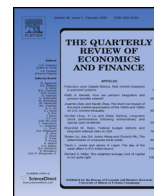




Contents lists available at ScienceDirect

The Quarterly Review of Economics and Finance

journal homepage: www.elsevier.com/locate/qref



Oil shocks and stock return volatility[☆]

Lance J. Bachmeier, Soheil R. Nadimi*

Department of Economics, Kansas State University, 327 Waters Hall, Manhattan, KS 66506, USA

ARTICLE INFO

Article history:

Received 30 October 2017
Received in revised form 9 January 2018
Accepted 14 January 2018
Available online xxx

JEL classification:

G17
Q41

Keywords:

Oil price
Stock return
Volatility
Prediction

ABSTRACT

Asset return volatility is important to the macroeconomy. This paper asks whether oil price volatility can be used as a predictor of stock return volatility. In contrast with previous research, we focus on the out-of-sample predictive power of oil price volatility rather than on in-sample inference. Formal tests of out-of-sample predictive ability find no evidence supporting the use of oil price volatility as a predictor of future stock return volatility. Further analysis using rolling window estimation and structural break tests shows that the coefficients of this relationship are very unstable. The coefficients can be positive, negative, or close to zero depending on the sample that is chosen. We discuss the implications of this finding for monetary policy.

© 2018 Board of Trustees of the University of Illinois. Published by Elsevier Inc. All rights reserved.

1. Introduction

The volatility of asset prices is believed by many to have important effects on the macroeconomy (see e.g. Phelps, 1999). This suggests that monetary and fiscal policy should be made taking into account the volatility of asset prices, and in particular, the volatility of stock prices. Farmer (2012) has advocated a policy of direct government intervention to reduce the volatility of the stock prices. If these views are correct, and the government should be offsetting or even preventing volatility of stock prices, it is important to find good predictors of stock price volatility. An obvious candidate is oil price volatility. There are many published estimates of the effect of oil shocks on macroeconomic variables.¹ A growing literature has found evidence that oil price shocks have an effect on stock prices,²

with most authors finding that higher oil prices have a negative effect on stock returns.

A natural question is whether oil price volatility is a useful predictor of stock market volatility. Several papers have considered this question and concluded that oil price volatility can be used to improve upon forecasts of stock return volatility. Elyasiani, Mansur, and Odusami (2011) estimated GARCH(1,1) models of industry stock returns that allowed the variance of the error term to depend on the previous day's oil price volatility. For the period from December 1998 to December 2006, they were able to reject the null hypothesis of a zero coefficient in the variance equation for five of thirteen industries. Sadorsky (1999) reported impulse response functions and forecast error variance decompositions for real stock returns following shocks to the price of oil and oil price volatility. Papers with a more specialized focus include Sadorsky (2003), which investigated the effect of oil price volatility on the volatility of technology stocks, and Hammoudeh, Diboglu, and Aleisa (2004), which estimated the effect of oil price volatility on the volatility of oil industry stock prices. The conclusion of all of these papers is that there is a useful forecasting relationship between lagged oil price volatility and stock return volatility.

This paper differs from the others by focusing on the out-of-sample forecast power of oil price volatility.³ As emphasized by Clark and McCracken (2013), "Forecasts need to be good to be

[☆] This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

* Corresponding author.

E-mail addresses: lanceb@ksu.edu (L.J. Bachmeier), soheil@ksu.edu (S.R. Nadimi).

¹ Some recent papers include Atems, Kapper, and Lam (2015), Edelstein and Kilian (2009), Hamilton (2011), Herrera and Pesavento (2009), Herrera, Lagalo, and Wada (2011), Kilian (2009), Kilian and Lewis (2010), Kilian and Vigfusson (2011), and Melichar (2016).

² See e.g. Alsalmán and Herrera (2015), Apergis and Miller (2009), Basher, Haug, and Sadorsky (2012), Chen (2010), Cunado and De Gracia (2014), Jones and Kaul (1996), and Kilian and Park (2009).

³ It is important to stress that the goal of this paper is not to estimate a model of stock return volatility. That has been done in many previous papers, and it would

useful for decision making. Determining if forecasts are good involves formal evaluation of the forecasts.” One reason in particular that a correlation identified in the full sample might not translate into good forecasts is parameter instability (Pettenuzzo & Timmerman, 2011). We build on the work done in the papers cited above by evaluating the out-of-sample forecast accuracy of stock return volatility models with and without oil price volatility. We investigate the stability of the parameters of the relationship through time. Full-sample Granger causality test results, along with the other in-sample evaluation techniques applied in the previous literature, can be misleading in the presence of parameter instability, and we find that to be the case.

The most important result to emerge from our analysis is that the relationship between oil price volatility and stock return volatility is unstable. Rolling window regression estimates show that the coefficients vary substantially over time. The variation in the parameter estimates is so substantial that it is possible to find any desired correlation between the variables – positive, negative, or zero – simply by choosing an appropriate subsample of the data. Structural break tests reject the null hypothesis of parameter stability for the S&P 500, the CRSP value-weighted index, and industry-level returns for 49 sectors that cover nearly all of the economy. Formal tests of out-of-sample predictive ability that exclude the 2008–2009 financial crisis period find no support for the use of oil price volatility as a predictor of stock return volatility. On the basis of our findings of parameter instability and the failure of models with oil price volatility to consistently improve out-of-sample forecasts of stock return volatility in the past, and in contrast to the existing literature, we conclude that there is no basis for using oil price volatility as a predictor of stock return volatility.

2. Data

Daily data on West Texas Intermediate (WTI) spot prices were downloaded from the Federal Reserve Economic Database (FRED) provided by the Federal Reserve Bank of St. Louis. We use two stock indexes. Data on the S&P 500 closing price were downloaded from Yahoo! Finance. The CRSP value-weighted index and industry-level value-weighted returns for 49 sectors were downloaded from the website of professor Kenneth French.⁴ In Table 1 are the complete names of all industry sectors and their shortened names that are used in the text. All data cover the period January 2, 1986 (the earliest available date for daily oil prices) to April 30, 2015. We use the natural log return of all variables.

Volatility of the oil price and stock return data are measured as the realized volatility of those series. The realized volatility of each series was calculated as the sample standard deviation for each month. Fig. 1 plots the realized volatility series of WTI price change as well as the S&P 500 and the CRSP returns for the period January 1986 to April 2015. Realized volatility has been used as a measure of volatility in the existing literature (see e.g. Andersen, Bollerslev, Diebold, & Labys, 2003; Schwert, 1989).

One might question the decision to use realized volatility measures rather than the popular GARCH family of volatility models. There is no obvious reason to prefer a GARCH model. The advantage of using a realized volatility measure is that it is consistent with the real-time nature of an actual forecasting exercise. That can be done with GARCH models, but only if one sacrifices efficiency, and it is unclear what would be gained from doing so. Second, even if

be straightforward to do so using a GARCH model or one of its many variants, but that would not by itself provide any information about out-of-sample stock return volatility prediction. Hypothesis testing and characterizing the dynamics of the process are important but distinct from forecast evaluation.

