OIL PRODUCTION IN THE LOWER 48 STATES: ECONOMIC, GEOLOGICAL, AND INSTITUTIONAL DETERMINANTS

Professor Robert K. Kaufmann
Center for Energy & Environmental Studies
Boston University
675 Commonwealth Avenue
Boston, MA 02215
kaufmann@bu.edu
617-353-3940

Professor J. Cleveland
Center for Energy & Environmental Studies
Boston University
675 Commonwealth Avenue
Boston, MA 02215
cutler@bu.edu
617-353-3083

ABSTRACT

In this paper, we establish an empirical model for oil production in the lower 48 states that represents its economic, physical, and institutional determinants. We estimate a vector error correction model for oil production in the lower 48 states that specifies real oil prices, average production costs, and prorationing by the Texas Railroad Commission. These modifications enable us to generate a model that accounts for most of the variation in oil production in the lower 48 states between 1938 and 1991. The inconsistencies between the historical record and the assumptions that underlie economic and Hubbert models imply that these models cannot be used to forecast future supplies. Furthermore, these inconsistencies indicate that the ability of a competitive market to make a smooth transition from oil to alternative energy sources may be less effective than currently believed.

I Introduction

Real oil prices have been relatively low for more than a decade, yet there has been a resurgence in concerns about supply (Kerr, 1998). Mirroring the debates of the 1960s and 1970s, attention is focused on extreme positions. On the optimistic side, Fisher (cited in Kerr, 1998) argues that global oil production won't peak for another 30 or 40 years. At the other end of the spectrum, Campbell and Laherrere (1998) forecast that global production will peak and turn down within five years.

There are many approaches to modeling oil supply, but two have received considerable development and application. The first is the standard economic model for the depletion of an exhaustible resource, the so-called Hotelling model (Hotelling, 1931). The Hotelling model, and its derivatives, assumes that producers rank fields by their cost of extraction and recover them in order of increasing costs, and that this profit maximizing behavior generates a consistent (and optimal) path for prices and production. This is the standard model described in all leading economic texts (e.g., Pearce and Turner, 1990; Tietenberg, 2000.).

The second approach is the so-called Hubbert curve, named after the petroleum geologist M. King Hubbert (1962). The Hubbert model assumes that cumulative production can be described by a logistic function. Under these conditions, the annual rate of production is defined by the first difference of the logistic curve, which traces a smooth, symmetric bell-shaped curve over time. The Hubbert curve and its variants are popular among physical scientists (Campbell and Laherrere, 1998).

The reliability of the economic and Hubbert models is undermined by restrictive assumptions that often conflict with the actual experience of the oil industry. The Hotelling model is not consistent with many aspects of the historical record of the US oil industry (Watkins, 1992). The assumption that firms extract resources starting with the highest quality deposits is contradicted by the empirical record of exploration and development. Oil exploration and discovery in the US was less successful than a random search until the 1930's (Menard and Sharman, 1975). Since the

1930’s, the effort required to produce a barrel of oil declined steadily until the 1960's and rose thereafter (Cleveland, 1991; 1993). The assumption that profit maximizing behavior generates a consistent and optimal path for prices and production is undermined by the seeming contradiction between production and real oil prices. Production in the lower 48 states doubled between 1945 and 1970 despite a steady decline in real oil prices. Between 1973 and 1985, real oil prices nearly tripled, but production declined nearly 20 percent.

Proponents of the Hubbert curve are heartened by the fact that the seeming contradiction between prices and production is consistent with the rising and falling portions of the bell-shaped curve. But the Hubbert model is undermined by the assumption that prices and other economic variables do not affect production, an error that causes some glaring mismatches between the bell-shaped curve and the historical record. For example, production in the lower 48 states stabilizes in the late 1970’s and early 1980’s, which contradicts the steady decline forecast by the Hubbert model. Several analysts have modified the Hubbert model to include economic and institutional determinants of production and have changed the estimation procedures to improve its statistical properties (Uri, 1980; Kaufmann, 1991; Pesaran and Samiei, 1995; Moroney and Berg, 1999). Nonetheless, a significant fraction of the revised models’ explanatory power is generated by the bell-shaped curve, which has no explicit connection to real world variables.

