14

Predicting the Price of Oil

Naci H. Mocan and Ahmet E. Kocagil

INTRODUCTION

The first OPEC shock occurred in 1973–1974, after the oil embargo following the Arab-Israeli Yom Kippur War. In 1980 the price of oil jumped again following the Iranian revolution. After 1981, the price of oil started to decline, and in 1986 oil prices plunged due to the near-collapse of OPEC. In this chapter we will investigate whether this observed erratic behavior of the oil price over time can be explained by economic theory. If we find a reasonably positive answer to that question, we will then investigate if there exists the promise of sound predictability in the market.

EMPIRICAL IMPLEMENTATION

Hotelling's rule states that the price of an exhaustible resource should rise at the rate of interest. If the price of the resource rises faster than the prevailing rate of interest, the resource owner will extract all of it. If the price grows slower than the prevailing rate of return, the owner will have the incentive to discontinue production. Thus in equilibrium the price must rise at the rate of interest.

More formally, Hotelling's model suggests the following price path for an exhaustible resource:

$$P_t = P_o e^{rt} \tag{1}$$

where P is the price of the exhaustible resource, r is the interest rate, and t is the time index. Taking logarithms of both sides of (1) we get

Table 14.1
Cochrane–Orcutt Estimates for Equation (1b)

Time Period	Estimated Equation[a]	R2	D-W	RHO[b]
1951:2-1989:12	Ln P=2.21+DUMMIES+0.004 T (3.35) (2.6)	0.99	1.09	0.99
1951:2-1973:12	Ln P=2.31+DUMMIES+0.002 T (19.30) (3.03)	0.99	1.44	0.99
1974:1-1981:2	Ln P=-2.27+DUMMIES+0.016T (-2.20) (5.79)	0.99	1.24	0.96
1951:2-1981:2	Ln P=2.03+DUMMIES+0.006 T (2.70) (2.75)	0.99	1.05	0.99
1981:3-1986:1	Ln P=7.06+DUMMIES-0.006 T (0.30) (0.09)	0.99	1.23	0.96
1986:2-1989:12	Ln P=4.03+DUMMIES-0.00002T (0.02) (-0.03)	0.77	1.39	0.97
1981:3-1989:12	Ln P=7.52+DUMMIES-0.007 T (3.06) (-1.28)	0.97	1.02	0.98

a P stands for the price of oil, T is the linear trend term, DUMMIES represents the eleven monthly seasonals, Ln stands for the natural logarithm, Numbers in parentheses are the t-ratios.

b R2 and DW denote the regression R-squared and the Durbin-Watson statistic respectively. RHO is the estimated first-order autocorrelation coefficient of the residuals.

$$1nP_t = 1nP_o + rt \tag{1a}$$

and its empirical counterpart

$$1nP_t = \alpha + \beta t + u_t \tag{1b}$$

where u_t is a zero mean, constant variance, serially uncorrelated stochastic error term. For the price variable we used the price index for crude petroleum (1967 = 100).[1] The data consisted of monthly observations from February 1951 through December 1989. Since there were disruptions in the oil price behavior in 1973, 1981, and 1986, we also estimated the models for the 1951–1973, 1974–1981, 1981–1986, 1986–1989 periods, as well as for the 1951–1981 period, where there was a continuous increase in the price of oil, and 1981–1989, where the oil price exhibited a decline.

The results of the estimation of equation (1b) using the whole sample and six subsamples are displayed in table 14.1. Since the preliminary regression results revealed high autocorrelation of the errors, we applied the Cochrane–Orcutt procedure to correct for serial correlation.

The coefficients of the eleven monthly dummies, which are included to control for seasonal variation, are not reported.

As can be seen, we have a positive and significant trend term for the whole sample (1951–1989). The trend terms for the 1951–1981 period, as well as all subsamples prior to 1981, are also positive and significant. The regression results for the post-1981 samples, however, reveal that the trend terms become negative and insignificant. These results imply that equation (1a), however simple, has a reasonable predictive power, the price of oil is governed by a long-term positive trend, and the observed decline in the price of oil after 1981 is not a reversal of that trend in the statistical sense.

Note that by differencing equation (1a), we obtain

$$\ln P_{t-1} = \ln P_o + r(t-1) \tag{1c}$$

Subtracting (1c) from (1a) yields

$$\ln P_t - \ln P_{t-1} = r \tag{2}$$

According to equation (2) the growth rate (logarithmic difference) of the price should be equal to the rate of interest. That is, if observations on the growth rate of price are plotted against the corresponding observations on the rate of interest, the points should lie along a line with a slope of one. For the nominal rate of interest we used three-month commercial paper rates as reported by the board of governors of the Federal Reserve System. To obtain the real rate of interest, we adjusted the values of the commercial paper by the producer price index, industrial commodities. In this chapter, "rate of interest" will stand for the real rate. We graphed the 1951–1987, 1951–1973, 1974–1981, and 1981–1987 periods. The series in all cases have been smoothed with a four-year moving average to reduce random variations. The smoothed value for any observation is obtained by taking an average of the value for the twenty-four monthly observations preceding it and the twenty-four observations following it. The values for February 1951 observations, for example, are the averages of the observations between February 1949 and February 1953. Each graph displayed the 45-degree line, the plots of the growth rate of oil prices with respect to the real rate of interest for every month, and a line that is fitted to the observations.[2] The points did not lie on the 45-degree line. They diverged from the 45-degree line mostly in the 1951–1987 sample. This is because between 1951 and 1973 there was not much fluctuation in the price of oil, generating growth rates close to zero, whereas there were considerable fluctuations in the rate of interest. This was portrayed by the graph for the 1951–1973 period, where the fitted line had a very small slope when