⁴ <http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data.library.html>.

Table 1
 List of industry sectors.

Name used in the text	Complete name
Agriculture	Agriculture
Food Prod	Food Products
Candy Soda	Candy & Soda
Beer	Beer & Liquor
Tobacco	Tobacco Products
Recreation	Recreation
Entertain	Entertainment
Printing	Printing and Publishing
Cons Goods	Consumer Goods
Apparel	Apparel
Healthcare	Healthcare
Med Equip	Medical Equipment
Pharma Prod	Pharmaceutical Products
Chemicals	Chemicals
Rubber Plas	Rubber and Plastic Products
Textiles	Textiles
Constr Mat	Construction Materials
Construct	Construction
Steel Works	Steel Works Etc
Fabric Prod	Fabricated Products
Machinery	Machinery
Electric Equip	Electrical Equipment
Autos	Automobiles and Trucks
Aircraft	Aircraft
Shipbuild	Shipbuilding, Railroad Equipment
Defense	Defense
Prec Metals	Precious Metals
Mining	Non-Metallic and Industrial Metal Mining
Coal	Coal
Petroleum	Petroleum and Natural Gas
Utilities	Utilities
Communic	Communication
Pers Serv	Personal Services
Bus Serv	Business Services
Computers	Computers
Comp Soft	Computer Software
Electro Equip	Electronic Equipment
Meas Control	Measuring and Control Equipment
Bus Suppl	Business Supplies
Ship Cont	Shipping Containers
Transport	Transportation
Wholesale	Wholesale
Retail	Retail
Rest Hotels	Restaurants, Hotels, Motels
Banking	Banking
Insurance	Insurance
Real Estate	Real Estate
Trading	Trading
Others	Others

one were willing to estimate a GARCH model using small subsamples of the data, the realized volatility measures would be able to take full advantage of the rich information available in the daily data, while the GARCH model would discard all intramonthly data. This was one of the motivations for introducing realized volatility (Andersen et al., 2003). If the goal of our paper were instead to estimate a volatility model using the full sample of data, a GARCH model would be a natural starting point.

3. Full-sample results

3.1. Contemporaneous relationship

Following Den Haan (2000), we measure the comovement between stock return volatility and oil price volatility as the correlation of the residuals of a vector autoregressive (VAR) model⁵:

$$s_t = \alpha_0 + \alpha_1 s_{t-1} + \alpha_2 w_{t-1} + \varepsilon_{st} \quad (1)$$

⁵ The results presented here are robust to the use of longer lag lengths.

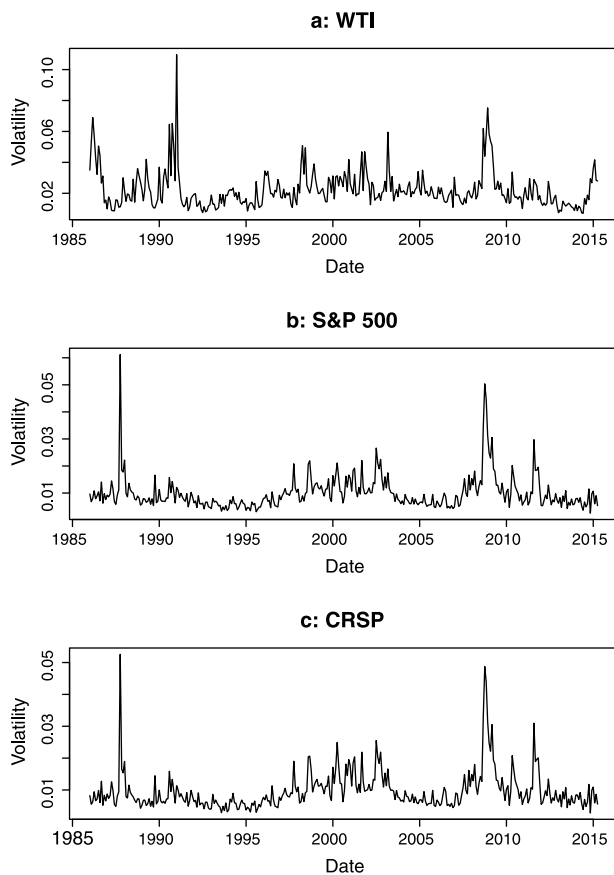


Fig. 1. Oil price and stock return volatility.

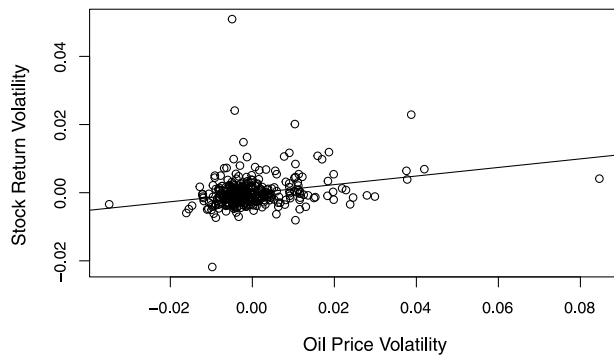


Fig. 2. VAR model residuals.

$$w_t = \beta_0 + \beta_1 s_{t-1} + \beta_2 w_{t-1} + \varepsilon_{wt} \quad (2)$$

where s_t and w_t are the realized volatility of the S&P 500 return and change in the price of WTI, respectively, in month t . Fig. 2 is a plot of ε_{st} against ε_{wt} for the period January 1986 to April 2015.

Overlaying the plot is the fitted regression line

$$\bar{\varepsilon}_{st} = -0.00 + 0.13\bar{\varepsilon}_{wt} + 0.05 I(1987 : 10) + \nu_t \quad (3)$$

(-0.71) (5.93) (13.40)

where $I(1987 : 10)$ is a dummy variable equal to 1 in October 1987 (the month of Black Monday) and 0 otherwise. t -Statistics are in parenthesis below the coefficient estimates.⁶

⁶ Estimating Eq. (3) with GARCH(1,1) estimates of the monthly volatility of the S&P 500 and WTI instead of the realized volatility measures yields a coefficient of 0.14, nearly identical to the reported estimate of 0.13.

There is a positive, statistically significant correlation between the two series, with an adjusted R -squared of 0.37. We can offer no interpretation beyond that without imposing additional assumptions; the correlation could be due to the effect of oil price volatility on the stock market, changes in the macroeconomy causing the two series to move together, or some combination of the two. The relationship in Fig. 2 is consistent with the negative effect of oil price shocks on stock returns that has been documented in the literature.

We have estimated the same regressions for the CRSP value-weighted index return and the 49 industry portfolio returns. Table 2 contains the estimated coefficients and adjusted R -squared values for all industries. The results for the CRSP value-weighted index are virtually identical to those for the S&P 500. All estimates of β reported in Table 2 are positive, covering a range from 0.05 to 0.18, and the t -statistics are less than 1.96 for only two industries. The largest adjusted R -squared, for the consumer goods portfolio, is 0.40. It is not surprising that oil price volatility explains so much of the volatility of the consumer goods sector, as higher energy prices often crowd out other forms of discretionary spending (see e.g. Gicheva, Hastings, & Villas-Boas, 2010).

