CF}$ or $\widehat{\rho}_t^{DR}$. Importantly, results replicate the main findings coming from the previous specification. In particular, cash flow news correlation explains about 84% of variation in stock-oil correlation, and positively predicts stock-oil correlation up to 12 months with high statistical significance.

[18] Macro-factors are available at Sydney Ludvigson's webpage.

Higher–Order VARs. To decompose unexpected excess returns for stock and oil in cash flow and discount rate news, we follow Campbell and Vuolteenaho (2004) and Campbell et al. (2018) and specify a VAR of order 1. However, the optimal number of lags used in the VAR specification can also be assessed by formal econometric tests. Specifically, we consider the following lag-order selection criteria: Akaike (AIC), Hannan-Quinn (HQ), Schwarz (SC), and the forecast prediction error (FPE). Appendix Table D.1 reports results for the different methods. The different selection criteria deliver uniform results for stocks, with an optimal lag-order of 2; in the case of oil, AIC and FPE select 2 lags, while HQ and SC select 1 lag. Thus, as an additional robustness check, we investigate whether our empirical findings are robust to a VAR specified with 2 lags.

Using a VAR(2) specification, the R2 for r_t^{eq} goes from 2.4% (in the case of the VAR(1)) to 3.1%, and for r_t^{oil} from 5.1% to 7.3%. Appendix Table D.2 reports results from regressing the stock-oil time-varying correlation on the two news components from cash flows and discount rates, obtained using a VAR of order 2. The results largely confirm our main findings that the stock-oil correlation is mainly explained by the time variation in the cash flow news of the two assets.

Speculative Pressure. Given the increased financialization of commodity markets, price pressure arising from speculative activities potentially affects the stock-oil correlation. In Table 9, we test whether speculative activity is related to the stock-oil correlation. As the measure for speculative activity, we use the difference between hedgers and speculators CFTC positions divided by open interest. Positions of speculators can be related to cash flow news and discount rate news. While column (1) in Table 9 provides evidence that speculative pressure is related to the overall stock-oil correlation, the variable is driven out once we control for the cash-flow news correlation.

Table 9: Stock-Oil and Stock-Oil News Correlation: Financialization?

This table reports the estimated coefficients for the regression $\widehat{\rho}_t^r = \gamma_0 + \gamma_1 \widehat{\rho}_t^{news} + \gamma_2 spec_t + \beta' \mathbf{x}_t + \varepsilon_t$. $\widehat{\rho}_t^r$ is the (adjusted) conditional correlation estimate (DCC) of equity and oil excess returns; $\mathbf{x}_t$ are the state variables used in VAR specifications for both equity and oil. $\widehat{\rho}_t^{news}$ is either $\widehat{\rho}_t^{CF}$ (column (2)) or $\widehat{\rho}_t^{DR}$ (column (3)) or both (column (4)). For U.S. equity, we use the CRSP value weighted index, and for oil WTI futures contracts. Values in parenthesis are heteroskedasticity and autocorrelation consistent (HAC) standard errors using Newey and West (1987) with automatic bandwidth selection procedure as described in Newey and West (1994). ***, **, and * indicates respectively 1%, 5%, and 10% level of significance. The sample period is March 1983 to December 2019.

	Stock-Oil Correlation			
	(1)	(2)	(3)	(4)
Cash Flow News Stock-Oil Correlation		2.527^{***}		2.489^{***}
		(0.336)		(0.387)
Discount Rate News Stock-Oil Correlation			-0.696	-0.337
			(0.556)	(0.351)
spec	1.000^{**}	0.212	1.111^{***}	0.275
	(0.458)	(0.228)	(0.406)	(0.235)
Constant	-0.733^{**}	-0.416^{***}	-0.915^{***}	-0.507^{***}
	(0.292)	(0.151)	(0.268)	(0.183)
Observations	441	440	440	440
Adjusted R^2	0.421	0.837	0.444	0.842

4 Conclusions

In this paper, we study the time-varying correlation between stocks and oil. We find that more than three quarters of total variation in unexpected returns comovement between the two assets is explained by the correlation in cash flow news. Thus, we provide new evidence to address what we denote as "Bernanke's puzzle", that is why the stock-oil correlation in recent years has been high and positive in the U.S.. Our findings suggest that changes in the structure of the real economy, specifically the increasing oil production of the U.S. over the past years, appear to be key drivers of such dynamics of the comovement.

References

Andrews, D. (1991). Heteroskedasticity and autocorrelation consistent covariant matrix estimation. *Econometrica 59*(3), 817–858.