Table 14.2
Cochrane–Orcutt Estimates for Equation (2a)

Time Period	Estimated Equation[a]	R2	D-W	RHO[b]
1951:2-1989:12	Ln P = 3.27+DUMMIES+0.01 R (5.06) (1.09)	0.99	1.08	0.99
1951:2-1973:12	Ln P = 2.63+DUMMIES+0.001 R (94.18) (0.40)	0.98	1.13	0.96
1974:1-1981:2	Ln P = 3.89+DUMMIES+0.016 R (25.73) (0.88)	0.98	1.00	0.95
1951:2-1981:2	Ln P = 3.23+DUMMIES+0.002 R (12.56) (0.39)	0.99	0.92	0.99
1981:3-1986:1	Ln P = 4.78+DUMMIES+0.09 R (48.91) (2.27)	0.97	0.88	0.91
1986:2-1989:12	Ln P = 4.98+DUMMIES+0.41 R (10.52) (2.27)	0.73	1.02	0.60
1981:3-1989:12	Ln P = 4.81+DUMMIES+0.18 R (12.86) (1.99)	0.97	1.11	0.99

a P stands for the price of oil, R is the real rate of interest, DUMMIES represents the eleven monthly seasonals, Ln stands for the natural logarithm, Numbers in parentheses are the t-ratios.

b R2 and DW denote the regression R-squared and the Durbin-Watson statistic respectively. RHO is the estimated first-order autocorrelation coefficient of the residuals.

the interest rate is plotted on the horizontal axis, and the rate of growth of oil price is plotted on the vertical axis. In periods 1974–1981 and 1981–1987 we observed that the fitted lines had slopes closer to one, but were located below the 45-degree line by a constant. In sum, even though the graphs did not perfectly agree with the outlook of the theory in the form of equation (2), they did not present a situation that is totally out of context.

To more formally test the empirical success of (2), we estimate

$$1nP_t = \alpha + \beta r_t + u_t \qquad (2a)$$

The results are reported in table 14.2. In this model, the rate of interest does not have explanatory power in the samples prior to 1981. For the 1981–1986, 1986–1989, and 1981–1989 periods, however, the *t*-ratios of the interest rate are greater than 1.96: interest rate explains the oil price well. Thus, the results depicted in table 14.2 imply that equation (2a) does have predictive power for the post-1981 period, and an extended version of it, such as a distributed lag model, is a good candidate for exploring the behavior of the oil price after 1981.

Although at first glance it seems that either equation (1b) or (2a) can be used as empirical counterparts of equation (1) and forecasts can be created using elaborated versions of them, in the next section we will show that empirical implementations based on either one would yield erroneous conclusions.

TIME-SERIES PROPERTIES OF THE OIL PRICE AND THE RATE OF INTEREST

Since Charles R. Nelson and Charles I. Plosser's influential article (1982), the importance of determining the trend properties of a time series employed in the empirical analysis is recognized. Some series may have deterministic time trends: the level of the series increases by some fixed amount each period. In this case trend can be eliminated by regressing the series on some polynomial of time trend. Some series, on the other hand, may be governed by stochastic trends: the level of the series increases each period by some fixed amount, but the rate of increase deviates from its average by some unforecastable random amount.

Denoting the natural logarithms of the time series by x_t and the deviations from trend by u_t, a series with a deterministic trend can be represented as

$$x_t = \alpha + \beta t + u_t \tag{3}$$

where t is the trend term and u_t is the error term with zero mean and constant variance. On the other hand, a series with a stochastic trend is generated by the model

$$x_t - x_{t-1} = \beta + \epsilon_t \tag{4}$$

where ϵ_t is a stationary series (one with time-independent mean and variance) with mean zero and constant variance. If the variable x_t is represented by (4) it is called *integrated* of order one (or integrated, for short). If the time series x_t is integrated, the first difference of it is stationary with mean β. Hence, differencing is the appropriate procedure for trend elimination.

A test developed by David A. Dickey and Wayne A. Fuller (1981) can be used to test the hypothesis that a particular series is governed by stochastic or deterministic trend. To implement the Dickey–Fuller test we estimate

$$\Delta x_t = \alpha_0 + \beta_0 t + \beta_1 x_{t-1} + \sum_{i=1}^{n} \delta_i \Delta x_{t-i} + \epsilon_t \tag{5}$$

where x is in natural logs and Δ is the difference operator. The variable x has a random growth component if $\beta_0 = \beta_1 = 0$.