3.2. Granger causality tests

The previous section established a strong contemporaneous relationship between oil price volatility and stock return volatility. We now turn to the question of whether there is a forecasting relationship between the two variables. We begin by testing for Granger causality from oil price volatility (σ_{oil}) to stock return volatility (σ_{stock}). We estimate regressions of the form

$$\sigma_{stock,t} = \alpha + \beta\sigma_{stock,t-1} + \gamma\sigma_{oil,t-1} + \varepsilon_t \quad (4)$$

and test $H_0 : \gamma = 0$.⁷

Intuitively, we expect oil price volatility to be useful as a predictor of stock return volatility ($\gamma \neq 0$) if (i) oil price volatility at time t has an effect on future observations of macroeconomic variables, interest rates, or other fundamental determinants of stock returns, and (ii) lagged stock return volatility does not adequately capture that information. A non-zero value of γ does not violate common definitions of market efficiency, which may rule out predictability of stock returns, but not stock return volatility.

Testing for predictability using a Granger causality test is a standard approach in the econometrics literature (see e.g. Hamilton, 1994). Alternatively, we could test for predictability by regressing the stock return volatility on only lagged oil price volatility:

$$\sigma_{stock,t} = \alpha_A + \gamma_A\sigma_{oil,t-1} + \varepsilon_t \quad (5)$$

This is similar to the model Driesprong, Jacobsen, and Maat (2008) used to evaluate predictability of stock returns after an oil price shock. The difficulty with interpreting the estimates of such a model is that it is possible to have $\gamma_A \neq 0$ but $\gamma = 0$ if lagged values of σ_{stock} capture the information in lagged values of σ_{oil} that is useful for predicting future stock return volatility. This situation could occur if σ_{stock} and σ_{oil} both reflect shocks to the macroeconomy, as in late 2008 and early 2009. Although the two volatility series would be strongly correlated across time, there is no reason to expect that to translate into a forecasting relationship, because lagged stock return volatility could fully account for macroeconomic volatility.

In practice, it is almost certain that γ_A will be different from zero regardless of the usefulness of oil price volatility as a predictor of stock return volatility. We have shown above that oil price volatility is strongly contemporaneously correlated with stock return volatility, and it is well-known (see below for additional evidence) that

⁷ The lag length was selected by the Schwarz information criterion.

Table 2
 Correlations of industry stock return volatility and oil price volatility.

Sector	Correlation	t-stat	\bar{R}^2	Sector	Correlation	t-stat	\bar{R}^2
CRSP	0.12	(5.89)	0.32	Shipbuild	0.05	(1.68)	0.07
Agriculture	0.05	(1.58)	0.03	Defense	0.15	(5.88)	0.29
Food Prod	0.09	(5.06)	0.32	Prec Metals	0.11	(2.91)	0.17
Candy Soda	0.11	(3.73)	0.24	Mining	0.09	(2.94)	0.19
Beer	0.10	(4.83)	0.26	Coal	0.12	(2.65)	0.04
Tobacco	0.12	(3.99)	0.12	Petroleum	0.11	(3.96)	0.24
Recreation	0.13	(4.76)	0.27	Utilities	0.06	(2.70)	0.16
Entertain	0.15	(4.35)	0.27	Communic	0.14	(6.18)	0.28
Printing	0.13	(4.72)	0.23	Pers Serv	0.11	(4.69)	0.25
Cons Goods	0.11	(5.14)	0.40	Bus Serv	0.12	(5.31)	0.29
Apparel	0.11	(4.45)	0.25	Computers	0.11	(3.46)	0.19
Healthcare	0.10	(4.09)	0.20	Comp Soft	0.14	(4.64)	0.30
Med Equip	0.11	(5.12)	0.24	Electro Equip	0.11	(3.70)	0.19
Pharma Prod	0.10	(4.66)	0.33	Meas Control	0.09	(3.69)	0.25
Chemicals	0.14	(5.79)	0.28	Bus Suppl	0.10	(4.93)	0.35
Rubber Plas	0.11	(5.33)	0.31	Ship Cont	0.10	(3.91)	0.26
Textiles	0.14	(4.61)	0.22	Transport	0.14	(5.75)	0.28
Constr Mat	0.13	(5.40)	0.27	Wholesale	0.09	(4.46)	0.23
Construct	0.16	(5.12)	0.21	Retail	0.12	(5.35)	0.32
Steel Works	0.15	(4.53)	0.26	Rest Hotels	0.10	(5.01)	0.33
Fabric Prod	0.14	(5.01)	0.17	Banking	0.15	(4.40)	0.14
Machinery	0.13	(4.82)	0.26	Insurance	0.13	(5.12)	0.19
Electric Equip	0.14	(5.69)	0.30	Real Estate	0.10	(3.51)	0.22
Autos	0.15	(5.41)	0.23	Trading	0.18	(5.61)	0.16
Aircraft	0.16	(5.62)	0.20	Others	0.18	(6.05)	0.22

Table 3
 Granger causality regressions.

	$\bar{\beta}$	$t(\bar{\beta})$	$\bar{\gamma}$	$t(\bar{\gamma})$	$\bar{\gamma}_A$	$t(\bar{\gamma}_A)$
S&P 500	0.60	13.43	0.04	1.58		
S&P 500					0.15	5.54
CRSP	0.63	14.36	0.03	1.38		
CRSP					0.14	5.56

stock return volatility is a serially correlated process, so we expect rejection of $\gamma_A = 0$ no matter the value of γ . $\bar{\gamma}_A$ will be picking up the predictive power of lagged stock return volatility even when lagged oil price volatility reveals nothing about future stock return volatility.

In Table 3 are the estimates of Eqs. (4) and (5) for the S&P 500 and CRSP. In both cases, $\bar{\gamma}_A$ is significant at a 5% level, with t-statistics greater than 5, but $\bar{\gamma}$ is not. The point estimates of γ_A are several times larger than the point estimates of γ in both cases. As explained above, it is not surprising to see significant estimates of γ_A , given estimates of β that show the importance of the autoregressive term in (4), and given the strong contemporaneous correlation of σ_{stock} and σ_{oil} .

It is possible that the aggregate results are masking predictability at the industry level. Table 4 presents results for all 49 industries. We are able to reject the null hypothesis of no Granger causality for six industries, with a positive coefficient in all cases, as expected, including several which are heavily energy-dependent or related to transportation: recreation, shipbuilding and railroad equipment, and shipping containers. This is broadly consistent with the findings of previous studies.

