An example of regression-based forecast
(CHEN, 2013)

Ka-fu WONG
University of Hong Kong

April 10, 2014

Contents

1. Paper

Chen, Shiu-sheng (2013): "Forecasting Crude Oil Price Movements with Oil-Sensitive Stocks," Economic Inquiry (forthcoming).

陳旭昇 [Shiu-Sheng Chen]
國立台灣大學經濟系教授
美國威斯康辛大學麥迪遜校區經濟博士
Professor of Economics, National Taiwan University
Ph.D. in Economics, University of Wisconsin-Madison

研究興趣 [Fields of Interest]
總體與貨幣, 國際金融, 能源經濟
Macro and Monetary Economics
International Finance
Energy Economics

2. Purpose

- To find a model that will beat "no change" forecast.
- Regression based models that incorporate oil-sensitive stock price indices can outperform the no-change model.

3. Model for forecasting nominal crude oil prices

Run the regression using the first R observations

$$y_{t+h} = \alpha + \beta x_t + u_{t+h} \qquad t = 2, \ldots, R-h$$

where

$$y_{t+h} \equiv \frac{op_{t+h} - op_t}{op_t}, \qquad x_t \equiv \frac{sp_t - sp_{t-1}}{sp_{t-1}}$$

t	sp_t	x_t	op_t	y_t
1	sp_1		op_1	–
2	sp_2	x_2	op_2	–
3	sp_3	x_3	op_3	–
				–
h			op_h	–
$h+1$			op_{1+h}	y_{1+h}

Forecast of $\hat{y}_{R+h}$

$$\hat{y}_{t+h} = \frac{\widehat{op}_{t+h} - op_t}{op_t} = \hat{\alpha} + \hat{\beta} x_t, \quad t = R$$

Forecast of $\widehat{op}_{R+h}$

$$\widehat{op}_{t+h} = (1 + \hat{\alpha} + \hat{\beta} x_t) \times op_t, \quad t = R$$

Repeat by adding one additional observations to the regression sample. So we will have forecast form sample with the following observations in the gression sample:

$$R, R+1, R+2, ..., T-h$$

Compute the forecast error of the regression-based model as

$$e_{t+h}(M^{SP}) = op_{t+h} - \widehat{op}_{t+h}, \quad t = R, R+1, ..., T-h$$

Compute the forecast error of a no-change model as

$$e_{t+h}(M^{NC}) = op_{t+h} - op_t, \quad t = R, R+1, ..., T-h$$

(sp_t, op_t)		(x_t, y_t)						
$(1, ..., R)$	$\longrightarrow$	$(2, ..., R-h)$	$\longrightarrow$	$\hat{y}_{R+h}$	$\longrightarrow$	$\widehat{op}_{R+h}$	$\longrightarrow$	e_{R+h}
$(1, ..., R+1)$	$\longrightarrow$	$(2, ..., R-h+1)$	$\longrightarrow$	$\hat{y}_{R+1+h}$	$\longrightarrow$	$\widehat{op}_{R+1+h}$	$\longrightarrow$	e_{R+1+h}
...	...	...		...		...		...
$(1, ..., T-h)$	$\longrightarrow$	$(2, ..., T-2h+1)$	$\longrightarrow$	$\hat{y}_T$	$\longrightarrow$	$\widehat{op}_T$	$\longrightarrow$	e_T

Compute the MSPE

$$MSPE(M^{SP}) = \frac{\sum_{t=R}^{T-h} e_{t+h}^2(M^{SP})}{T-h-R+1},$$

$$MSPE(M^{NC}) = \frac{\sum_{t=R}^{T-h} e_{t+h}^2(M^{NC})}{T-h-R+1}$$

Compute the ratio of MSPEs as

$$\frac{MSPE(M^{SP})}{MSPE(M^{NC})}$$

4. Model for forecasting REAL crude oil prices

Run the regression using the first R observations

$$y_{t+h} = \alpha + \beta x_t + u_{t+h}, \quad t = 2, \ldots, R - h$$

where

$$y_{t+h} \equiv \frac{rop_{t+h} - rop_t}{rop_t}, \qquad x_t \equiv \frac{rsp_t - rsp_{t-1}}{rsp_{t-1}}, \qquad rop_t = \frac{op_t}{cpi_t}$$