In this paper, we relax the restrictive assumptions that underlie economic and Hubbert models of production and estimate an empirical model for oil production in the lower 48 states that represents its economic, physical, and institutional determinants. The model includes a variable that represents prorationing decisions by the Texas Railroad Commission, which relaxes the assumption that oil producers operate in a competitive market. The model includes a variable that represents the cost of production, which relaxes the assumption that firms rank and produce their fields in order of decreasing quality. We relax Hubbert model’s assumption that prices play no role in production by including a price variable. This price variable is decomposed in a way that differentiates the effect of price increases from price decreases, which relaxes the assumption of perfect price reversibility that underlies economic models. These modifications enable us to generate a model that accounts for most of the variation in oil production in the lower 48 states between 1938 and 1991. The inconstancies between the historical record and the assumptions that underlie economic and Hubbert models imply that these models cannot be used to forecast future supplies. Furthermore, the models' inconsistencies indicate that the ability of a competitive market to make a smooth transition from oil to alternative energy sources may be less effective than currently believed.

II Methodology

We estimate a model for oil production in the lower 48 states that specifies real oil prices, average production costs, and prorationing by the Texas Railroad Commission as explanatory variables. Univariate tests developed by Dickey and Fuller (1979) indicate that these data are nonstationary (Table 1). Each series contains a stochastic trend that can be removed only after the data are differenced once (or perhaps twice for the prorationing variable). Such nonstationary data are said to be integrated order 1, or I(1). This property is critical to the choice of technique used to estimate the model. Granger and Newbold (1974) find that the diagnostic statistics generated by ordinary least squares indicate a meaningful relation among unrelated I(1) variables more often than implied by random chance. These relations are termed spurious regressions.

To differentiate between spurious and meaningful regressions, a large part of the recent econometric time series literature deals with estimation and inference in the presence of I(1) variables. Engle and Granger (1987) define the notion of cointegration: a property displayed by variables that share a common stochastic trend. To determine whether two or more variables cointegrate, Johansen (1988) and Johansen and Juselius (1990) describe a full information likelihood procedure based on the principle that variables which share the same stochastic trend can be combined linearly in a way that eliminates the stochastic trend. The coefficients that generate a stationary combination of nonstationary variables is termed a cointegrating vector (CV).

The procedures to estimate the cointegrating vector(s) are derived from a vector autoregression (VAR) in levels, which can be represented by equation 1:

$$x_t \quad = \quad A_1 y_{t-1} + \ldots + A_k y_{t-k} + \mu + \delta t + \Phi D_t + \varepsilon_t \qquad (1)$$

in which y is a vector of p variables whose behavior is being modeled, k is the number of lags, the A's and Φ are matrices of regression coefficients, μ and δ are a vector of constants, and ε_t is a vector of error terms each of which is niid. Some variables may be weakly exogenous and therefore excluded from the LHS. As a result, the dimensions of the vector x_t may not be identical with y_t. D_t is a vector of dummy variables and also possibly some stochastic variables that are found to be weakly exogenous.

To identify groups of variables that constitute a cointegrating relation from a vector of stochastic variables (y) which contain two or more I(1) variables, the cointegration procedure specifies the VAR as a vector error correction model (VECM):

$$\Delta x_t \quad = \quad \Gamma_1 \Delta y_{t-1} + \ldots + \Gamma_{k-1} \Delta y_{t-k+1} + \Pi y_{t-1} + \mu + \Phi \Delta D_t + \varepsilon_t \qquad (2)$$

Equation 2 specifies the first difference of the nonstationary variables, which is stationary, as a function of linear lagged values of the first difference of the nonstationary variables, which also are stationary, and linear combinations of the nonstationary variables, which are termed the cointegrating relations.

The term $\Pi\ y_{t-1}$ is the error correction mechanism (ECM). If there are one or more CV's, the ECM can be reformulated using equation 3:

$$\Pi\ y_{t-1} = \alpha\beta'[1,\ t, y_{t-1}\ ']' \qquad (3)$$

The term $[1, t, y_{t-1}\ ']'$ indicates that a constant and/or a deterministic trend may be included in the cointegrating relation. The deterministic trend is intended to capture the behavior of trend stationary variables, i.e. variables that are stationary after removal of a deterministic trend rather than after differencing (Hansen and Juselius, 1995). β' is the matrix of coefficients that creates a stationary combination of nonstationary variables, i.e. the cointegration relation. This linear combination of variables is termed a cointegrating relation and represents the long-run equilibrium relation among the variables included in the cointegrating relation. α is a matrix of coefficients that indicate whether a cointegrating relation (or ECM) affects a particular dependent variable. The number of cointegrating vectors, the variables that make-up a cointegrating vector, the coefficients associated with these variables, and the relation between an ECM and the dependent variables all are evaluated from diagnostic statistics that are generated by the estimation process.

We specify a VECM (equation 2) to estimate the relation among oil production in the lower 48 states (Prod), the average real cost of oil production (Cost), real oil prices (Poil), and the fraction of capacity allowed to operate by the Texas Railroad Commissions (Ration) between 1938 and 1991. The sample period is determined the availability of data used to calculate average cost, which are not available before 1936 or after 1991 (Cleveland, 1991).