One drawback of the results reported in Table 4 is that it treats all oil price movements the same. Following the pioneering work of Kilian (2009) and Kilian and Park (2009), it has been common for researchers to allow shocks to oil supply and oil demand to have different effects on the macroeconomy and on stock returns. There is no reason to believe that oil price volatility due to oil demand shocks (which reflect shocks to world economic activity) will have the same effect on stock return volatility as oil price volatility due

to concerns about present and future oil supplies. We have reestimated Eq. (4) including a measure of the volatility of oil demand:

$$\sigma_{stock,t} = \alpha + \beta\sigma_{stock,t-1} + \gamma\sigma_{oil,t-1} + \delta\sigma_{REA,t-1} + \varepsilon_t \quad (6)$$

where $\sigma_{REA,t-1}$ is the fitted GARCH(1,1) volatility of the real economic activity index in Kilian (2009), updated to the end of our dataset.⁸ The results, reported in Table 5, suggest that the oil supply/oil demand distinction has no effect on our conclusions.

4. Parameter stability

A central concern when forecasting is that the parameters of the model should be stable over time (see e.g. Clark & McCracken, 2013). There are two reasons to be concerned about parameter instability with the stock return volatility-oil price volatility relationship in particular. First, a one-time event like the U.S. financial crisis could make it harder to find evidence of predictability in the full sample (due to outlier behavior) or easier (if there was a strong correlation between the variables only during the crisis). Second, the price of oil is driven by multiple shocks, including oil supply and aggregate demand shocks, and the relative importance of those shocks will change over time. A Granger causality test using the full sample might reject $\gamma = 0$ in Eq. (4), even though oil price volatility is not useful as a predictor of stock return volatility, or vice versa.

To get an overview of the degree of stability of the coefficient estimates, Fig. 3 presents plots of 10-year rolling window estimates of γ in Eq. (4) for the S&P 500 and CRSP. The estimates of γ follow similar patterns for the two indices. For the subsamples that end before 2001, $\bar{\gamma}$ is about zero. $\bar{\gamma}$ is negative and falling for the samples ending in the period 2003–2005. In 2008, after the fall of Lehmann Brothers, and at a time of extreme volatility of oil prices and stock returns, $\bar{\gamma}$ jumped sharply. The variation in the parameter estimates over time is so great that it is possible to find evidence for any desired correlation – positive, negative, or zero – by simply choosing an appropriate window of data. This calls into question the reliability of Granger causality tests as a tool for evaluating

⁸ Updated values of the index can be downloaded from Kilian's website at <http://www-personal.umich.edu/~lkilian/reaupdate.txt>.

Table 4
 Granger causality tests.

Sector	$\bar{\beta}$	$t(\bar{\beta})$	$\bar{\gamma}$	$t(\bar{\gamma})$	Sector	$\bar{\beta}$	$t(\bar{\beta})$	$\bar{\gamma}$	$t(\bar{\gamma})$
Agriculture	0.65	15.94	0.05	1.86	Defense	0.55	11.67	0.04	1.58
Food Prod	0.44	8.83	0.04	1.88	Prec Metals	0.63	14.82	0.02	0.50
Candy Soda	0.40	8.08	0.11	3.91	Mining	0.79	23.11	0.01	0.18
Beer	0.55	11.86	0.03	1.67	Coal	0.81	25.25	0.02	0.64
Tobacco	0.56	12.45	0.04	1.61	Petroleum	0.66	15.86	0.03	1.31
Recreation	0.45	9.11	0.07	2.38	Utilities	0.60	13.67	0.04	1.84
Entertain	0.63	14.59	0.06	1.84	Communic	0.65	14.94	0.03	1.31
Printing	0.66	15.68	0.04	1.41	Pers Serv	0.52	11.05	0.04	1.75
Cons Goods	0.47	9.61	0.04	1.69	Bus Serv	0.62	14.06	0.03	1.44
Apparel	0.63	14.69	0.05	2.13	Computers	0.71	18.31	0.02	0.71
Healthcare	0.49	10.18	0.04	1.54	Comp Soft	0.58	12.81	0.03	1.07
Med Equip	0.51	10.55	0.03	1.20	Electro Equip	0.74	19.91	0.02	0.78
Pharma Prod	0.48	10.06	0.02	1.03	Meas Control	0.74	19.87	0.03	1.19
Chemicals	0.66	15.72	0.03	1.28	Bus Suppl	0.58	12.98	0.03	1.51
Rubber Plas	0.62	14.35	0.03	1.15	Ship Cont	0.50	10.70	0.06	2.25
Textiles	0.68	16.86	0.04	1.43	Transport	0.54	11.40	0.04	1.64
Constr Mat	0.66	15.74	0.03	1.12	Wholesale	0.59	13.37	0.04	1.87
Construct	0.69	17.26	0.04	1.17	Retail	0.60	13.28	0.03	1.38
Steel Works	0.73	19.57	0.04	1.17	Rest Hotels	0.51	10.75	0.05	2.57
Fabric Prod	0.66	15.97	0.02	0.66	Banking	0.76	21.04	0.05	1.65
Machinery	0.68	16.81	0.04	1.66	Insurance	0.73	18.63	0.03	1.35
Electric Equip	0.63	14.76	0.05	1.95	Real Estate	0.77	22.18	0.04	1.31
Autos	0.64	15.11	0.04	1.54	Trading	0.79	22.50	0.02	0.68
Aircraft	0.54	11.42	0.05	1.93	Others	0.57	12.32	0.05	1.58
Shipbuild	0.55	12.29	0.07	2.98					

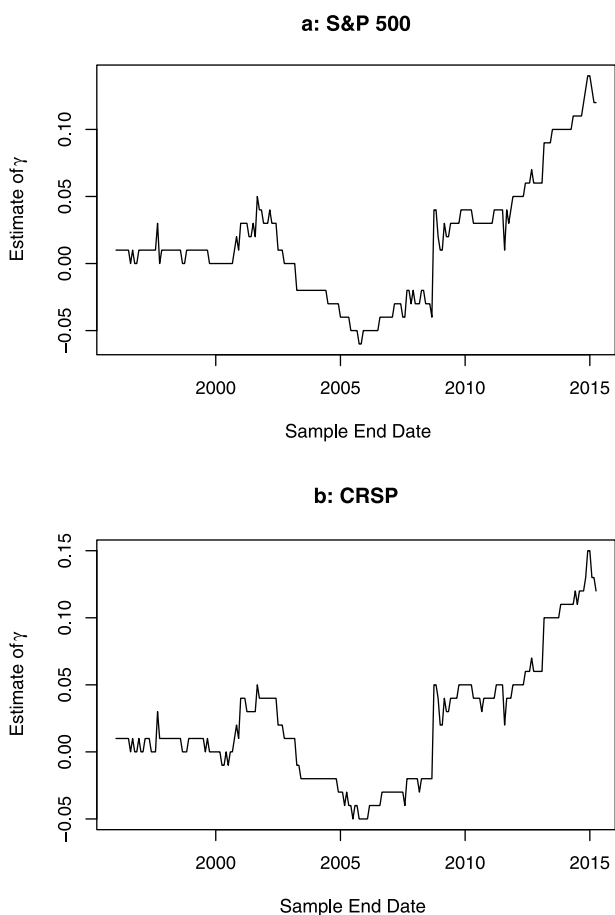


Fig. 3. Rolling window estimates.

the usefulness of oil price volatility as a predictor of stock return volatility.