After estimating equation (5) for the price of oil, we found that the *F*-statistic of the null hypothesis that $\beta_0 = \beta_1 = 0$ is 1.42 when $n = 1$, 1.38 when $n = 6$, and 1.57 when $n = 12$. The critical values tabulated by Dickey and Fuller are 6.34 when the sample size is 250, and 6.30 when the sample size is 500. Therefore, we could not reject the null hypothesis of a stochastic trend for the oil price variable. Consequently, its trend elimination should be done by taking the first differences. For the rate of interest, the *F*-statistics were 15.57, 7.53, and 6.94 for $n = 1$, $n = 6$, and $n = 12$, respectively. In the case of the rate of interest, we distinctly rejected the hypothesis that the series is governed by a stochastic trend.

It has been recognized that the usual techniques of regression analysis can result in highly misleading conclusions when the variables contain stochastic trends (Stock and Watson 1988). It can be shown that spurious regression results can occur, in the sense of overestimation of the *t* values (so that parameters that are actually insignificant from zero may appear to be significant) if the levels of variables are used in the model. This can be distinguished by low Durbin–Watson values, even though the R^2 values may seem to be satisfactory or even high (Granger 1989). Note that the results reported in tables 14.1 and 14.2 reveal high R^2, but low Durbin–Watson statistics, giving clues of spurious regressions.

When specifying regression models in time series, one has to make sure that the different variables are integrated to the same degree (Maddala 1988). Since the price of oil has a stochastic trend, equation (1b) is mis-specified, because it regresses the price of oil, which is integrated of order one, on a linear time trend that is not an integrated process.[3] Similarly, equation (2a) is also mis-specified since the rate of interest is not an integrated process.

Given these considerations, the correct empirical specification to investigate the behavior of the oil price in the context of equation (1) should be the following:

$$\Delta \mathrm{Ln}\, P_t = \alpha + \beta\, \mathrm{RES}_t + u_t \tag{6}$$

where RES_t is the residual of a regression of the interest rate on time.

In equation (6) the left-hand side is the first difference of the logarithm of the price of oil, hence it is stationary (integrated of order zero). The right-hand side variable is the trend deviations of the rate of inter-

Table 14.3
Cochrane–Orcutt Estimates for Equation (6)

Time Period	Estimated Equation[a]	R2	D-W	RHO[b]
1951:2-1989:12	ΔLn P = 0.007+DUMMIES+0.01 RES (1.25) (1.68)	0.23	1.95	0.45
1951:2-1973:12	ΔLn P = 0.005+DUMMIES+0.003 RES (2.40) (1.20)	0.09	1.99	0.25
1974:1-1981:2	ΔLn P = 0.01+DUMMIES+0.004 RES (0.88) (0.17)	0.21	1.98	0.37
1951:2-1981:2	ΔLn P = 0.006+DUMMIES+0.01 RES (1.44) (2.19)	0.22	1.97	0.44
1981:3-1986:1	ΔLn P = -0.005+DUMMIES-0.004 RES (-0.91) (-0.26)	0.37	1.86	0.37
1986:2-1989:12	ΔLn P = 0.03+DUMMIES-0.028 RES (0.68) (-0.15)	0.41	2.14	0.47
1981:3-1989:12	ΔLn P = 0.01+DUMMIES+0.01 RES (0.51) (0.22)	0.31	1.88	0.48

a P stands for the price of oil. RES is the residual of the regression of the rate of interest on a constant, a linear, and a quadratic term. Δ is the difference operator. DUMMIES represents the eleven monthly seasonals. Ln stands for the natural logarithm, Numbers in parentheses are the t-ratios.

b R2 and DW denote the regression R-squared and the Durbin-Watson statistic respectively. RHO is the estimated first-order autocorrelation coefficient of the residuals.

est, which is also integrated of order zero. Therefore, both the sides of the equation are of the same order, and equation (6) is well specified.

The results reported in table 14.3 indicate that this model is indeed better specified. The Durbin–Watson statistics are close to 2 in all samples, but the R-square values are considerably lower than the ones reported in tables 14.1 and 14.2. The interest rate coefficient is significant at the 2 percent level in the 1951–1981 sample, but it is not significant in the post-1981 samples. As a result, the *t*-statistic of the interest rate in the whole sample is 1.68, which is significant only at the 9 percent level. This seems to indicate that even though the versions (1b) and (2a) of the relationship between the rate of interest and the price of oil reveal a statistical link between them that can be further investigated and developed for forecasting purposes, the correct specification depicted in equation (6), which takes into account the time-series properties of the trend terms of the series, fails to capture a causal link from the rate of interest to the price of oil that spans the post-1981 period.