Forecast of

$$\hat{y}_{t+h} \equiv \frac{\widehat{rop}_{t+h} - rop_t}{rop_t} = \hat{\alpha} + \hat{\beta} x_t, \quad t = R$$

Forecast of $\widehat{rop}_{t+h}$

$$\widehat{rop}_{R+h} = (1 + \hat{\alpha} + \hat{\beta} x_t) \times rop_t, \quad t = R$$

Repeat by adding one additional observations to the regression sample. So we will have forecast form sample with the following observations in the gression sample:

$$R, R+1, R+2, ..., T-h$$

Compute the forecast error of the regression-based model as

$$e_{t+h}(RM^{SP}) = rop_{t+h} - \widehat{rop}_{t+h}, \quad t = R, R+1, ..., T-h$$

Compute the forecast error of a no-change model as

$$e_{t+h}(RM^{NC}) = rop_{t+h} - rop_t, \quad t = R, R+1, ..., T-h$$

Compute the MSPE

$$MSPE(RM^{SP}) = \frac{\sum_{t=R}^{T-h} e_{t+h}^2(RM^{SP})}{T-h-R+1},$$

$$MSPE(RM^{NC}) = \frac{\sum_{t=R}^{T-h} e_{t+h}^2(RM^{NC})}{T-h-R+1}$$

Compute the ratio of MSPEs as

$$\frac{MSPE(RM^{SP})}{MSPE(RM^{NC})}$$

5. Data

- Frequency of data: monthly
- Sample period: 1984:M10 to 2012:M8
- First regression sample: 1984:M10 to 1990:M12
- Out-of-sample forecast period: 1991:M1 to 2012:M8

Table 2: Data Description

Variables	Code	Source
Baseline (1984M10–2012M8)		
WTI	14176AAZZF.M17	IFS
Brent	46276AAZZF...	IFS
Dubai	46676AAZZF...	IFS
World Average	00176AAZZF...	IFS
AMEX Oil Index	^XOI	Yahoo Finance
S&P 500	^GSPC	Yahoo Finance
U.S. Consumer Price Index	CPIAUCSL	FRED
MSCI World Sector Index (1995M1–2012M8)		
(1) Energy	M1DWE1$	Datastream
(1.1) Energy Equipment & Services	M3DWES$	Datastream
(1.2) Oil & Gas	M3DWOG$	Datastream
(2) Transportation	M2DWTR$	Datastream
(2.1) Air Freight & Logistics	M3DWAF$	Datastream
(2.2) Airlines	M3DWAL$	Datastream
(2.3) Marine	M3DWMA$	Datastream
(2.4) Road & Rail	M3DWRR$	Datastream

Figure 1: Crude Oil Prices.

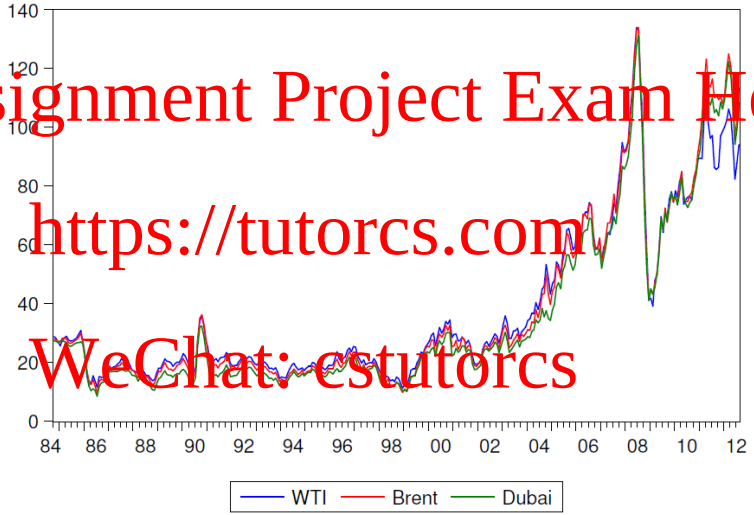
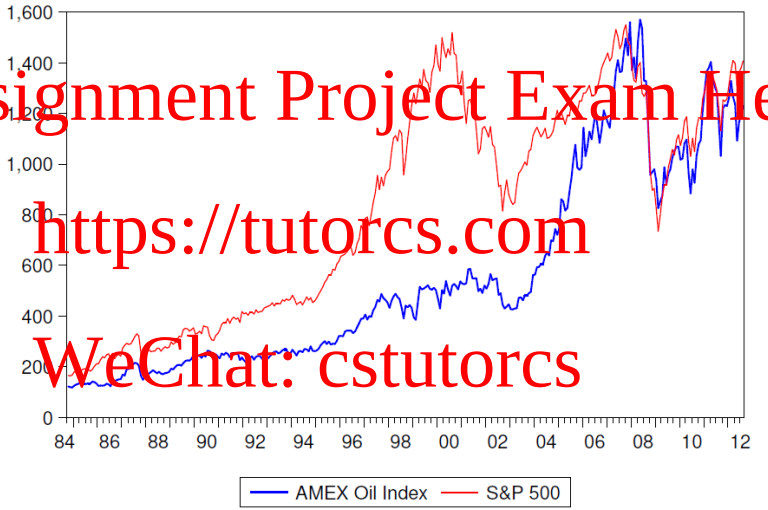