We repeated the above exercise for each of the 49 industry portfolios. The results are summarized in Table 6, with the S&P 500 and CRSP included for comparison purposes. For each industry portfo-

lio, we report the smallest $\bar{\gamma}$, the largest $\bar{\gamma}$, and the p -value for the sup- F test of Andrews (1993) for a structural break at an unknown date. Our findings for the S&P 500 and CRSP indices carry through to all of the sectors, with $\bar{\gamma}$ changing signs in all but four cases, and even for those four industries, the estimates of γ cover a range of similar width to the estimates for the S&P 500. The null hypothesis of no structural break is rejected for every return volatility series. This is further evidence that Granger causality test results for the full sample are not a reliable way to assess the usefulness of oil price volatility as a predictor of stock return volatility.

5. Out-of-sample forecast evaluation

Motivated by the parameter instability found in the previous section, we do a comparison of the out-of-sample forecast performance of an autoregressive model:

$$\sigma_{stock,t} = \alpha + \beta\sigma_{stock,t-1} + \varepsilon_t^{AR} \tag{7}$$

against that of an ARX model that includes lagged oil price volatility⁹:

$$\sigma_{stock,t} = \alpha + \beta\sigma_{stock,t-1} + \gamma\sigma_{oil,t-1} + \varepsilon_t^{ARX} \tag{8}$$

We split the sample into an initial estimation period of January 1986 to December 1999, and a validation period of January 2000 to April 2015. The two models were estimated recursively, using all data that would have been available at the time a forecast was made. To make the initial forecasts, both models were estimated using data from January 1986 to December 1999, and the estimated models were used to make forecasts of σ_{stock} in January 2000. The dataset was updated to include data through January 2000, the two models were reestimated, and forecasts were made of σ_{stock} in February 2000. The process was repeated until forecasts of σ_{stock} were produced for all 184 observations in the validation period.

⁹ This is equivalent to a vector autoregressive forecast of stock return volatility.

Table 5
 Granger causality tests with the volatility of real economic activity index.

Sector	$\bar{\beta}$	$t(\bar{\beta})$	$\bar{\gamma}$	$t(\bar{\gamma})$	$\bar{\delta}$	$t(\bar{\delta})$
S&P 500	0.60	13.14	0.04	1.58	0.00	0.33
CRSP	0.62	13.91	0.03	1.39	0.00	0.55
Agriculture	0.65	15.43	0.05	1.86	0.00	0.19
Food Prod	0.44	8.79	0.04	1.88	0.00	-0.03
Candy Soda	0.39	7.85	0.12	4.11	0.00	-2.29
Beer	0.53	11.55	0.04	1.90	0.00	-2.37
Tobacco	0.54	11.80	0.05	1.91	0.00	-2.66
Recreation	0.43	8.69	0.07	2.40	0.00	1.36
Entertain	0.60	13.45	0.07	1.92	0.00	1.69
Printing	0.61	13.51	0.04	1.63	0.00	2.47
Cons Goods	0.47	9.59	0.04	1.69	0.00	-0.13
Apparel	0.62	13.93	0.05	2.17	0.00	1.00
Healthcare	0.49	10.07	0.04	1.54	0.00	0.25
Med Equip	0.50	10.46	0.03	1.20	0.00	0.12
Pharma Prod	0.48	10.02	0.02	1.10	0.00	-1.30
Chemicals	0.65	14.69	0.03	1.34	0.00	1.05
Rubber Plas	0.59	12.96	0.03	1.27	0.00	2.15
Textiles	0.62	14.08	0.05	1.72	0.00	3.22
Constr Mat	0.59	12.97	0.03	1.42	0.00	3.12
Construct	0.68	16.07	0.04	1.24	0.00	1.13
Steel Works	0.72	18.15	0.04	1.23	0.00	0.97
Fabric Prod	0.60	13.20	0.02	0.92	0.00	3.08
Machinery	0.66	15.69	0.05	1.72	0.00	1.21
Electric Equip	0.63	14.35	0.05	1.96	0.00	0.56
Autos	0.63	14.23	0.04	1.59	0.00	1.19
Aircraft	0.54	11.24	0.05	1.93	0.00	0.35
Shipbuild	0.52	11.40	0.07	3.01	0.00	2.10
Defense	0.55	11.66	0.04	1.59	0.00	-0.40
Prec Metals	0.63	14.54	0.02	0.50	0.00	0.10
Mining	0.77	20.25	0.01	0.26	0.00	1.03
Coal	0.81	23.33	0.02	0.67	0.00	0.40
Petroleum	0.66	15.27	0.03	1.32	0.00	0.21
Utilities	0.59	13.23	0.04	1.84	0.00	0.32
Communic	0.65	14.92	0.03	1.32	0.00	-0.59
Pers Serv	0.50	10.20	0.04	1.81	0.00	1.71
Bus Serv	0.61	13.53	0.03	1.46	0.00	0.91
Computers	0.70	18.14	0.02	0.82	0.00	-1.40
Comp Soft	0.56	12.40	0.04	1.28	0.00	-2.19
Electro Equip	0.73	19.69	0.03	0.90	0.00	-1.62
Meas Control	0.74	19.80	0.03	1.20	0.00	-0.42
Bus Suppl	0.57	12.33	0.03	1.53	0.00	1.30
Ship Cont	0.50	10.64	0.06	2.24	0.00	-0.08
Transport	0.51	10.59	0.04	1.71	0.00	1.76
Wholesale	0.58	12.49	0.04	1.93	0.00	1.40
Retail	0.59	13.26	0.03	1.45	0.00	-1.21
Rest Hotels	0.51	10.72	0.05	2.57	0.00	-0.26
Banking	0.75	19.37	0.06	1.75	0.00	1.16
Insurance	0.71	17.08	0.04	1.44	0.00	1.02
Real Estate	0.74	19.30	0.04	1.45	0.00	1.73
Trading	0.79	21.87	0.02	0.67	0.00	-0.08
Others	0.57	12.30	0.05	1.58	0.00	-0.05

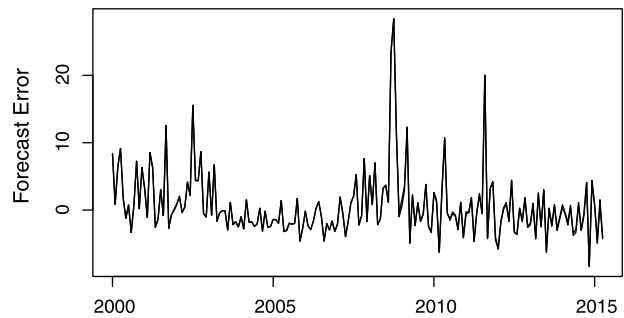


Fig. 4. Out-of-sample forecast errors.