Before jumping to the conclusion that the dynamics of the interest rate series has no predictive power on the dynamics of the price of oil, we estimate a more sophisticated version of equation (6). We employ the vector-autoregression (VAR) methodology of Christopher A. Sims

(1980). Vector-autoregressions can be interpreted as the dynamic reduced forms that arise from dynamic stochastic structural models. In a VAR system, each variable is regressed on its own lagged values, lagged values of the other endogenous variables, and lagged values of the relevant exogenous variables. Thus, the VAR technique captures the dynamic inter-relations among variables. Moreover, consistent and efficient estimates can be obtained by ordinary least-squares since the right-hand-side variables are the same in each equation (Zellner 1962).

The VAR representation of model (6) consists of the following equations:

$$\Delta LnP_t = \alpha_1 + \sum_{i=1}^{k} \beta_{1i} \Delta LnP_{t-i} + \sum_{i=1}^{k} \gamma_{1i} RES_{t-i} + \epsilon_{1t} \qquad (7a)$$

$$RES_t = \alpha_2 + \sum_{i=1}^{k} \beta_{2i} \Delta LnP_{t-i} + \sum_{i=1}^{k} \gamma_{2i} RES_{t-i} + \epsilon_{2t} \qquad (7b)$$

According to equation (7a), the current value of the rate of growth of the oil price is explained by its own lagged values, as well as the lagged values of the trend deviations of the rate of interest. Similarly, equation (7b) states that the current value of the deviation of the interest rate from its trend is explained by its past behavior and the past behavior of the growth rate of the oil price. α_1 and α_2 represent the constant and the seasonal dummies, ϵ_{1t} and ϵ_{2t} are the white noise disturbance terms. By including a sufficient set of lags, the VAR technique eliminates the autocorrelation of the disturbances.

TESTS FOR CAUSALITY IN THE VAR SYSTEM

Once the VAR system has been estimated, causality tests can be performed. If the set of coefficients of a particular right-hand-side variable is significant, then that variable "causes" the dependent variable in the sense of C. W. J. Granger (1969). As an example, consider the following null hypothesis pertaining to equation (7a): $H_0: \gamma_{11} = \gamma_{12} = \ldots = \gamma_{1k} = 0$. Rejection of the null hypothesis H_0 would indicate causality from the interest rate to the price of oil (because it would indicate that the coefficients on *RES* in equation [7a] are significantly different from zero; that is, the inclusion of lagged interest rate variable does improve the prediction of the oil price).

Table 14.4 exhibits the estimation results of the VAR models. Since our primary concern is to explain the price of oil, we report only the coefficients of the lagged interest rate variable in the oil price equation (7a), and suppress the constant, seasonal dummies and the lagged values of the price of oil. The models are estimated with different lag

Table 14.4
Estimation Results of Equation (7A) of the VAR Model

	Coefficients of (RES) in the sample[a]						
Lag	1954-1989	1954-1973	1974-1981	1951-1981	1981-1986	1986-1989	1981-1989
1							0.161 (1.58)
2							-0.262 (-2.10)
3					-0.1217 (-1.20)		0.152 (1.51)
4					-0.0999 (-0.69)		
5					-0.0899 (-0.59)		
6	0.0021 (0.18)		-0.0195 (-0.64)	0.0081 (1.04)	-0.5107 -(2.8)	-0.193 (-0.33)	-0.174 (-1.91)
7	-0.0075 (-0.56)	-0.0016 (-0.41)		-0.0110 (-1.20)	0.1645 (1.36)		0.196 (1.75)
8	0.0031 (0.24)	-0.0029 (-0.69)		-0.0075 (-0.86)	0.0895 (0.88)		0.159 (1.69)
9			-0.0038 (-0.06)		0.1413 (1.43)	0.254 (0.51)	-0.216 (-2.21)
10	0.0142 (1.11)	0.0053 (1.20)	0.0976 (1.58)	0.0182 (2.06)	0.0602 (0.52)	0.700 (0.94)	0.076 (0.75)
11	0.0021 (0.16)	-0.0010 (-0.23)	0.0112 (0.16)	0.0107 (1.18)			-0.011 (-0.14)
12	-0.0019 (-0.17)	-0.0005 (-0.13)	0.0038 (0.06)	-0.0079 (-1.00)		-1.167 (-1.60)	
13					0.1711 (1.75)	-0.338 (-0.69)	-0.104 (-1.37)
14			-0.0843 (-1.50)		0.0870 (1.02)	1.676 (1.80)	0.221 (2.35)
15						-0.116 (-1.21)	
16		0.0011 (0.31)	0.0073 (0.13)			-0.5606 (-0.76)	-0.038 (-0.50)
18	-0.0022 (-0.21)	-0.0007 (-0.19)		-0.0076 (-1.0)	-0.1387 (-2.46)	0.3325 (0.52)	0.027 (0.63)
19	-0.0185 (-1.73)	-0.0031 (-0.81)		-0.0063 (-0.79)			
20			-0.1176 (-1.70)	-0.0031 (-0.41)		-0.9228 (-1.90)	
21			.0511 (.722)				
22			.0709 (1.09)			.0095 (.02)	
23	.0022 (.194)	.0052 (1.50)					
24	-.0035 (-.39)	-.0014		(-2.58)	-.1287	-.011 (.37)	
25		-.0110 (-2.98)					
26	-.0045 (-.39)		-.0249 (-.44)				
27						-.1024 (-.203)	
28		-.0013 (-.37)				1.2138 (1.798)	
29	-.0020 (-.18)	.0003 (.108)		-.0066 (-.90)		-1.126 (-1.73)	
30	.0065 (.572)		.0374 (.652)	.0077 (.999)	-.0178 (-.447)		
31							
32			.0019 (.028)	.0034 (.427)	.0164 (.227)	.8971 (.649)	
33	-.0130 (-1.10)		-.0885 (-1.50)	-.0149 (-1.70)	.2286 (3.12)	-.8833 (-1.41)	
34	-.0053 (-.51)	.0022 (.687)		-.0031 (-.39)			
36		-.0018 (-.57)	.0402 (.934)	.0004 (.065)			
F-stat	1.75 (.04)	1.56 (.08)	1.66 (.09)	1.91 (.09)	1.92 (.09)	3.97 (.09)	1.65 (.08)
R2	.15	.21	.53	.22	.75	.95	.56