Previous studies of the oil industry hint at an asymmetric relation between real oil prices and production (Frankel, 1946). To evaluate this possibility, we use the method developed by Gately (1992) to decompose the real price data into three series. One series is the maximum real price of oil between 1936 and any given date (P_{max}). This series increases monotonically because P_{max} remains at the previous all-time high when real prices fall below this maximum (Figure 1). Real prices often decline following a new all-time high. Such declines are termed price cuts. These prices cuts are accumulated to form a second series that decreases monotonically, which is termed P_{cut}. Following such declines, real prices may rise towards, but not exceed, the previous maximum. These price recoveries are accumulated to form a third series that increases monotonically, which is termed P_{rec}. The sum of P_{max}, P_{cut}, and P_{rec} at any time is the real price of oil.

We also include a variable (Ration) that represents the fraction of capacity allowed to operate by the Texas Railroad Commission (TRC). Kaufmann (1991) was the first to quantify the effects of prorationing on production. This relationship was subsequently confirmed by Pesaran and Sameie

(1995) and Moroney and Berg (1999). The TRC opened and closed capacity between 1935 and 1973 to damp the boom and bust cycle in oil production and prices (Prindle, 1981). By the early 1960's, the TRC allowed owners to operate their wells at less than 30 percent of capacity. Texas was the largest producing state during the sample period, and the effect of prorationing decisions by the TRC was amplified by other state commissions that followed the TRC's lead. Data for the fraction of operable capacity allowed to operate between 1938 and 1986 in Texas are obtained from the Texas Railroad Commission. Since 1978, the TRC allowed owners to operate at capacity and we extrapolate this value (1.00) through 1991.

III Results

We analyze a VECM that contains a time trend in cointegration space. We choose this specification based on a method developed by Pantula (1989). This specification also is the least restrictive, which implies we have not imposed any arbitrary restrictions on the VECM *a priori*. We choose the number of lagged first differences to include in equation 2, which is given by k - 1, by comparing results generated by VECMs that have 1- 4 lags and are estimated over the same period, 1941 - 1991. The performance of each VECM is evaluated using the Schwartz and Hannon Quinn information criterion (Hansen and Juselius, 1995). These tests evaluate an increase in the lag length as measured by the decrease in the determinant of the residual covariance matrix of the model while making an adjustment for the decrease in the degrees of freedom. The Schwartz criterion indicates that the shortest lag length 1 cannot be rejected, while the Hannon Quinn criterion indicates that the longest lag length 4 is appropriate. We choose the shorter lag length (1), which is indicated by the Schwartz criterion, consistent with the statistical preference for the most parsimonious model and the need to conserve degrees of freedom in the relatively short sample period.

The λ_{trace} and λ_{max} statistics are used to choose the rank of Π in equation 2, which corresponds to the number of cointegrating vectors (Johansen, 1988, 1991; Johansen and Juselius, 1990). The λ_{trace} statistic tests the null hypothesis that the number of cointegrating vectors is less than or equal to r against a general alternative that the number of cointegrating vectors is greater than r. The λ_{max} statistic tests the null hypothesis that the number of cointegrating vectors is r against the alternative r+1 cointegrating vectors. Both the λ_{trace} and λ_{max} statistics indicate that assigning Π a rank less than three is rejected strongly (Table 2). Based on this result, we estimate a model that has three cointegrating vectors.

Identifying the long run structure of the VECM requires several types of restrictions (Pesaran and Shin, 1994). Greenslade *et al.*, (1997) suggest that the efficiency of identifying the long run structure can be improved by first attempting to reduce the number of endogenous variables. We evaluate whether variables can be excluded from the x vector in equation 2 by testing a restriction that makes all elements of α associated with an individual variable equal to zero. Results indicate that we can reject the null hypothesis that Prod ($\chi^2(3) = 21.43$, $p < 0.00$), P_{cut} ($\chi^2(3) = 25.36$, $p < 0.00$), P_{rec} ($\chi^2(3) = 22.25$, $p < 0.00$), P_{max}, ($\chi^2(3) = 11.02$, $p < 0.01$) or Cost ($\chi^2(3) = 16.90$, $p < 0.00$) are exogenous. We cannot reject the null hypothesis that Ration ($\chi^2(3) = 4.98$, $p < 0.27$) is exogenous. Based on these results, we exclude Ration from the x vector. With Ration exogenous, repeating the tests for lag length and the rank of Π does not change the choices described above.