$\sigma_{stock,it}^A$ is observed volatility of stock index i in month t . A second metric is the maximum absolute percentage difference:

$$MaxAPD_i = \max \left(\frac{|\sigma_{stock,it}^{OOS} - \sigma_{stock,it}^{IS}|}{\sigma_{stock,it}^A} \right) \quad (10)$$

$MaxAPD$ gives a measure of the riskiness of doing inference using in-sample predictions rather than out-of-sample forecasts, by providing information about the worst possible outcome. It is in the cases where the two forecasts diverge substantially that portfolio decisions would be affected the most.

The results, presented in Table 7, can be summarized as follows. For some sectors, the mean difference in forecasts is quite small. That is not surprising, because the weight put on the oil price volatility term is close to zero for those sectors, so it does not make much difference how the forecasts are constructed. For other sectors the difference is important, representing an average deviation of about 10% across all observations. The maximum difference in the forecasts exceeds 20% of volatility in most sectors, with some as high as 100%. Looking at the average difference across all observations masks the large differences that occur in the time periods when the parameters are changing. Focusing exclusively on in-sample predictions is a dangerous proposition in the presence of parameter instability.

5.2. Forecast evaluation results

Letting $\bar{\sigma}_{stock,it}$ be the out-of-sample forecast of stock return volatility from model i in month t , the out-of-sample forecast error for model i in month t , e_{it} , can be calculated as

$$e_{it} = \sigma_{stock,t} - \bar{\sigma}_{stock,it} \quad (11)$$

The forecast errors for the two models (scaled by 1000 for readability) can be found in Fig. 4. The two series are so similar as to be almost indistinguishable. There is no obvious advantage to using one or the other of the models based on Fig. 4.

Following Diebold and Mariano (1995), a common approach to comparing forecasting models is to calculate the loss differential series. For our models, assuming MSE loss, the loss differential series is

$$d_t = e_{AR,t}^2 - e_{ARX,t}^2 \quad (12)$$

A positive value of d_t indicates that the forecast loss associated with the month t AR model forecast was greater, a negative value indicates that the ARX model forecast loss was greater, and a value of zero indicates that the models forecast equally well. Fig. 5 is a plot of the loss differential series. It is in most cases small relative to the squared forecast errors and there is no obvious tendency for

5.1. How different are in-sample and out-of-sample forecasts?

We begin by asking how different the in-sample and out-of-sample stock return volatility forecasts are. On one hand, it is more convenient to do inference on in-sample predictions. In practice, however, all stock return forecasts are by definition made in an out-of-sample fashion, so it is only reasonable to draw conclusions from in-sample analysis if the two methods produce forecasts that are about the same. One metric for measuring the quality of approximation provided by in-sample forecasts is the mean absolute percentage difference in the forecasts:

$$MAPD_i = \text{mean} \left(\frac{|\sigma_{stock,it}^{OOS} - \sigma_{stock,it}^{IS}|}{\sigma_{stock,it}^A} \right) \quad (9)$$

where $\sigma_{stock,it}^{OOS}$ is the out-of-sample forecast of stock index i in month t , $\sigma_{stock,it}^{IS}$ is the in-sample forecast of stock index i in month t , and

Table 6
 Rolling window estimates by industry.

Sector	min($\bar{\gamma}$)	max($\bar{\gamma}$)	sup-F	Sector	min($\bar{\gamma}$)	max($\bar{\gamma}$)	sup-F
S&P 500	-0.06	0.14	0.00	Shipbuild	-0.02	0.19	0.00
CRSP	-0.05	0.15	0.00	Defense	-0.04	0.21	0.00
Agriculture	0.00	0.20	0.00	Prec Metals	-0.11	0.24	0.01
Food Prod	-0.04	0.13	0.00	Mining	-0.17	0.23	0.00
Candy Soda	0.03	0.26	0.00	Coal	-0.11	0.25	0.00
Beer	-0.02	0.18	0.00	Petroleum	-0.08	0.22	0.00
Tobacco	-0.01	0.20	0.00	Utilities	-0.03	0.16	0.00
Recreation	-0.04	0.27	0.00	Communic	-0.05	0.17	0.00
Entertain	-0.01	0.22	0.00	Pers Serv	-0.01	0.15	0.00
Printing	-0.06	0.20	0.00	Bus Serv	-0.02	0.14	0.00
Cons Goods	-0.08	0.18	0.00	Computers	-0.03	0.19	0.00
Apparel	0.01	0.22	0.00	Comp Soft	-0.04	0.17	0.00
Healthcare	-0.03	0.22	0.00	Electro Equip	-0.03	0.17	0.00
Med Equip	-0.08	0.14	0.00	Meas Control	-0.02	0.16	0.00
Pharma Prod	-0.08	0.12	0.00	Bus Suppl	-0.03	0.18	0.00
Chemicals	-0.03	0.20	0.00	Ship Cont	-0.03	0.23	0.00
Rubber Plas	-0.03	0.20	0.00	Transport	-0.02	0.24	0.00
Textiles	-0.04	0.26	0.00	Wholesale	-0.01	0.16	0.00
Constr Mat	-0.03	0.19	0.00	Retail	-0.03	0.13	0.00
Construct	-0.05	0.20	0.00	Rest Hotels	-0.01	0.15	0.00
Steel Works	-0.03	0.29	0.00	Banking	-0.07	0.31	0.00
Fabric Prod	-0.03	0.24	0.00	Insurance	-0.04	0.18	0.00
Machinery	-0.01	0.25	0.00	Real Estate	-0.08	0.27	0.00
Electric Equip	0.00	0.19	0.00	Trading	-0.08	0.19	0.00
Autos	-0.05	0.26	0.00	Others	-0.07	0.17	0.00
Aircraft	-0.05	0.21	0.00				

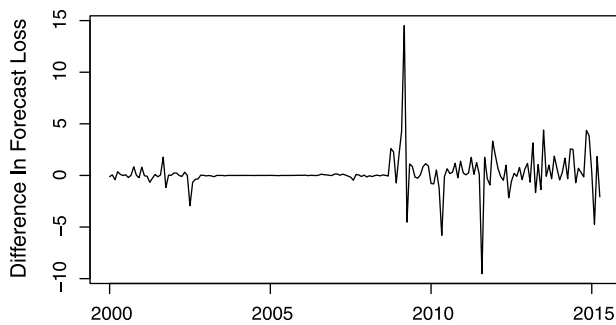


Fig. 5. Forecast loss differential series.

it to be positive.¹⁰ The pattern of the loss differential series in Fig. 5 suggests that the financial crisis and the period that followed may have been different from the rest of the sample. We accommodate this by using dummies to allow the relative forecast performance of the two models to be different during the crisis period.

In Table 8 are estimates of the regressions

$$d_t = \alpha + \varepsilon_t \tag{13}$$

$$d_t = \alpha + \beta_1 I_{1t} + \varepsilon_t \tag{14}$$

$$d_t = \alpha + \beta_1 I_{1t} + \beta_2 I_{2t} + \varepsilon_t \tag{15}$$

where $I_{1t} = 1$ in the period 2008–2009 and zero otherwise, and $I_{2t} = 1$ in the period 2010–2011 and zero otherwise.