a The numbers in parentheses below the coefficients are the t-values
The numbers in parentheses below the F-statistics are the marginal significance levels.

specifications for each sample. Each subset is estimated with fifteen lags. In all samples the F-statistics of lagged interest rate variable are significant at the 9 percent level or better. Thus, in the VAR specifications we sharply reject the hypotheses that the lagged values of the rate of interest do not increase the prediction of the price of oil. Put differently, we conclude that interest rate Granger-causes the price of oil.

If the VAR models reported in table 14.4 are well specified, the residuals should be white noise. The Durbin–Watson statistic is not appropriate in this case. It is biased toward a finding of no serial correlation in errors when the model contains lagged dependent variables. Therefore we analyzed the correlograms of the residual series and the ARIMA structure of their first differences. In all cases the residuals appeared to be white noise. Also, we applied a Lagrange multiplier test to the errors, where the residuals from each VAR specification were regressed on the same set of regressors and a set of lagged residuals and the significance of the lagged residuals are checked. Again, we found no error structure different from white noise.[4]

PROJECTIONS USING VECTOR-AUTOREGRESSIONS

We used estimated VAR models displayed in table 14.4 for projecting the price of oil. It was seen that the actual price of oil between January 1989 and December 1989, and the projected price of oil for the same period obtained from the VAR model displayed in column 6 of table 14.4 were very close to each other.[5] Similarly, to generate the forecasts for the period March 1985 to February 1986, we used the VAR model estimated using sample 1981–1986 depicted in column 5 of table 14.4. Once again the estimated VAR model was seen to yield highly reliable in-sample projections. When we used the same model to create out-of-sample forecasts, however, the projections deviated significantly from the actual price of oil. The forecasts for the period March 1986 to February 1987 based on the model estimated using the data from 1981 to 1986 were consistently below the actual prices. The model increasingly underestimated the price of oil. Similarly, we calculated the in-sample and out-of-sample forecasts of the oil price using the VAR model estimated within the period January 1974 through February 1981. Once again, we had very accurate within-sample, but imprecise out-of-sample forecasts. We reached the same conclusion in projections based on the model estimated between 1954 and 1973. In sum, even though the vector-autoregressive specifications depict the dynamics of the oil price accurately and generate reliable within-sample forecasts for the 1954–1973, 1974–1981, 1981–1986, and 1986–1989 periods, they are not helpful in predicting out-of-sample values.

If we were not able to capture the dynamics of the system, we would observe signs of model mis-specifications in terms of non-white noise errors, and we would not be able to create accurate in-sample projections. Given that this is not the case, the most likely reason for the inaccurate out-of-sample projections is the possibility of having structural changes in 1973, 1981, and 1986. The relationship between the price of oil and its past values as well as past values of the interest rate might have been altered in 1973, 1981, and 1986, invalidating the coefficients calculated using previous data generating structure and hence making it impossible to generate credible projections.