Bakshi, G., X. Gao, and A. G. Rossi (2019). Understanding the sources of risk underlying the cross section of commodity returns. *Management Science 65*(2), 619–641.

Balduzzi, P. and E. Chiang (2020). Real exchange rates and foreign currency risk premia. *Review of Asset Pricing Studies 10*, 94–121.

Basak, S. and A. Pavlova (2016). A model of financialization of commodities. *The Journal of Finance 71*(4), 1511–1556.

Bernanke, B. (2016). The relationship between stocks and oil prices. Ben Bernanke's Blog on Brookings posted on February 19, 2016.

Bianchi, D. (2021). Adaptive expectations and commodity risk premiums. *Journal of Economic Dynamics and Control*, 104078.

Boons, M. and M. Szymanowska (2014). The price of commodity risk in stock and futures markets. Working paper.

Brennan, M. J. (1958). The supply of storage. *The American Economic Review 48*(1), 50–72.

Brogaard, J., M. C. Ringgenberg, and D. Sovich (2019). The economic impact of index investing. *The Review of Financial Studies 32*(9), 3461–3499.

Campbell, J. Y. (1991). A variance decomposition for stock returns. *The Economic Journal 101*(405), 157–179.

Campbell, J. Y. and J. Ammer (1993). What moves the stock and bond markets? A variance decomposition for long-term asset returns. *The Journal of Finance 48*, 3–37.

Campbell, J. Y., S. Giglio, C. Polk, and R. Turley (2018). An intertemporal capm with stochastic volatility. *Journal of Financial Economics 128*(2), 207–233.

Campbell, J. Y. and R. J. Shiller (1988). The dividend-price ratio and expectations of future dividends and discount factors. *The Review of Financial Studies 1*(3), 195–228.

Campbell, J. Y. and T. Vuolteenaho (2004). Bad beta, good beta. *American Economic Review 94*(5), 1249–1275.

Cassacus, J. and P. Collin-Dufresne (2005). Stochastic convenience yield implied from commodity futures and interest rates. *The Journal of Finance 65*(5), 2283–2331.

Chari, V. V. and L. Christiano (2017). Financialization in commodity markets. Working Paper.

Chen, L. and X. Zhao (2009). Return decomposition. *The Review of Financial Studies 22*(12), 5213–5249.

Cheng, I.-H. and W. Xiong (2014). Financialization of commodity markets. *Annual Review of Financial Economics 6*, 419–441.

Chiang, E. and X. N. Mo (2019). A 'bad beta, good beta' anatomy of currency risk premiums and trading strategies. Working paper.

Christoffersen, P. and X. N. Pan (2018). Oil volatility risk and expected stock returns. *Journal of Banking & Finance 95*, 5–26.

Cortazar, G., P. Liedtke, H. Ortega, and E. S. Schwartz (2018). Time-varying term structure of oil risk premiums. Working Paper.

David, A. (2019). Exploration activity, long-run decisions, and the risk premium in energy futures. *The Review of Financial Studies 32*(4), 1536–1572.

Davis, J. L., E. F. Fama, and K. R. French (2000). Characteristics, covariances, and average returns: 1929 to 1997. *The Journal of Finance 55*(1), 389–406.

Deaton, A. and G. Laroque (1992). On the behaviour of commodity prices. *Review of Economic Studies 59*(1), 1–23.

Draghi, M. (2017). Monetary policy and the economic recovery in the euro area. Speech given at ECB and its Watchers XVIII Conference, Frankfurt, April 6, 2017.

Engle, R. (2002). Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models. *Journal of Business & Economic Statistics 20*(3), 339–350.

Fama, E. F. and K. R. French (1988). Business cycles and the behavior of metals prices. *The Journal of Finance 43*(5), 1075–1093.

Fernandez-Perez, A., A.-M. Fuertes, and J. Miffre (2017). Commodity markets, long-run predictability, and intertemporal pricing. *Review of Finance 21*(3), 1159–1188.

Gibson, R. and E. S. Schwartz (1990). Stochastic convenience yield and the pricing of oil contingent claims. *The Journal of Finance 45*(3), 959–976.

Gilje, E., R. Ready, and N. Roussanov (2016). Fracking, drilling, and asset pricing: Estimating the economic benefits of the shale revolution. Technical report, National Bureau of Economic Research.

Goldstein, I. and L. Yang (2019). Commodity financialization and information transmission. Working paper.

Gorton, G. and K. G. Rouwenhorst (2006). Facts and fantasies about commodity futures. *Financial Analysts Journal 62*(2), 47–68.