The estimates in column 1 of Table 8 confirm that the AR model has a larger MSE for the full out-of-sample period. The estimates in columns 2 and 3 suggest that the difference in MSE is driven largely by the 2008–2009 time period, in the immediate aftermath of the financial crisis, when stock returns and oil prices were experiencing high volatility.

A Diebold and Mariano (DM: 1995) comparison of the models could be done using the estimates of α in Table 8. The downside of

Table 7
 In-sample vs. out-of-sample predictions.

Sector	MAPD	MaxAPD	Sector	MAPD	MaxAPD
S&P 500	0.04	0.26	Shipbuild	0.03	0.19
CRSP	0.04	0.26	Defense	0.03	0.17
Agriculture	0.07	0.46	Prec Metals	0.01	0.13
Food Prod	0.04	0.17	Mining	0.09	0.40
Candy Soda	0.07	0.41	Coal	0.03	0.14
Beer	0.08	0.38	Petroleum	0.04	0.21
Tobacco	0.06	0.25	Utilities	0.06	0.61
Recreation	0.03	0.21	Communic	0.05	0.31
Entertain	0.05	0.42	Pers Serv	0.04	0.21
Printing	0.05	0.52	Bus Serv	0.04	0.27
Cons Goods	0.04	0.22	Computers	0.03	0.24
Apparel	0.03	0.21	Comp Soft	0.07	0.33
Healthcare	0.04	0.22	Electro Equip	0.03	0.29
Med Equip	0.04	0.18	Meas Control	0.02	0.24
Pharma Prod	0.05	0.18	Bus Suppl	0.04	0.24
Chemicals	0.05	0.35	Ship Cont	0.04	0.25
Rubber Plas	0.06	0.37	Transport	0.06	0.58
Textiles	0.09	1.01	Wholesale	0.04	0.35
Constr Mat	0.06	0.55	Retail	0.03	0.13
Construct	0.06	0.28	Rest Hotels	0.03	0.22
Steel Works	0.06	0.47	Banking	0.06	0.31
Fabric Prod	0.07	0.46	Insurance	0.07	0.46
Machinery	0.05	0.29	Real Estate	0.09	0.57
Electric Equip	0.03	0.17	Trading	0.03	0.14
Autos	0.06	0.57	Others	0.06	0.33
Aircraft	0.03	0.21			

Table 8
 Loss differential across subsamples.

	Eq. (13)	Eq. (14)	Eq. (15)
α	0.19	0.07	0.13
$t(\alpha)$	1.49	0.51	0.88
β_1		0.95	0.89
$t(\beta_1)$		2.51	2.32
β_2			-0.41
$t(\beta_2)$			-1.06

that approach is that the distribution of the DM statistic is nonstandard when the models are nested. We instead apply the ENC-NEW test of Clark and McCracken (2001) to test the null hypothesis that

¹⁰ The loss differential series is positive (a smaller loss associated with the ARX model forecast) 55% of the time. The MSE for the ARX model is 23.2, the MSE for the AR model is 23.4, and the MSE ratio for the two models is 0.99.

Table 9
ENC-NEW test statistics by industry.

Sector	Full sample	1 lag	2 lags	Sector	Full sample	1 lag	2 lags
S&P 500	0.93	0.52	0.20	Shipbuild	3.44	-1.28	-0.78
CRSP	0.51	0.21	-0.06	Defense	0.80	0.42	0.54
Agriculture	0.56	-0.50	-0.31	Prec Metals	-1.94	-0.89	-0.15
Food Prod	1.48	1.02	0.72	Mining	-0.86	0.20	0.14
Candy Soda	7.33	3.12	1.35	Coal	-0.28	-0.11	-0.15
Beer	0.81	0.88	0.61	Petroleum	-0.14	-0.16	0.06
Tobacco	1.09	0.89	0.38	Utilities	0.33	0.17	0.11
Recreation	3.20	1.28	1.84	Communic	0.12	0.29	0.02
Entertain	1.18	-0.21	-0.35	Pers Serv	1.20	-0.11	-0.39
Printing	0.32	0.01	-0.27	Bus Serv	0.60	0.09	-0.14
Cons Goods	1.10	0.76	0.65	Computers	-0.39	-0.36	0.04
Apparel	1.65	0.40	-0.39	Comp Soft	-0.29	-0.21	0.17
Healthcare	-0.09	-0.69	0.46	Electro Equip	-0.23	-0.18	-0.01
Med Equip	-0.25	0.45	1.05	Meas Control	0.16	0.08	0.01
Pharma Prod	-0.13	0.54	0.53	Bus Suppl	0.39	0.22	0.18
Chemicals	0.22	-0.25	-0.37	Ship Cont	2.26	0.92	0.55
Rubber Plas	0.00	-0.24	-0.29	Transport	0.56	-0.05	0.20
Textiles	0.53	-0.72	-1.00	Wholesale	1.02	-0.05	-0.40
Constr Mat	0.26	-0.87	-1.05	Retail	0.63	0.48	-0.05
Construct	0.27	-0.57	-0.34	Rest Hotels	3.12	1.29	0.81
Steel Works	-0.12	-0.39	-0.07	Banking	0.64	-0.30	-0.25
Fabric Prod	-0.42	-0.59	-0.42	Insurance	0.42	0.01	-0.06
Machinery	0.74	-0.18	-0.23	Real Estate	0.13	-0.19	0.34
Electric Equip	1.44	0.42	0.07	Trading	-0.03	-0.10	-0.32
Autos	0.52	-0.04	-0.01	Others	0.97	0.31	0.39
Aircraft	1.00	0.26	0.27				

the two models forecast equally well against the alternative that the ARX model forecasts have a lower MSE. We do the test for the S&P 500, CRSP, and each of the 49 industry portfolios. The AIC and SIC select a lag length of one, but to confirm that our results are not sensitive to this choice, we also report results for a lag length of two. The tabulated 95% critical values provided in [Clark and McCracken \(2001\)](#) are 2.234 for one lag and 2.709 for two lags.

The ENC-NEW test statistics can be found in [Table 9](#). The columns titled “Full Sample” report the ENC-NEW statistic calculated using the full out-of-sample period. Five industries have test statistics greater than the critical value of 2.234: Candy and Soda, Recreation, Shipbuilding, Shipping Containers, and Restaurants and Hotels. As shown in the previous section, however, the 2008–2009 time period was one in which both volatility series were larger than normal due to the U.S. financial crisis, and there is evidence of parameter instability as a result. Therefore, the evidence that oil price volatility is useful as a predictor of stock return volatility for these industries can be called into question. The columns titled “1 Lag” and “2 Lag” are the ENC-NEW statistics calculated by dropping the forecasts for the 2008–2009 time period. Dropping those observations causes the ENC-NEW test statistics to drop in nearly all cases. The only industry for which we can reject the null hypothesis is Candy and Soda. Our results cannot be attributed to the use of an overly parsimonious model, as we are not able to reject the null hypothesis for any of the industries when using a longer lag length.