Figure 2: AMEX Oil Index and S&P 500 Index



6. In-sample fit

Table 3 (a): In-sample Predictability of Nominal Oil Prices

$$y_{t+h} = \alpha + \beta x_t + u_{t+h} \quad t = 2, ..., R - h$$

	WTI			Brent		
	$\hat{\beta}$	t-stat	p-value	$\hat{\beta}$	t-stat	p-value
Baseline (1984:M10–2012:M8)						
AMEX Oil Index	0.43	2.95	**0.00**	0.48	3.95	**0.00**
S&P 500	0.01	0.07	0.95	0.03	0.14	0.89
MSCI (1995:M1–2012:M8)						
Energy	0.28	2.51	**0.01**	0.36	3.17	**0.00**
Energy Equip. & Services	0.22	2.82	**0.00**	0.27	3.28	**0.00**
Oil & Gas	0.28	2.61	**0.01**	0.34	3.03	**0.00**
Transportation	0.16	0.95	0.34	0.25	1.49	0.14
Air Freight & Logistics	0.05	0.46	0.65	0.09	0.80	0.43
Airlines	0.10	1.03	0.30	0.19	1.99	**0.05**
Marine	0.20	1.69	0.09	0.23	1.92	**0.05**
Road & Rail	0.03	0.19	0.85	0.10	0.63	0.53

Table 3 (b): In-sample Predictability of Nominal Oil Prices

$$y_{t+h} = \alpha + \beta x_t + u_{t+h} \quad t = 2, ..., R - h$$

	Dubai			Average		
	$\hat{\beta}$	t-stat	p-value	$\hat{\beta}$	t-stat	p-value
Baseline (1984:M10–2012:M8)						
AMEX Oil Index	0.51	4.53	**0.00**	0.47	4.20	**0.00**
S&P 500	0.03	0.15	0.88	0.02	0.11	0.92
MSCI (1995:M1–2012:M8)						
Energy	0.34	3.25	**0.00**	0.32	2.99	**0.00**
Energy Equip. & Services	0.25	3.33	**0.00**	0.25	3.17	**0.00**
Oil & Gas	0.33	3.11	**0.00**	0.31	2.85	**0.00**
Transportation	0.22	1.39	0.16	0.20	1.28	0.20
Air Freight & Logistics	0.05	0.81	0.41	0.08	0.69	0.49
Airlines	0.17	1.91	0.06	0.15	1.63	0.10
Marine	0.22	1.93	**0.05**	0.22	1.87	0.06
Road & Rail	0.07	0.47	0.64	0.06	0.42	0.67