Gorton, G. B., F. Hayashi, and K. G. Rouwenhorst (2013). The fundamentals of commodity futures returns. *Review of Finance 17*(1), 35–105.

Hamilton, J. D. and J. C. Wu (2014). Risk premia in crude oil futures prices. *Journal of International Money and Finance 42*, 9–37.

Hitzemann, S. (2016). Macroeconomic fluctuations, oil supply shocks, and equilibrium oil futures prices. Working Paper.

Kang, W., K. G. Rouwenhorst, and K. Tang (2020). A tale of two premiums: The role of hedgers and speculators in commodity futures markets. *The Journal of Finance 75*(1), 377–417.

Känzig, D. R. (2021). The macroeconomic effects of oil supply news: Evidence from opec announcements. *American Economic Review 111*(4), 1092–1125.

Kilian, L. and C. Park (2009). The impact of oil price shocks on the us stock market. *International Economic Review 50*(4), 1267–1287.

Ludvigson, S. C. and S. Ng (2009). Macro factors in bond risk premia. *The Review of Financial Studies 22*(12), 5027–5067.

Mohaddes, K. and M. H. Pesaran (2017). Oil prices and the global economy: Is it different this time around? *Energy Economics 65*, 315–325.

Newey, W. K. and K. D. West (1987). A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica 55*(3), 703–708.

Newey, W. K. and K. D. West (1994). Automatic lag selection in covariance matrix estimation. *The Review of Economic Studies 61*(4), 631–653.

Obstfeld, M., R. Arezki, and G. M. Milesi-Ferretti (2016). Oil prices and the global economy: It's complicated. VOXEU / CEPR. April 13.

Obstfeld, M., G. M. Milesi-Ferretti, and R. Arezki (2016). Oil prices and the global economy: It's complicated. IMF Blog. March 24.

Ready, R. C. (2018a). Oil consumption, economic growth, and oil futures: The impact of long-run oil supply uncertainty on asset prices. *Journal of Monetary Economics 94*, 1–26.

Ready, R. C. (2018b). Oil prices and the stock market. *The Review of Finance 22*(1), 155–176.

Routledge, B. R., D. J. Seppi, and C. S. Spatt (2000). Equilibrium forward curves for commodities. *Journal of Finance*, 1297–1338.

Shiller, R. J. (2000). *Irrational exuberance*. Princeton: Princeton University Press.

Singleton, K. J. (2014). Investor flows and the 2008 boom/bust in oil prices. *Management Science 60*(2), 300–318.

Szymanowska, M., F. De Roon, T. Nijman, and R. Van Den Goorbergh (2014). An anatomy of commodity futures risk premia. *The Journal of Finance 69*(1), 453–482.

Appendix

A Additional Results

Table A.1: Stock-Oil and Stock-Oil News Correlation

This table reports the estimated coefficients for the regression $\widehat{\rho}_t^r = \gamma_0 + \gamma_1 \widehat{\rho}_t^{news} + \varepsilon_t$. $\widehat{\rho}_t^r$ is the (adjusted) conditional correlation (DCC) estimate of equity and oil excess returns. For U.S. equity, we use the CRSP value weighted index, and for oil WTI futures contracts. $\widehat{\rho}_t^{news}$ is either the (adjusted) conditional correlation (DCC) estimate of cash flow news for equity and oil (column 1) or the (adjusted) conditional correlation (DCC) estimate of discount rate news for equity and oil (column 2) or both (column 3). Values in parenthesis are heteroskedasticity and autocorrelation consistent (HAC) standard errors using Newey and West (1987) with optimal truncation lag as suggested by Andrews (1991). ***, **, and * indicates respectively 1%, 5%, and 10% level of significance. The sample period is March 1983 to December 2019.

	Stock-Oil Correlation		
	(1)	(2)	(3)
Cash Flow News Stock-Oil Correlation	2.865***		2.864***
	(0.331)		(0.320)
Discount Rate News Stock-Oil Correlation		0.102	0.039
		(0.942)	(0.206)
Constant	−0.196***	0.120	−0.196***
	(0.049)	(0.166)	(0.051)
Observations	440	440	440
Adjusted R^2	0.788	−0.001	0.787