Excluding Ration from the x vector does not ease the process of identifying the cointegrating relations greatly. Just identifying restrictions eliminate one variable from each cointegrating relation. Because there are six variables, there are 20 sets of restrictions that would just identify the system. Because there are three cointegrating relations, tests that eliminate variables from cointegration space are not definitive. Finally, there is no *a priori* theory regarding which variables constitute the cointegrating relations. As a result, we cannot use a mechanistic process to identify the most parsimonious model of the relation among oil prices, the decomposed price series, average cost of production, and pro-rationing decisions by the TRC. Instead, we use *a priori* theory as a guide to identify model 1, which has the following form for the equation for the first difference in production (the first equation in the VECM):

$$\Delta \mathrm{Pr}od_t = \tilde{\alpha}_{11}(\mathrm{Pr}od_{t-1} + \frac{\beta_{21}}{\beta_{11}} Pcut_{t-1} + \frac{\beta_{31}}{\beta_{11}} + \mathrm{Pr}\,ec_{t-1} + \frac{\beta_{41}}{\beta_{11}} P\,\mathrm{max}_{t-1}$$

$$+\tilde{\alpha}_{12}\,(\frac{\beta_{52}}{\beta_{11}}\,\mathrm{Cost}_{t-1} + \mathrm{Ration}_{t-1})$$

$$+\tilde{\alpha}_{13}\,(\frac{\beta_{23}}{\beta_{53}}\,Pcut_{t-1} + \frac{\beta_{33}}{\beta_{53}}\,\mathrm{Pr}ec_{t-1} + \frac{\beta_{43}}{\beta_{53}}\,P\,\mathrm{max}_{t-11} + Cost_{t-1)}) + \varepsilon_{1+t}$$

in which $\tilde{\alpha}_{11} = \alpha_{11}/\beta_{11}$, $\tilde{\alpha}_{12} = \alpha_{12}/\beta_{62}$, and $\tilde{\alpha}_{13} = \alpha_{13}/\beta_{53}$. The degree to which the process used to identify model 1 biases the results is described in section V, which reports on the results of a sensitivity analysis.

The restrictions that identify this model are rejected at the 5 percent level but not at the 1 percent level ($\chi 2(7) = 12.32$, $p < 0.03$). Model 1 contains a cointegrating relation (CR #1) that includes Prod, P_{max}, P_{cut}, and P_{rec}, a second cointegrating relation (CR #2) that includes Ration and Cost, and a third cointegrating relation (CR #3) that includes P_{max}, P_{cut}, $P_{rec,}$ and Cost (Table 3). The elements of α indicate that CR #1 and CR #2 'load' into the equation for the first difference of oil production (Table 3).

Next, we attempt to identify whether the price effects in CV #1 and CV #3 are asymmetric and the nature of that asymmetry. The asymmetry can take many forms. The elements of β associated with P_{max}, P_{cut}, and P_{rec} in each of the cointegrating vectors may be different. Alternatively, two of the elements of β associated with P_{max}, P_{cut}, and P_{rec} may be equal. Wollfram (1971) describes an asymmetry in which a price increase to an all-time high has a greater effect than a price cut or a price recovery. Trail *et al.*, (1978) describe an asymmetry in which price increases, either an increase to an all time high or a recovery back towards an all-time high, have a greater effect than a price cut.

To differentiate among these and other categories of price asymmetry, we test restrictions on the elements of β associated with P_{max}, P_{cut}, and P_{rec} in two steps. In the first step, we impose restrictions on CV #1 or CV #3 that equalize two of the three elements of β associated with P_{max},

P_{cut}, and $P_{rec.}$ Six sets of restrictions are possible--three for each cointegrating vector. Each of these restrictions is evaluated individually relative to model 1 using a likelihood ratio test. The results indicate that only one set of restrictions can be rejected (Table 4). The elements of β associated with P_{cut} and P_{rec} are not equal in CV #3.

Next, we test restrictions that equalize two elements of β associated with P_{max}, P_{cut}, and P_{rec} in each of the cointegrating vectors. There are six combinations. The restriction on two of the three elements of β associated with P_{max}, P_{cut}, and P_{rec} in the other cointegrating vector is evaluated two ways. We test whether the individual restriction on the other cointegrating vector is rejected. We also evaluate whether the entire set of restrictions is rejected. This second evaluation is necessary because the set of restrictions that constitutes model 1 is rejected at the 5 percent level. The results indicate that three of the six possible restrictions cannot be rejected and generate three over-identified models (model 2, model 3, and model 4) that cannot be rejected at the 5 percent level (Table 4).