Table 4 (a): In-sample Predictability of Real Oil Prices

$$y_{t+h} = \alpha + \beta x_t + u_{t+h} \quad t = 2, ..., R - h$$

	WTI			Brent		
	$\hat{R}^2$	t-stat	p-value	$\hat{R}^2$	t-stat	p-value
Baseline (1984:M10–2012:M8)						
AMEX Oil Index	0.42	4.00	**0.00**	0.47	4.03	**0.00**
S&P 500	0.00	0.03	0.98	0.02	0.12	0.90
MSCI (1995:M1–2012:M8)						
Energy	0.26	2.41	**0.02**	0.34	3.15	**0.00**
Energy Equip. & Services	0.21	2.79	**0.01**	0.26	3.29	**0.00**
Oil & Gas	0.24	2.23	**0.03**	0.33	3.00	**0.00**
Transportation	0.14	0.89	0.37	0.24	1.50	0.13
Air Freight & Logistics	0.04	0.40	0.69	0.09	0.79	0.43
Airlines	0.09	1.00	0.32	0.18	2.01	**0.04**
Marine	0.20	1.71	0.09	0.23	1.96	**0.05**
Road & Rail	0.01	0.06	0.95	0.09	0.58	0.56

Table 4 (b): In-sample Predictability of Real Oil Prices

$$y_{t+h} = \alpha + \beta x_t + u_{t+h} \quad t = 2, ..., R - h$$

	Dubai			Average		
	$\hat{\beta}$	t-stat	p-value	$\hat{\beta}$	t-stat	p-value
Baseline (1984:M10–2012:M8)						
AMEX Oil Index	0.49	4.64	**0.00**	0.45	4.29	**0.00**
S&P 500	0.02	0.11	0.91	0.01	0.08	0.94
MSCI (1995:M1–2012:M8)						
Energy	0.32	3.22	**0.00**	0.30	2.95	**0.00**
Energy Equip. & Services	0.24	3.32	**0.00**	0.24	3.17	**0.00**
Oil & Gas	0.31	3.07	**0.00**	0.29	2.79	**0.01**
Transportation	0.20	1.37	0.17	0.19	1.25	0.21
Air Freight & Logistics	0.09	0.80	0.42	0.07	0.66	0.51
Airlines	0.16	1.91	0.06	0.14	1.63	0.10
Marine	0.22	1.96	**0.05**	0.21	1.90	0.06
Road & Rail	0.05	0.38	0.70	0.05	0.33	0.74

7. Out-of-sample Forecast Comparison

Table 5 (a): Out-of-sample forecast comparsion (Nominal Oil Prices)

$$\frac{MSPE(M^{SP})}{MSPE(M^{NC})}$$

	WTI		Brent		Dubai		Average	
Horizon	MSPE Ratio	Success Ratio	MSPE Ratio	Success Ratio	MSPE Ratio	Success Ratio	MSPE Ratio	Success Ratio
$h = 1$	**0.78**	**0.63**	**0.76**	**0.62**	**0.72**	**0.61**	**0.74**	**0.63**
3	**0.95**	0.57	**0.94**	0.57	**0.92**	0.60	**0.94**	0.57
6	1.02	0.54	1.01	0.56	1.00	0.55	1.01	0.55
9	1.04	0.53	1.01	0.55	1.01	0.56	1.02	0.54
12	1.04	0.53	0.99	0.54	0.98	0.58	1.00	0.56

The success ratio is constructed using the proportion of forecasts that correctly predict the sign of the change in the oil price.

Table 5 (b): Out-of-sample forecast comparsion (Real Oil Prices)

$$\frac{MSPE(RM^{SP})}{MSPE(RM^{NG})}$$

	WTI		Brent		Dubai		Average	
Horizon	MSPE Ratio	Success Ratio	MSPE Ratio	Success Ratio	MSPE Ratio	Success Ratio	MSPE Ratio	Success Ratio
$h = 1$	**0.80**	**0.62**	**0.79**	**0.65**	**0.74**	**0.63**	**0.77**	**0.63**
3	**0.95**	0.54	**0.94**	0.53	**0.93**	0.55	**0.94**	0.52
6	1.01	0.52	1.00	0.51	0.99	0.51	1.00	0.52
9	1.02	0.54	1.01	0.55	1.00	0.58	1.01	0.57
12	1.03	0.58	0.99	0.58	0.98	0.57	1.00	0.57

The success ratio is constructed using the proportion of forecasts that correctly predict the sign of the change in the oil price.

8. Conclusion

- Using the NYSE Arca (AMEX) oil index as a predictor, the one-month-ahead forecasts for nominal crude oil prices reduce the mean squared prediction error by between 22% (for the West Texas Intermediate oil price) and 28% (for the Dubai oil price).
- Moreover, the directional forecast based the AMEX oil index is significantly better than a 50:50 coin toss.