Table A.2: Stock-Oil and Stock-Oil News Correlation

This table reports the estimated coefficients for the regression $\widehat{\rho}_t^{r} = \gamma_0 + \gamma_1 \widehat{\rho}_t^{news} + \beta' \mathbf{x}_t + \varepsilon_t$. $\widehat{\rho}_t^{r}$ is the (adjusted) conditional correlation (DCC) of equity and oil excess returns; $\mathbf{x}_t$ are the state variables used in VAR specifications for both equity and oil. For U.S. equity, we use the CRSP value weighted index, and for oil WTI futures contracts. $\widehat{\rho}_t^{news}$ is either the (adjusted) conditional correlation (DCC) estimate of cash flow news for equity and oil (column 1) or the (adjusted) conditional correlation (DCC) estimate of discount rate news for equity and oil (column 2) or both (column 3). Values in parenthesis are heteroskedasticity and autocorrelation consistent (HAC) standard errors using Newey and West (1987) with automatic bandwidth selection procedure as described in Newey and West (1994). ***, **, and * indicates respectively 1%, 5%, and 10% level of significance. The sample period is March 1983 to December 2019.

	Stock-Oil Correlation		
	(1)	(2)	(3)
Cash Flow News Stock-Oil Correlation	2.612***		2.602***
	(0.296)		(0.317)
Discount Rate News Stock-Oil Correlation		−0.391	−0.256
		(0.610)	(0.310)
Constant	−0.388***	−0.735**	−0.450***
	(0.144)	(0.351)	(0.163)
Observations	440	440	440
Adjusted R^2	0.834	0.333	0.837

B U.S. Oil Trade (EIA)

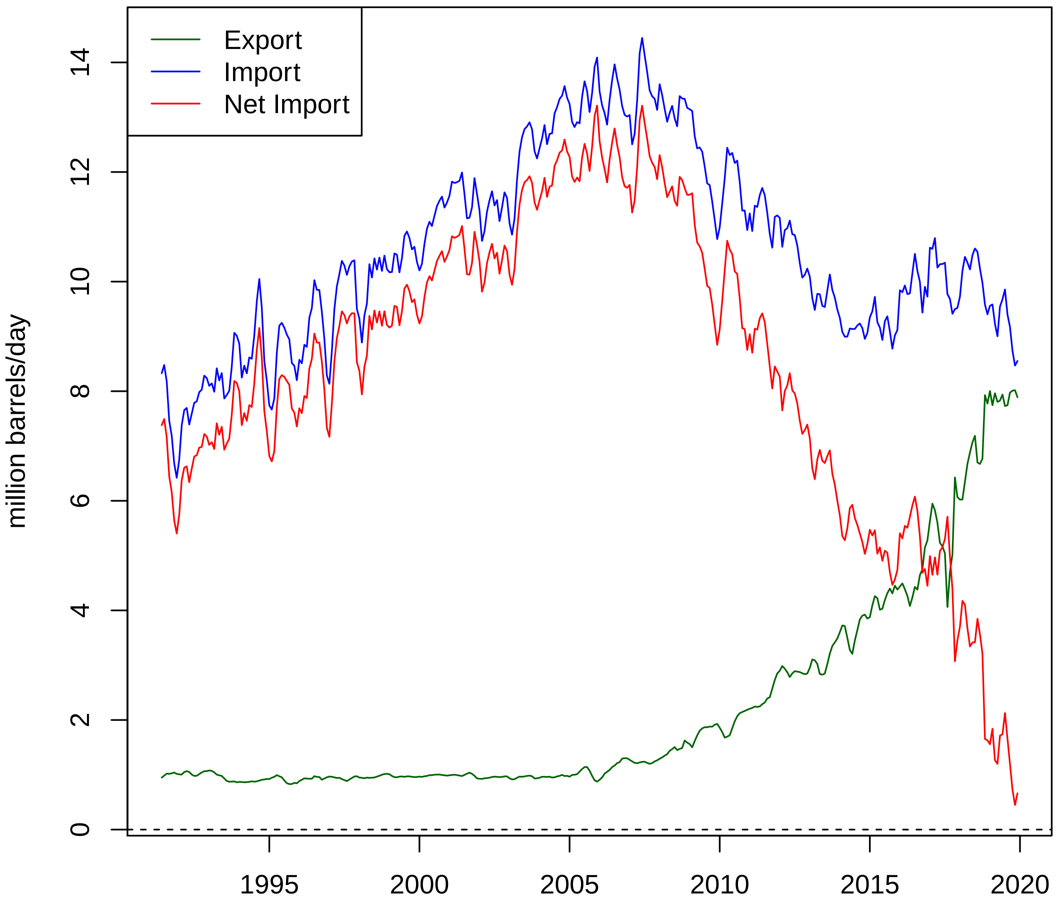

Figure B.1: U.S. Oil Import-Export. This figure shows U.S. oil total export, import, and net import for the period 1991 to 2020. Monthly observations, quarterly MA smoothed.