TESTING FOR STRUCTURAL CHANGE

In this section we will test the hypothesis that the dynamics of the oil price is governed by one structure. Rejection of the hypothesis would mean that the structural relationship among variables is being altered over time, thus the poor forecasting performance of the models can be attributed to these structural breaks. To perform the tests in the VAR context, we employed the following algorithm. We took the VAR model estimated between 1954 and 1989 as our basic model. This system can be called a "restricted system," as it does not allow the estimated coefficients to change over time. We estimate eighty-four coefficients in this VAR system using data from 1954 to 1989 and assume that these coefficients remain intact over time.[6] On the other hand, estimating two VAR models, say, one for 1954–1973 and one for 1974–1989, allows the coefficients to vary between 1954 and 1989; hence there are no restrictions on the coefficients, and that system can be termed an "unrestricted system." If the observed volatility of the price of oil is only a random variation around a given structure, then the residuals obtained from the restricted system should not be different from the residuals of the unrestricted system. Since we use VAR models, we should compare the error properties of the systems as a whole. Therefore, after estimating the basic model as presented in column 1 of table 14.4, we obtained the variance-covariance matrix of the residuals. We then estimated VAR systems for 1954–1973 and 1974–1989 periods, created error series for the price of oil and the rate of interest, and obtained their variance-covariance matrix. To test whether a structural change took place in 1973, we compared the systemwide variance-covariance matrices of the restricted and unrestricted models. Along the same lines, we tested structural changes in 1981, 1986, as well as 1973 and 1981, and 1981 and 1989. We could not accept the hypothesis that there were no structural changes in the system between 1954 and 1989. Rather, we found that structural changes took place in December 1973, February 1981, and February 1986. The chi-square values were 151.49, 201.28, and 366.53 for December 1973, February 1981, and February 1986, respectively.[7] We also tested for two concurrent structural changes during the 1954–1989 period. The data revealed that there were structural breaks in December 1973 and February 1981, and in February 1981 and February 1986. The chi-square statistics were 372.67 and 531.77, respectively. Based on the VAR model estimated between 1981 and 1989, we tested and rejected the hypothesis of no structural change in 1986; there was not *one* structure prevailing between 1981 and 1989 (the chi-square was 264.11). In sum, regardless of how we constructed the test, we distinctly rejected the hypothesis of no structural change. The data revealed that in December 1973, February 1981, and February 1986 structural changes took place that altered the whole dynamics of the system. Note that the results were not sensitive to the choice of a particular month. In fact, they were extremely robust.

AN INVENTORY MODEL

There exists the common belief that the oil consumers and companies increase their inventories and stocks in anticipation of price increases. An observed increase in the price of oil due to a shift of the supply curve for oil to the left triggers a shift in the demand curve to the right, which yields further price increases. This idea makes the price of oil a function of the level of inventories along with the rate of interest. Obviously, in a system that consists of oil price, rate of interest, and inventories, there are mutual causalities among variables. For instance, while the price of oil depends on the rate of interest and inventories, inventories are a function of the oil price and the rate of interest as well. Therefore, once again vector-autoregressions are the best means of capturing these dynamic inter-relationships among the variables.

As a measure of oil inventories, we used end-of-month real petroleum inventories reported by the U.S. Department of Commerce, Bureau of Economic Analysis. The range of variable, which will be called INV in the rest of the text, is 1967 to 1989. To check the trend structure of the inventory variable, we estimated equation (5) for INV and found that the *F*-statistic for the null hypothesis of a stochastic trend was 2.20 when $n=1$, 1.96 when $n=6$, and 1.49 when $n=12$. Thus we accepted the hypothesis that the inventory variable is governed by a variable trend. Note that both the oil price and the inventory are integrated variables, hence we will employ them in differenced form. If there is a long-run equilibrium relationship between our two integrated variables P and INV, then the following must hold: $\phi \text{ P} + \lambda \text{ INV} = \omega$, where ϕ and λ are the coefficients of the price and inventories, respectively, and ω is the equilibrium error that would seldomly deviate from zero and should be stationary. In this case the integrated variables are sharing a common stochastic trend, and they are called co-integrated. Robert F. Engle and Clive William John Granger (1987) have shown that vector-autoregressions in differences are mis-specified if the variables are co-integrated. To see if P and INV are co-integrated, we estimated

$$\mathrm{LnP}_t = \alpha + \beta\, \mathrm{LnINV}_t + \omega_t \tag{8}$$

The residuals ω_t are then used in

$$\Delta\omega_t = \eta\, \omega_{t-1} + \nu_t \tag{8a}$$

and

$$\Delta\omega_t = \xi\omega_{t-1} + \sum_{i=1}^{4} \delta\omega_{t-i} + \nu_t \tag{8b}$$

where ν_t is the white noise error term. We test the null hypothesis of $\eta=0$ in (8a) and $\xi=0$ in (8b). Accepting the null hypothesis would mean that P and INV are not co-integrated. We found that the coefficient η was .051 with a *t*-statistic of -2.55 in equation (8a) and ξ was $-.097$ with a *t*-statistic of -4.54 in equation (8b). The critical *t*-values tabulated by Robert F. Engle and Byung Sam Yoo (1987) are 3.37 and 3.25 for equations (8a) and (8b), respectively, at 5 percent level when the sample size is 200. Hence we concluded that the price of oil and inventories are not co-integrated. As a result, the VAR model is estimated in the following form:

$$\Delta\mathrm{LnP}_t = \alpha_1 + \sum_{i=1}^{k} \beta_{1i}\Delta\mathrm{LnP}_{t-i} + \sum_{i=1}^{k} \gamma_{1i}\mathrm{RES}_{t-i} + \sum_{i=1}^{k} \delta_{1i}\Delta\mathrm{LnINV}_{t-i} + \epsilon_{1t} \tag{9a}$$

