Finding Costs in the U.S. Petroleum Industry: Assessing the Opposing Effects of Technological Change and Depletion with Error Correction Modeling

John T. Cuddington*

November 14, 2001 (9/2/03) **

* Economics Department, Georgetown University. Helpful discussions with Bob Cumby and Diana Moss are gratefully acknowledged.

Abstract

Finding Costs in the U.S. Petroleum Industry: Assessing the Opposing Effects of Technological Change and Depletion with Error Correction Modeling

There is a common claim in the nonrenewable resource literature that improvements in technology may largely offset increasing scarcity over time as the resource is exploited. Yet there is little empirical evidence on this crucial matter. This paper makes two contributions to the literature on finding costs in the exploration and development process. First, it makes the methodological point that previous methods for estimating finding costs are inappropriate due to the likely nonstationarity of the variables involved. Second, using a variable capturing technological innovation in the U.S. petroleum industry, it estimates a vector error correction model (VECM) for average finding costs for nonassociated natural gas and crude oil. The separate impacts of ongoing technological change and resource depletion on finding costs are isolated. Counterfactual simulations are then used to examine the extent to which improvements in technology have offset the effects of increasing scarcity over time. The results suggest that technological advance has been very significant in offsetting what would otherwise have been sharply rising costs for additions to U.S. natural gas reserves. Technology's impact on crude oil finding costs has been more modest, presumably because that segment of the U.S. industry is more mature.

JEL Classification: D24 Production; Capital and Total Factor Productivity
Q31 Nonrenewable Resources and Conservation: Supply and Demand
L71 Industry Studies: Primary Products (Mining, Extraction, and Refining: Hydrocarbon Fuels)

key words: technological change, productivity growth, cost functions, cointegration, error correction models, nonrenewable resource depletion.

Finding Costs in the U.S. Petroleum Industry: Assessing the Opposing Effects of Technological Change and Depletion with Error Correction Modeling

The economics and public policy literatures have long been concerned with the determinants of supply and demand for nonrenewable resources, as well as long-term trends and cycles in the prices of primary commodities and resource-based products. The 2000-1 'energy crisis' in the U.S. has renewed policymakers' interest in such matters, as they attempt to explain the sharp run-up in energy prices as well as the longer-term prognosis.

This paper is concerned with longer-term supply considerations in the U.S. petroleum industry, although the methodological issues addressed also apply for other nonrenewable resources (e.g. minerals and primary metals). The petroleum industry is of general interest, not just because of its pivotal role in economic activity. It has also benefitted from dramatic changes in state-of-the-art E&D and production technology. Technological advances such as three-dimensional seismic techniques, polycrystalline diamond compact drill bits, horizontal drilling, and offshore platforms capable of operating in hostile, deep-water environments are widely acknowledged to have had significant impact on productivity in E&D.[1] Bohi (1999) provides an excellent discussion of these developments over the last decade, and their impact on various exploration and development (E&D) productivity measures.

The process of supplying petroleum products involves (i) E&D activity to find economically viable additions to the stock of proven reserves, (ii) extraction of those reserves and (3) distribution to downstream producers or final consumers. Presumably, as proven

[1] See, e.g., American Gas Association (1990), National Petroleum Council (1987), and Tippee and Beck (1991). Moss (1993) provides a detailed discussion of various technological advances in the industry.

reserves are exploited, it will become increasingly costly to uncover additional reserves through E&D -- the so-called 'resource depletion effect.' Improvements in technology may, to a greater or lesser extent, offset the price effects of increasing scarcity over time as the resource is exploited.[2] Yet there is little *empirical* evidence on this matter. Recent exceptions are Bohi (1999) and Cuddington-Moss (CM) (2001). This paper, like CM, studies the E&D process for additional U.S. petroleum reserves by estimate finding cost functions including both a proxy for technology and the depletion effect. The E&D process is presumably where depletion has its initial impact.

The present paper has two objectives. The first is methodological and involves the statistical methods used in estimating finding cost functions. It is argued that there is a strong theoretical presumption that the variables that appear in finding cost functions are nonstationary. Therefore, modern time series econometric methods – including unit root and cointegration tests, and vector error correction models (VECMs) – are appropriate for estimating such relationships. As they have not been used to date in the finding cost literature, previously estimated finding cost functions may, in fact, be spurious regressions (á la Granger-Newbold (1974)).

The second objective is empirical. The paper uses the aforementioned techniques to re-examine the extent to which improvements in technology have offset the effects of increasing scarcity over time in the U.S. petroleum industry. Specifically, we use the Cuddington-Moss

---

[2] Technological advances improve the efficiency of E&D by reducing costs of geological and geophysical (G&G) activity, drilling activity, and/or increasing success ratios in the discovery process. (In addition, technological innovations may lead to synthetic substitutes for nonrenewable resources, which affects the demand for primary products.)

dataset to re-estimate finding costs for nonassociated natural gas[3] and crude oil reserves using VECMs.