We also test whether the elements of β associated with P_{max}, P_{cut}, and P_{rec} are equal (perfect price reversibility) in CV # 1 or CV # 3. Restrictions that equalize the elements of β associated with P_{max}, P_{cut}, and P_{rec} in CV # 3 are rejected relative to any set of restrictions that would equalize two of the elements of β associated with the decomposed price series in CV #3 (Table 3). On the other hand, it is not possible to reject restrictions that equalize elements of β associated with P_{max}, P_{cut}, and P_{rec} in CV #1 if the elements of β associated with P_{rec} and P_{max} are equal in CR #3. Nonetheless, this model is not acceptable because the entire set of restrictions is rejected at the 5 percent level.

IV Discussion

Models 2 - 4 generally are similar. In these models, CR # 1 can be interpreted as the long run relation between price and the rate of oil production. This interpretation is consistent with the sign and statistical significance of CV #1. The elements of β associated with P_{max}, P_{cut}, and P_{rec} are statistically significant and have the sign opposite that associated with Prod. These results indicate that production increases as the real price of oil; (1) rises to an all-time high, (2) recovers back towards that high or, (3) that production declines as oil prices fall way from a previous high. Models 2 - 4 include every possible combination in which two of the elements of β associated with P_{max}, P_{cut}, and P_{rec} are equal. These results indicate that the relation between price and production is asymmetric, but the nature of the asymmetry is uncertain.

The element of α associated with CR #1 in model 2-4 indicates that disequilibrium in CR #1 has a statistically significant effect on the annual change in oil production. The negative sign indicates that disequilibrium in the long run relation among Prod, P_{max}, P_{cut}, and P_{rec} has a self-correcting effect on production in the long run. For example, an increase in price reduces the value of CR #1. This reduction has a positive effect on production because the element of α associated with CR #1 is negative, which increases production towards the long run value implied by CR #1. The rate of adjustment is slow. The element of α associated with CR #1 and the equation for the first difference for production in model 2-4 implies that 6-7 percent of the disequilibrium in the long run relation among Prod and P_{max}, P_{cut}, and P_{rec} is eliminated each year.

Cointegrating relation #2 can be interpreted as the long run relation between prorationing and the average cost of production. The elements of β associated with Ration and Cost indicate reducing the fraction of capacity allowed to operate is related negatively to average costs. This result is consistent with the economic inefficiency of prorationing decisions by the TRC (Prindle, 1981; Adelman, 1964). Rather than shutting-in high cost fields, decisions to proration capacity tended to favor small, independent, high cost producers. By allowing wells owned by high cost producers to remain open, shutting in capacity had a positive effect on average cost.

Production is not included in CR #2 (the restriction that eliminates production from CR #2 in model 2 cannot be rejected $\chi^2(1) = 1.56$, $p < 0.22$). Nonetheless, disequilibrium in CR #2 has a statistically significant effect on the annual change in production. The element of α associated with CR #2 and the annual change in oil production in models 2 - 4 indicates that disequilibrium in the long run relation between average cost and prorationing decisions by the TRC has a negative effect on production. This is consistent with economic theory--an increase in the average cost of production has a negative effect on production.

The effect of Ration on production is more difficult to interpret. Combined with the negative value for the element of α associated with CR #2 and the annual change in oil production, the element of β associated with Ration indicates that increasing the fraction of capacity allowed to operate has a negative effect on production in the long run. This negative effect may be generated by the institutional response to prorationing decisions by the TRC. Owners countered TRC mandates to shut-in production by increasing the rate at which they drilled wells and decreasing the space between wells (Adelman, 1964; Prindle, 1981). Both of these responses tend to increase production in the longer term by increasing the absolute level of capacity.

In the short term, opening capacity should increase production. This effect is represented by the element of Φ associated with the first difference of Ration (Table 3). This element of Φ in models 2-4 is positive (0.626 - 0.638) and is statistically significant ($p < 0.001$), which indicates that an increase in the fraction of capacity allowed to operate by the TRC has a positive effect on the first difference of production. This positive short run effect outweighs the negative long-run effect on production. Increasing the fraction of capacity allowed to operate by one percentage point (0.01) reduces production by 0.00243 via CR #2 (1.000 * 0.01 * -0.243) but increases production by 0.0063 via the first difference of Ration in the D matrix (.01 * 0.63). Together, these effects indicate that opening one percent of capacity increases production by about 4 million barrels per year.

CR #3 can be interpreted as the long run relation between the price of oil and the average cost of production. As the real price of oil increases, profit increases, which attracts investment to the industry. At the margin, increased investment forces firms to develop lower quality deposits, which increase average costs. The price asymmetry in CR # 3 is clear. Price recoveries have a greater effect than price cuts and price increases to an all-time high. Price increases to new all-time highs spur innovations that allow the industry to recover oil from fields which were previously uneconomic. Subsequent price reductions spur innovations that reduce the cost of the new technology. The lower cost allows the new technology to penetrate the market rapidly when the prices recover.