$$\mathrm{RES}_t = \alpha_2 + \sum_{i=1}^{k} \beta_{2i}\Delta\mathrm{LnP}_{t-i} + \sum_{i=1}^{k} \gamma_{2i}\mathrm{RES}_{t-i} + \sum_{i=1}^{k} \delta_{2i}\Delta\mathrm{LnINV}_{t-i} + \epsilon_{2t} \tag{9b}$$

$$\Delta\mathrm{LnINV}_t = \alpha_3 + \sum_{i=1}^{k} \beta_{3i}\Delta\mathrm{LnP}_{t-i} + \sum_{i=1}^{k} \gamma_{1i}\mathrm{RES}_{t-i} + \sum_{i=1}^{k} \delta_{3i}\Delta\mathrm{INV}_{t-i} + \epsilon_{1t} \tag{9c}$$

Tables 14.5 and 14.6 report the coefficients of inventory and interest rate variables in the oil price equation (9a). Note that since the models are estimated with different lag lengths, the starting points of the samples may differ. As can be seen, in this VAR specification the R^2 values are higher than the VAR specification which consisted only of the oil price and the interest rate. Note also that all the *F*-statistics are significant, demonstrating that in all samples the past values of the interest rate and the inventory changes help explain the dynamics of the oil

Table 14.5
Estimation Results of the VAR Model with Inventory Variable

	Period							
Lags	1969-1989		1968-1973		1974-1981		1970-1981	
	RES	INV	RES	INV	RES	INV	RES	INV
1			-.035	.529				
			(-.83)	(2.57)				
2			-.0004	.919				
			(-.01)	(5.35)				
3			.055	.028				
			(1.78)	(.15)				
4			-.012	-.681				
			(-.24)	(-3.21)				
5			-.049	-.584				
			(-.87)	(-3.80)				
6	.0139	.3051	.048	.113	-.064	-.094		
	(.47)	(1.76)	(1.53)	(.68)	(-1.88)	(.31)		
7	-.0237	.1281	.021	.183				
	(-.58)	(.74)	(.48)	(1.09)				
8	.0126	-.2538	-.049	.523			-.037	-.621
	(.31)	(-1.47)	(-1.35)	(1.74)			(-1.43)	(-3.07)
9	-.0109	.5305	-.017	.703	.020	1.301	-.021	.372
	(-.23)	(3.06)	(-.47)	(1.68)	(.34)	(3.54)	(-.58)	(1.84)
10	.0539	.1608	.058	-.608	.059	-.149	.047	.085
	(1.30)	(.93)	(1.81)	(-2.17)	(.97)	(-.47)	(1.35)	(.46)
11	-.0082	.2111			.047	.393	.058	.295
	(-1.97)	(1.18)			(.68)	(1.23)	(1.49)	(1.45)
12	-.0079	-.0295	-.069	.083	.023	-.985	-.0003	-.282
	(-.24)	(-.16)	(-2.31)	(.37)	(.40)	(-3.18)	(-.008)	(-1.49)
13			.003	.400				
			(.10)	(1.83)				
14					-.155	.350	-.050	-.179
					(-2.77)	(.82)	(-1.79)	(-.90)
15			.085	-.033				
			(3.76)	(-.21)				
16			-.044	-.801	.097	-.012		
			(-1.01)	(-3.77)	(1.39)	(-.04)		
17			-.025	-.836				
			(-.58)	(-3.99)				
18	.0194	.2581						
	(.60)	(1.41)						
19	-.0592	.1425						
	(-1.46)	(.76)						
20	-.0004	.0704			-.197	.225	-.034	-.404
	(-.01)	(.38)			(-2.59)	(.63)	(-1.04)	(-1.97)
21					.109	-.051	.034	-.198
					(1.71)	(-.13)	(1.04)	(-.98)
22					.153	-.156		
					(2.35)	(-.43)		
25	-.023	-.186						
	(-.72)	(-1.03)						
26	.010	.111			-.127	.939	-.023	.415
	(.33)	(.61)			(-2.32)	(2.68)	(1.16)	(2.14)
30					.088	.420		
					(1.48)	(1.29)		
32					.045	.533		
					(.76)	(1.71)		
33					-.103	.251		
					(-1.89)	(.70)		
36					-.021	-.179	-.013	-.140
					(-.49)	(-.54)	(-.84)	(-.71)
F-stat	1.71	1.82	2.23	6.12	2.44	2.22	2.19	1.98
	(.07)	(.05)	(.11)	(.005)	(.02)	(.03)	(.03)	(.04)
R2	.27		.95		.78		.46	

a The double columns under each period present the coefficients of the interest rate and inventory variable in the oil price equation represented by (9a).
Other coefficients are suppressed for clarity of presentation. The numbers in parantheses below the F-statistics are the marginal significance levels.