C VAR Return Decomposition with Macro-Factors

Table C.1: VAR Parameter Estimates for Stock and Oil using Macro-Factors

This table reports the OLS parameter estimates for a first-order VAR model including a constant, the first four macro-factors constructed as in Ludvigson and Ng (2009), plus *rme* and *roe* repsectively in Column (1) and (2). Constant term etimates are not tabulated. Values in parenthesis are OLS standard errors. ***, **, and * indicates respectively 1%, 5%, and 10% level of significance. The sample period is March 1983 to December 2019.

	r^{eq}_{t+1} (1)	r^{oil}_{t+1} (2)
r^{eq}_t	0.047	
	(0.052)	
r^{oil}_t		0.197^{***}
		(0.051)
$F1_t$	-1.011	-4.197^{***}
	(0.712)	(1.509)
$F2_t$	2.163^{*}	3.735
	(1.136)	(2.396)
$F3_t$	0.815	1.667
	(0.993)	(2.036)
$F4_t$	1.483	4.231
	(1.358)	(2.752)
R^2	0.018	0.069

Table C.2: Stock-Oil and Stock-Oil News PCA Correlation

This table reports the estimated coefficients for the regression $\widehat{\rho}_t^r = \gamma_0 + \gamma_1 \widehat{\rho}_t^{news} + \varepsilon_t$. For U.S. equity, we use the CRSP value weighted index, and for oil WTI futures contracts. $\widehat{\rho}_t^{news}$ is the the (adjusted) conditional correlation (DCC) of cash flow news for equity and oil (column 1) or the (adjusted) conditional correlation (DCC) of discount rate news for equity and oil (column 2) or both (column 3), obtained using in the VAR return decomposition specification and the first four macro-factors constructed in Ludvigson and Ng (2009) as state variables. Costant term estimates are not tabulated. Values in parenthesis are heteroskedasticity and autocorrelation consistent (HAC) standard errors using Newey and West (1987) with automatic bandwidth selection procedure as described in Newey and West (1994). ***, **, and * indicates respectively 1%, 5%, and 10% level of significance. The sample period is March 1983 to December 2019.

	Stock-Oil Correlation		
	(1)	(2)	(3)
CF News Stock-Oil Factors Correlation	2.335***		2.590***
	(0.240)		(0.227)
DR News Stock-Oil Factors Correlation		−0.861	0.659***
		(0.625)	(0.117)
Constant	−0.119***	−0.177	0.082
	(0.022)	(0.173)	(0.052)
Observations	440	440	440
Adjusted R^2	0.839	0.073	0.873

D Higher–Order VAR Return Decomposition

Table D.1: Lag-Order Selection Criteria

This table reports the optimal number of lags in the VAR specification for stock and oil for different selection criteria. The lag-order selection criteria considered are: Akaike (AIC), Hannan-Quinn (HQ), Schwarz (SC), and the forecast prediction error (FPE). The sample period is March 1983 to December 2019.

	AIC	HQ	SC	FPE
stock	2	2	2	2
oil	2	1	1	2

Table D.2: Stock-Oil and Stock-Oil News Correlation using a VAR(2)

This table reports the estimated coefficients for the regression $\widehat{\rho}_t^r = \gamma_0 + \gamma_1 \widehat{\rho}_t^{news} + \varepsilon_t$. $\widehat{\rho}_t^r$ is the (adjusted) conditional correlation (DCC) estimate of equity and oil excess returns. For U.S. equity, we use the CRSP value weighted index, and for oil WTI futures contracts. $\widehat{\rho}_t^{news}$ is the the (adjusted) conditional correlation (DCC) of cash flow news for equity and oil (column 1) or the (adjusted) conditional correlation (DCC) of discount rate news for equity and oil (column 2) or both (column 3), obtained using a VAR(2) specification for the return decomposition. Costant term estimates are not tabulated. Values in parenthesis are heteroskedasticity and autocorrelation consistent (HAC) standard errors using Newey and West (1987) with automatic bandwidth selection procedure as described in Newey and West (1994). ***, **, and * indicates respectively 1%, 5%, and 10% level of significance. The sample period is March 1983 to December 2019.

	Stock-Oil Correlation		
	(1)	(2)	(3)
CF News Stock-Oil Correlation	2.917***		2.924***
	(0.190)		(0.202)
DR News Stock-Oil Correlation		−0.462	0.060
		(0.787)	(0.194)
Constant	−0.287***	0.142	−0.291***
	(0.037)	(0.115)	(0.045)
Observations	439	439	439
Adjusted R^2	0.839	0.012	0.839