Production is not included in CR #3 (the restriction that eliminates production from CR #3 in model 2 cannot be rejected $\chi^2(1) = 0.01$, $p < 0.92$). Similarly, the element of α associated with CR #3 in models 1 -3 indicates that this proxy for economic profit does not affect production. For model 4, the element of α associated with CR #3 is statistically significant and positive. The positive sign is opposite that expected from theory. On the other hand, model 4 is the most restrictive of the models accepted at the 5 percent level.

V Sensitivity Analysis

Despite results that generally are consistent with the behavior of the oil industry, we cannot reject other sets of restrictions that overidentify three cointegrating relations. For example, we can identify a model in which CR #1 can be replaced with a cointegrating relation that includes Cost and the time trend (model 5). Although the set of restrictions associated with model 5 are less restrictive than those associated with models 1-4, the first cointegrating relation in model 5 implies that the average cost of production is trend stationary. This result seems inconsistent with the results of the augmented Dickey Fuller Statistic, which indicates that this variable contains a stochastic trend. We recognize that the power of the Dickey Fuller statistic is limited (Enders, 1995; Kim and Schmidt, 1990; Schwert, 1989) but a visual examination of the data also seems to contradict the notion that the data for average cost are stationary about a deterministic trend (Figure 2).

To further evaluate the results described in the previous section, we increase the efficiency of the identification process by estimating a partial system in which Prod is the only endogenous variable, and P_{cut}, P_{rec}, P_{max}, Cost, and Ration are exogenous (Greenslade *et al.*, 1997). This reduces the number of cointegrating relations possible, a maximum of one cointegrating relation, and makes it easier to identify it. To estimate this partial system, we repeat the tests for lag length and the rank of Π. Both the Schwartz and Hannon Quinn criteria indicate that a lag length of one is appropriate and both the λ_{max} and λ_{trace} statistics (41. 23) clearly reject the null hypothesis that there are no cointegrating relations. The consistency of the Schwartz and Hannon Quinn criteria reinforce our decision to use the shortest lag length for the full system.

The results of models 6-8 indicate that equalizing the elements of β associated with P_{rec} and P_{max} is the least restrictive form of price asymmetry (model 6). Nonetheless, we cannot reject alternative sets of restrictions that equalize two elements of β associated with the decomposed price series. Relative to model 6, we reject model 9, which postulates a symmetric relation between prices and production ($\chi^2(1) = 4.25$, $p < 0.04$). We cannot reject a restriction that eliminates the deterministic trend from model 6 (model 10). All of the remaining variables are statistically significant and consistent with the t statistics for the elements of β, we reject strongly additional restrictions that would eliminate any of the remaining variables from the single cointegrating relation in model 10.

The results for model 10 generally are consistent with model 2. The single cointegrating relation contains the same variables as three cointegrating relations in models 2 - 4. The variables in the single cointegrating relation in model 10 have the same sign and the same effect on production as the variables in CR #1 - CR #3 in models 2-4. As in models 2-4, the relation between price and production is asymmetric. In model 10, price increases (P_{max}, P_{rec}) have a greater effect on production than price reductions. Increases in average cost have a negative effect on production. Increases in the Ration variable have a negative long run effect on production (via the cointegrating relation), but this effect is more than offset by the positive short run effect. The short run effects of the P_{cut}, P_{max}, and Cost also are statistically significant, and their sign is consistent with the effect of these variables on production indicated by the cointegrating relation. The negative value for alpha indicates that disequilibrium in this long term relation has a self-correcting effect on production in the long run. Together, these results indicate that the restrictions used to over-identify the three cointegrating relations in the full system may be a reasonable representation of the economic, technical, geological, and institutional determinants of oil production in the lower 48 states.

VI Conclusions

Any lingering uncertainty about the identification of the cointegrating relations does not affect the conclusion that the assumptions which underlie the economic and Hubbert models of oil production are inconsistent with the historical record, and that these inconsistencies prevent these models from making an accurate assessment of the quantity of oil that remains in the lower 48 states.