Except for the Candy and Soda industry, stock return volatility forecasts cannot be improved by accounting for oil price volatility. The relationship is plagued by instabilities following from the greater volatility of the financial crisis period.

6. Conclusions

This paper has revisited the question of whether oil price volatility is useful as a predictor of stock price volatility. There is a strong, positive contemporaneous relationship between the two volatility series. Consistent with previous studies, there is clear evidence of a predictive relationship when doing inference on the full sample.

The results are different when we evaluate of the ability of oil price volatility to improve out-of-sample forecasts of stock return volatility. Formal out-of-sample predictive ability tests find no evidence that oil price volatility can be used to improve forecasts of stock return volatility.¹¹ Further investigation reveals that the relationship between the two volatility series fluctuates wildly through time.¹² The changes in the relationship are not just in magnitude, but also in sign. One could find a strong positive relationship, a strong negative relationship, or no relationship at all, simply by choosing an appropriate subsample. Therefore, in spite of the reasonableness of the argument that there should be a link between stock market volatility and oil market volatility, we conclude that it cannot be exploited in practice.¹³

Our results provide no support for the hypothesis that oil price volatility should be used as a predictor of stock return volatility. Thus, monetary and fiscal policy authorities should not adjust policy in response to high oil price volatility, unless there are other concerns about oil price volatility beyond the effects on stock price volatility.

Conflict of interest

None.

Acknowledgement

We thank the anonymous reviewer for the insightful comments and suggestions.

¹¹ Further analysis (results are available upon request) showed that our out-of-sample forecast findings are not sensitive to the choice of time period.

¹² Note that we are referring here to a forecasting relationship, not a conditional correlation.

¹³ We repeated the analysis using Brent spot prices. The results, available upon request, do not change any of our conclusions.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <https://doi.org/10.1016/j.qref.2018.01.001>.

References

- Alsalman, Z., & Herrera, A. M. (2015). Oil price shocks and the U.S. stock market: Do sign and size matter? *The Energy Journal*, 36(3), 171–188.
- Andersen, T. G., Bollerslev, T., Diebold, F. X., & Labys, P. (2003). Modeling and forecasting realized volatility. *Econometrica*, 71(2), 579–625.
- Andrews, D. W. K. (1993). Tests for parameter instability and structural change with unknown change point. *Econometrica*, 61(4), 821–856.
- Apergis, N., & Miller, S. M. (2009). Do structural oil-market shocks affect stock prices? *Energy Economics*, 31(4), 569–575.
- Atems, B., Kapper, D., & Lam, E. (2015). Do exchange rates respond asymmetrically to shocks in the crude oil market? *Energy Economics*, 49, 227–238.
- Basher, S. A., Haug, A. A., & Sadorsky, P. (2012). Oil prices, exchange rates and emerging stock markets. *Energy Economics*, 34(1), 227–240.
- Chen, S. (2010). Do higher oil prices push the stock market into bear territory? *Energy Economics*, 32(2), 490–495.
- Clark, T. E., & McCracken, M. W. (2001). Tests of equal forecast accuracy and encompassing for nested models. *Journal of Econometrics*, 105, 85–110.
- Clark, T. E., & McCracken, M. W. (2013). *Advances in forecast evaluation. Handbook of economic forecasting* (Vol. 2) Amsterdam: North Holland.
- Cunado, J., & De Gracia, F. P. (2014). Oil price shocks and stock market returns: Evidence for some European countries. *Energy Economics*, 42, 365–377.
- Den Haan, W. J. (2000). The comovement between output and prices. *Journal of Monetary Economics*, 46(1), 3–30.
- Diebold, F. X., & Mariano, R. S. (1995). Comparing predictive accuracy. *Journal of Business and Economic Statistics*, 13(3), 253–263.
- Driesprong, G., Jacobsen, B., & Maat, B. (2008). Striking oil: Another puzzle? *Journal of Financial Economics*, 89(2), 307–327.
- Edelstein, P., & Kilian, L. (2009). How sensitive are consumer expenditures to retail energy prices? *Journal of Monetary Economics*, 56(6), 766–779.
- Elyasiani, E., Mansur, I., & Odusami, B. (2011). Oil price shocks and industry stock returns. *Energy Economics*, 33(5), 966–974.
- Farmer, R. E. A. (2012). The stock market crash of 2008 caused the great recession: Theory and evidence. *Journal of Economic Dynamics and Control*, 36(5), 693–707.
- Gicheva, D., Hastings, J., & Villas-Boas, S. (2010). Investigating income effects in scanner data: Do gasoline prices affect grocery purchases? *American Economic Review*, 100(2), 480–484.
- Hamilton, J. D. (1994). *Time series analysis* (Vol. 2) Princeton, NJ: Princeton University Press.
- Hamilton, J. D. (2011). Nonlinearities and the macroeconomic effects of oil prices. *Macroeconomic Dynamics*, 15(S5), 364–378.
- Hammoudeh, S., Dibooglu, S., & Aleisa, E. (2004). Relationships among U.S. oil prices and oil industry equity indices. *International Review of Economics & Finance*, 13(4), 427–453.
- Herrera, A. M., & Pesavento, E. (2009). Oil price shocks, systematic monetary policy and the “great moderation”. *Macroeconomic Dynamics*, 13(1), 107–137.
- Herrera, A. M., Lagalo, L. G., & Wada, T. (2011). Oil price shocks and industrial production: Is the relationship linear? *Macroeconomic Dynamics*, 15(S3), 472–497.
- Jones, C. M., & Kaul, G. (1996). Oil and the stock markets. *Journal of Finance*, 51(2), 463–491.
- Kilian, L. (2009). Not all oil price shocks are alike: Disentangling demand and supply shocks in the crude oil market. *American Economic Review*, 99(3), 1053–1069.
- Kilian, L., & Park, C. (2009). The impact of oil price shocks on the U.S. stock market. *International Economic Review*, 50(4), 1267–1287.
- Kilian, L., & Lewis, L. T. (2010). Does the Fed respond to oil price shocks? *The Economic Journal*, 121(555), 1047–1072.
- Kilian, L., & Vigfusson, R. J. (2011). Are the responses of the U.S. economy asymmetric in energy price increases and decreases? *Quantitative Economics*, 2(3), 419–453.
- Melichar, M. (2016). Energy price shocks and economic activity: Which energy price series should we be using? *Energy Economics*, 54, 431–443.
- Pettenuzzo, D., & Timmerman, A. (2011). Predictability of stock returns and asset allocation under structural breaks. *Journal of Econometrics*, 164(1), 60–78.
- Phelps, E. S. (1999). Behind this structural boom: The role of asset valuations. *American Economic Review*, 89(2), 63–68.
- Sadorsky, P. (1999). Oil price shocks and stock market activity. *Energy Economics*, 21(5), 449–469.
- Sadorsky, P. (2003). The macroeconomic determinants of technology stock price volatility. *Review of Financial Economics*, 12(2), 191–205.
- Schwert, W. G. (1989). Why does stock market volatility change over time? *Journal of Finance*, 44(5), 1115–1153.