Our estimated VECMs are able to isolate separate contributions of ongoing technological change and resource depletion on average finding costs. The VECMs are then used to construct counterfactual simulations for two key variables: average finding costs and annual discoveries, i.e. additions to the stock of proven reserves. In one scenario, technology change is assumed to continue at its historical rate. In a second, there is no further technological change as of a specified date (1973 or 1991). A comparison of the two scenarios suggests that technological advance has been very significant in offsetting what would otherwise have been sharply rising costs for additions to U.S. natural gas reserves. Technology's impact on crude oil finding costs have been more modest, presumably because that segment of the U.S. industry is more mature.

Section I reviews a typical specification of average finding cost functions in the literature, including possible proxies for resource depletion and technological change. It discusses the presumption that several of the variables typically used in such functions are nonstationary. Section II discusses the pros and cons of using OLS and/or the instrumental variable (IV) approach to estimating finding cost functions in such situations. Section III discusses the VECM as an alternative approach for estimating long-term equilibrium relationships among nonstationary variables. Section IV estimates finding cost functions for nonassociated natural gas and oil. Section V reports simulations for average finding cost based on forecasted rates of technological change, as well as alternative time paths based on the assumption of no technological change. Comparing these scenarios allows us to assess the long-

[3] Nonassociated gas is gas *not* occurring in association with oil.

term effects of technological change on finding costs. Section VI concludes.

I. FINDING COST FUNCTIONS and NONSTATIONARY VARIABLES

In the nonrenewable resource literature, it is common to estimate finding costs of additional reserves taking into account the depletion effect and, to the extent possible, technological change. A standard approach is to estimate an average finding cost function of the following form:[4]

$$\ln(C/PQ)_t = \beta_0 + \beta_1 TECH_t + \beta_2 lnQ_t + \beta_3 \ln(Z_t) + \epsilon_t \qquad \textbf{(1)}$$

where C_t is the total cost of E&D activity in period t. P_t is an appropriate sectoral price deflator picking up the effects of factor input prices on total costs. Q_t is the (stochastic) output of E&D process, namely the discovery of new, economically viable reserves in period t. $TECH_t$ is a measure of the level of technology currently in widespread use. Z_t is a proxy for increasing scarcity as existing reserves are exploited; it captures the depletion effect. The construction of TECH and Z series is discussed below.

Regarding the expected signs of the regression coefficients, β_1 is presumably negative, as improvements in technology should reduce average finding costs. β_2 would be zero if the E&D process exhibits constant returns to scale; $\beta_2 < 0$ (>0) indicates scale (dis)economies. β_3 is positive reflecting the depletion effect. Livernois (1988, 383) expresses scepticism about

[4] See CM (2001) for citations. They provide two models that yield such a cost function where TECH is a count variable for the total number of implemented technologies. In their 'quality ladders' model , the TECH variable enters in level form; in the 'technological varieties' model, log(TECH) appears.

satisfactorily identifying all of the parameters in (1): "One cannot expect to capture the *separate* effects of depletion and technological change [which he envisaged being captured by a time trend] by including two highly correlated trend variables in the cost function. Rather, the best one can do is to capture the *net* effects of depletion and technological change."

Typical proxies for the depletion effect, Z_t, are cumulative past reserve additions or cumulative past drilling effort (see, e.g., Livernois (1988), Pindyck (1978), Uhler (1979)).[5] In this paper, we'll use cumulative past reserve additions as a proxy for the depletion effect; the same conceptual issues arise with other proxies. Thus, Z_t is defined as:

$$Z_t \equiv \sum_{s=0}^{t-1} Q_s \quad . \tag{2}$$

The proxy for the level of technology in widespread use is more problematic. Some authors have used a simple time trend at capture the effect of ongoing technological change: $TECH_t = a + b\mathbf{t}$. An alternative is to cumulate past expenditures on R&D or patents as a proxy for the level of technology. Recently, Moss (1993) and CM (2001) have used a measure of technology based on a detailed analysis of articles in trade publications.[6] New technologies used in E&D in the petroleum sector are identified as of the year that they come into widespread use. Defining $NEWTECH_t$ as the number of new technologies implemented in period t, the

---

[5] Kenneth Arrow introduced a similar variable, the cumulation of past output or labor input, in his work on 'learning by doing' in production. Thus, β_3 in (1) might be interpreted as picking up the *net* impact of the depletion and learning by doing effects.