Hubbert’s models (1956,1962) have generated the most accurate forecast of oil production in the lower 48 states (Cleveland and Kaufmann, 1991), but this accuracy is spurious. The cointegration analysis indicates that oil production in the lower 48 states shares stochastic trends with the decomposed price series, average costs, and prorationing decisions by the TRC. These stochastic trends are not present in the deterministic bell-shaped curve, therefore the first difference of the bell-shaped curve drifts away from the annual change in oil production for extended periods (Figure 3). Hubbert was able to predict the peak in US production accurately because real oil prices, average real cost of production, and decisions by the TRC co-evolved with production in a way that traced a seemingly symmetric bell-shaped curve over time. A different evolutionary path for any of these variables could have produced a pattern of production that is significantly different from a bell-shaped curve. For example, if the TRC did not shut in production, or did not favor high cost producers, production may not have followed a bell-shaped curve and production may not have peaked in 1970. In effect, Hubbert got lucky. Thus, the Hubbert’s model ability to forecast production in the lower 48 states accurately probably cannot be extrapolated to other regions, as done by Colin and Laherrere (1998).

Similarly, there is little reason to believe that the forecasts generated by economic models are any more accurate. Both the full system and the partial system have a cointegrating vector that includes Cost. This variable does not increase monotonicaly, as assumed by the Hotelling model. Consequently, the lack of an accurate assessment for how the cost of production will evolve is a critical roadblock to an accurate estimate of oil production in the lower 48 states. This difficulty is exacerbated by the asymmetric price effect on production. The lack of perfect price reversibility implies that there is no unique or optimal path for price to follow towards the choke price, which would make it impossible to determine recoverable supply from information about the choke price.

These inconsistencies with the economic model also undermine the *de facto* policy for managing the depletion of conventional oil supplies—a belief that the competitive market will generate a smooth transition from oil. The asymmetric relation between prices and production undermines the assumed self-correcting market mechanism that implies rational behavior by rent maximizing individuals will generate an optimal path for price and production. The asymmetry implies that prices may not rise smoothly towards the choke price. Without this smooth price increase, the market may not generate the signals needed to spur timely development and investment in alternative energy sources. Without these efforts, alternative fuels may not be available in quantities sufficient to prevent prices from exceeding the choke/trigger price.

This effect may be compounded by a compression of the period during which prices rise. The results of the cointegration analysis indicate that the cost of production is negatively related to production. If the long run cost curve for oil production does not increase monotonically, there may be an extended period during which production increases and prices decline. This will shorten the period during which rising prices and declining production spur the development of alternatives relative to a steadily increasing cost curve. This too, will diminish the likelihood that alternative fuels will be available in quantities sufficient to prevent oil prices from exceeding the choke and/or trigger price. The negative economic effects associated with high prices and energy shortages imply that the inconsistencies with the economic model identified by this analysis may be sufficient to warrant a greater degree of government intervention in the transition from oil than currently envisioned by most policy makers.

Literature Cited

Cleveland, C.J. 1991. Physical and economic aspects of resource quality: the cost of oil supply in the lower 48 United States, 1936-1988. *Resources and Energy* **13:** 163-188

Cleveland, C.J. and R.K. Kaufmann. 1991. Forecasting ultimate oil recovery and its rate of production: Incorporating economic forces into the models of M. King Hubbert. *The Energy Journal.* **12**(2):17-46.

Colin, C.J. and J.H. Laherrere, 1998. The end of cheap oil. *Scientific American* March 78-83.

Cuthbertson K., S. G. Hall, and M. P. Taylor, 1992. *Applied Econometric Techniques*, University of Michigan Press, Ann Arbor MI.

Dickey, D.A. and Fuller, W.A., 1979. Distribution of the estimators for auto regressive time series with a unit root. *Journal of American Statistics Association.* **7**: 427-431.

Enders, W. 1995, *Applied Econometric Time Series*, John Wiley, New York.

Engle, R.E. and Granger, C.W.J. 1987. Cointegration and error correction: representation, estimation, and testing. *Econometrica* **55**: 251-276.

Fisher, A.C. 1981. *Resource and Environmental Economics* Cambridge University Press, Cambridge).

Frankel, P.H. 1946. *Essentials of Petroleum* (Frank Cass, London).

Gately, D. 1992. Imperfect price-reversibility of U.S. gasoline demand: Asymmetric responses to price increases and declines. *The Energy Journal* 13(4): 179-207.

Greenslade, J.V. S.G. Hall, and S.G.B. Henry. 1997. On the identification of cointegrated systems in small samples: practical procedures with an application to UK wages and prices. London Business School Mimeo.

Harris, D. 1977. Conventional crude oil resources of the United States: recent estimates, methods, for estimation and policy considerations. *Materials and Society* **1**:263-286.

Hotelling, H. 1931. The economics of exhaustible resources *Journal of Political Economy* **39**:137.

Hubbert, M. K., 1956. Nuclear energy and the fossil fuels. In: American Petroleum Institute (Editor), Drilling and Production Practice. American Petroleum Institute, New York.

Hubbert, M. K., 1962. Energy Resources. A Report to the Committee on Natural Resources, National Academy of Sciences. Government Printing Office. Publication No. 1000-D.