Table 14.6
Estimation Results of the VAR Model with Inventory Variable

	PERIOD					
Lags	1981-1986		1986-1989		1981-1989	
	RES	INV	RES	INV	RES	INV
1	-.114	.707			-.031	.294
	(-1.64)	(2.64)			(-.32)	(1.00)
2	-.060	.323			-.144	-.063
	(-.50)	(1.06)			(1.18)	(1.23)
3	-.156	-.486	-.455	-1.012	.132	-.615
	(-1.19)	(-1.64)	(-2.08)	(-1.79)	(1.51)	(-2.18)
4	.200	-.517	.422	-1.271		
	(1.73)	(-1.84)	(1.68)	(-2.50)		
6	-.019	.316				
	(-.27)	(1.01)				
8					.204	-.146
					(2.71)	(.53)
9	.030	.569	.40	2.649	-.250	1.011
	(.30)	(1.40)	(2.88)	(4.35)	(-2.71)	(3.60)
10	-.036	.182			.119	-.235
	(-.40)	(.62)			(1.57)	(.76)
12	.071	.288	-.587	.726	-.017	-.108
	(1.09)	(1.16)	(-2.62)	(1.18)	(-.23)	(.35)
13			-.110	-1.139	-.126	.628
			(-.51)	(-1.58)	(-1.44)	(2.18)
14	.016	.955			.130	.236
	(.36)	(3.21)			(1.90)	(.74)
16	-.124	-.218	.359	.548	-.160	-.031
	(-2.48)	(-.58)	(1.75)	(.79)	(-2.21)	(-.11)
17					.104	.027
					(1.08)	(.09)
18	-.096	.735			-.016	.342
	(-1.96)	(2.46)			(-.02)	(1.16)
24	.037	-.547				
	(.77)	(-1.69)				
F-stat	2.25	3.42	3.19	4.88	1.68	2.36
	(.09)	(.02)	(.03)	(.005)	(.10)	(.02)
R2	.91		.82		.71	

a The double columns under each period present the coefficients of the interest rate and inventory variable in the oil price equation represented by (9a).
Other coefficients are suppressed for clarity of presentation. The numbers in parentheses below the F-statistics are the marginal significance levels.

price. There is significant Granger-causality from interest rate and inventory changes to the price of oil.

We obtained the in-sample and out-of-sample forecasts of the VAR models with the inventory variable presented in tables 14.5 and 14.6. In all cases, the predicted and actual values were very close for in-sample projections. When the same models were used to create out-of-sample projections, however, forecasts deviated from the actual price of oil as was the case before.

We then performed the structural change tests as described earlier in the text. As before, we rejected the hypothesis of no structural change in 1973, 1981, 1986, 1973 and 1981, and 1981 and 1986. The calculated chi-square statistics were 616.03, 321.95, 245.71, 576.24, and 591.57, respectively, which were all significant at the one percent level or better.

CONCLUSION

Considering the time-series properties of the variables and using vector-autoregressions to control for mutual causalities, we have demonstrated that a system between the price of oil and interest rate is well defined and has explanatory power. Adding the inventory variable to this system increases the system's explanatory power even further. The interest rate and inventory variables strongly Granger-cause the price of oil. The lagged values of the interest rate and inventories explain the current value of the oil price. Consequently, the estimated models generate very accurate in-sample projections of the oil price. The same models fail when used to create out-of-sample projections. We have shown that this is due to the fact that the structural relationships between the price of oil and other variables have been altered over time. More precisely, we have found that there is *not* one structure governing the behavior of the price of oil over time, rather the structure is changed in 1973, in 1981, and in 1986. This implies that the observed erratic behavior of the price of oil cannot be attributed to random fluctuations around a given structure. Rather, interruptions in the market created, we believe, mainly by OPEC, yield permanent structural changes.

To the extent that the oil market is exposed to disturbances by OPEC influenced by political as well as economic factors, predicting the price of oil seems a very difficult task, even though the vector-autoregressive model used in this study generates extremely accurate projections of the price within the intervention periods.

NOTES

1. All variables used in this study are obtained from *Citibase*.
2. To conserve space we chose not to display the graphs in the text. They are, however, available from the authors upon request.
3. Charles R. Nelson and Hee Joon Kang (1981) analyzed theoretically a similar regression. In our case, as they reported, the residuals of equation (1b) showed distinct cyclical behavior, rather than being a white noise.
4. A detailed description of these tests can be found in Mocan (1990) and Mocan and Baytas (1991).
5. The graphs of the projections, which are not displayed, are available upon request.

6. The price of oil and the interest rate enter with fifteen lags; plus constant and eleven seasonal dummies which adds up to forty-two coefficients in each equation.

7. $n(\ln|A| - \ln|B|)$ is distributed chi-square, where $\ln|A|$ is the ln determinant of the variance-covariance matrix of the residuals of the restricted system, and $\ln|B|$ is the ln determinant of the variance-covariance matrix of the residuals obtained from the unrestricted system. The degrees of freedom is the number of restrictions, that is, the number of coefficients in the restricted system.