Johansen, S., 1988: Statistical analysis of cointegration vectors. *Journal of Economic Dynamics and Control*, 12, 231-254.

Johansen, S. and K. Juselius, 1990: Maximum likelihood estimation and inference on cointegration with application to the demand for money. *Oxford Bulletin of Economics and Statistics*, 52, 169-209.

Johansen, S. and K. Juselius, 1994: Identification of the long-run and the short run structure: an application to the ISLM model. *Journal of Econometrics*, 63, 7-36.

Kaufmann, R.K. 1991. Oil production in the lower 48 states: Reconciling curve fitting and econometric models. *Resources and Energy* 13:111-127.

Kerr, R. 1998. The next oil crisis looms large--and perhaps close *Science* **281**:1128-1131.

Kim, K. and P. Schmidt. 1990, 'Some evidence on the accuracy of Phillips-Perron tests using alternative estimates of nuisance parameters', *Economics Letters* **34**, 345-350.

Menard, H.W. and G. Sharman. 1975. Scientific uses of random drilling models *Science* 190:337-343.

Osterwald-Lennum, M. 1992. A note with quantiles of the asymptotic distribution of the maximum likelihood cointegration rank test statistics. *Oxford Bulletin of Economics and Statistics* **54**:461-471.

Pantula, S.G. 1989. Testing for unit roots in time series data. *Econometric Theory* **5**:256-271.

Pearce, D. W. and Turner, R. K., 1990. Economics of Natural Resources and the Environment. The Johns Hopkins University Press, Baltimore.

Pesaran, M.H. and H. Samiei. 1995 Forecasting ultimate resource recovery. *International Journal of Forecasting* 11(4) 543-555.

Pesaran, M.H. and Shin, Y. 1994. Long run structural modelling. University of Cambridge, Mimeo.

Prindle, D.F. 1981. *Petroleum Politics and the Texas Railroad Commission* University of Texas Press. Austin.

Ryan, J.H. 1965. National academy of sciences report on energy resources: discussion of limitations on logistic projections, *Bulletin of the American Association of Petroleum Geologists* **49**(1) 1713-1720.

Schwert, G. W. 1989, 'Tests for unit roots: a Monte Carlo investigation', *Journal of Business and Economic Statistics* **7**, 147-159.

Tietenberg, T., 2000. Environmental and Natural Resource Economics. Addison-Wesley, Reading.

Traill, B. D. Colman, and T. Young. 1978. Estimating irreversible supply functions *American Journal or Agricultural Economics* **60**:528-531.

Uri, N.D. 1980. Crude oil appraisal of the United States. *Energy Journal* 1(3):65-75.

Watkins, G.C. 1992. The Hotelling principle: autobahn or cul de sac? *The Energy Journal* 13(1) 1-24.

Wolffram, R. 1971. Positivistic measures of aggregate supply elasticities: Some new approaches--some critical notes. *American Journal of Agricultural Economics* **53**:356-359.

Figure Captions

Figure 1 Decomposed series for real oil prices (P_{max}, P_{cut}, P_{rec}) and real oil prices (Poil).

Figure 2 The average cost of producing oil in the US.

Figure 3 The annual change in oil production (solid line), the annual change generated by model 2 (dotted line) and the annual change generated by Hubbert's bell shaped curve (dashed line)

Table 1

Time Series Properties of Indicated by the Augmented Dickey Fuller Statistic

	Augmented Dickey Fuller Statistic for:		Classification
Variable	Level	First Difference	I(0), I(1), I(2)
Prod	-0.40	**-4.84**	I(1)
Pmax	-1.42	**-4.83**	I(1)
Pcut	0.18	**-4.16**	I(1)
Prec	-0.21	**-4.53**	I(1)
Roil	-2.64	**-4.53**	I(1)
Ration	-2.25	-2.26	I(2)
Cost	-1.95	**-3.92**	I(1)

Table 2
Lambda statistics used to choose the rank of Π

HO:r	p-r	λ_{max}	λ_{trace}	λ_{max}*	λ_{trace}*
0	6	**192.68**	**67.27**	114.90	43.97
1	5	**125.41**	**48.99**	87.31	37.52
2	4	**76.42**	**37.57**	62.99	31.46
3	3	38.85	20.89	42.44	25.54
4	2	17.96	13.08	25.32	18.96
5	1	4.88	4.88	12.25	12.25

λ_{max}* critical value of λ_{max} at $p < .05$ Values from Osterwald-Lenum (1992)

λ_{trace}* critical value of λ_{trace} at $p < .05$ Values from Osterwald-Lenum (1992)
Values that exceed the $p < .05$ threshold shown in bold