[6] Cuddington-Moss emphasize that this measure of technology has merits relative to cumulating past R&D expenditure or patents. It is temporally more accurate, and avoids the problem that all R&D does not produce viable technologies (in the sector where the R&D occurs or at all, in some cases). Some useful technological innovations are never patented; many patent inventions are never implemented.

cumulative number of technologies in widespread use becomes the proxy for the technology level:

$$TECH_t \equiv \sum_{i=0}^{t} NEWTECH_i \tag{3}$$

All of the aforementioned technology measures involve the use of cumulation functions like (3).

If the variables being cumulated in (2) and (3), Q_t and $NEWTECH_t$, are stationary, then their cumulations Z_t and $TECH_t$ will almost surely have unit roots.[7] They are said to integrated of order one or I(1). That is, they need to be first-differenced to achieve stationarity. To see that cumulating a stationary time series will typically produce a nonstationary one, suppose Q_t in (2), is mean stationary process with an ARMA representation. That is, $A(L)Q_t = B(L)\epsilon_t$, where A(L) is an invertible lag polynomial, B(L) is stationary and ϵ_t is i.i.d.. Taking the first difference of (2) yields:

$$\Delta Z_t \equiv (1-L)\ Z_t = \sum_{s=0}^{t-1} Q_s - \sum_{s=0}^{t-2} Q_s = Q_{t-1} \ . \tag{4}$$

Thus, if Q_t is stationary, ΔZ_t must be stationary. If ΔZ_t is stationary, Z_t itself must have a unit root.

Thus, there is a strong presumption that the technology and depletion variables used in

7 An exception is the special case where B(L) below includes an MA root equal to -1.

estimating finding cost functions are nonstationary.[8],[9] *A priori*, one would not expect marginal or average finding costs to be stationary variables either. After all, they depend on a diminishing stock of an exhaustible natural resource. The simplest variant of the Hotelling model, for example, predicts that marginal costs will rise at a rate equal to the real interest rate (rather than being stationary). Of course, technological improvements can have a countervailing effect, but it would be a fluke indeed if ongoing resource depletion and technological change just offset each other period after period so as to produce a stationary average finding cost series.[10]

II. ESTIMATION METHODS

The literature on finding costs contains many papers with cost functions that are similar to equation (1). For applications to the U.S. petroleum sector, the data samples are invariably quite short. The CM study, for example, uses annual observations from 1967 through 1990. The same constraint faced earlier authors since their studies use basically the same series on U.S. reserve additions.[11]

What estimation techniques have been used to estimate equation (1)? To what extent do

---

[8] Note that if Q and NEWTECH are I(1) rather than I(0) processes, then Z and TECH would be I(2).

[9] Also the statement may not hold if the technology or depletion variables enter the cost function in log form, ln(TECH) or ln(Z). By considering various Box-Cox transformation of potentially I(1) variables, Franses and Koop (1998, p.8), show that "unit root inference can be very sensitive to the transformation chosen..."

[10] See Kamien and Schwarz (1978) for a model of resource depletion with endogenous technological change.

[11] As CM (2001) explain: "The estimation period is unfortunately limited because reserve additions data prior to 1966 do not disaggregate reserve 'revisions' from the other components of reserve additions. Revisions and 'extensions' are lumped together. To some extent revisions reflect information gained through development drilling. They also reflect, in large part, a reassessment of the economic viability of *existing* reserves as market prices change. This component of revisions does not reflect 'output' from the E&D process. Thus, it is generally argued that reserve revisions should be excluded from the remaining components of reserve additions, as is done for the post 1966 data we use."

the properties of the estimators rely on asymptotic as opposed to small sample justification? To what extent are they valid in the presence of nontationary variables? These are the questions to which we now turn.

Most finding cost functions (or their underlying production functions) have been estimated using OLS and/or instrumental variables (IV) methods. Setting stationarity issue aside for the moment, it is well-known that OLS estimates are biased when regressors are correlated with the error term, as would be the case if reserve additions or technology are endogenous. Hence, IV estimates are typically reported to address possible simultaneity bias. While IV estimates are consistent, it is well-known that they may exhibit considerable small-sample bias depending on the quality of the instruments. Some authors (e.g. Epple, and Livernois) have used full information maximum likelihood estimation (FIML) methods. These methods also rely on asymptotic justification (assuming stationary variables) rather than estimators' small sample properties. The same comment applies for regression equations including lagged dependent variables or corrections for AR(1) errors.

All previous work on finding cost functions in the nonrenewable resource literature ignores the important stationarity issue highlighted in Section I above. It is now well known from the time-series econometrics literature, however, that standard regression techniques (i.e., OLS, IV, FIML) may produce spurious regressions -- and hence invalid statistical inferences -- if variables under consideration are nonstationary due to the presence of unit roots. Suppose that the individual series in finding cost equations in the published literature are nonstationary unit-root processes. How can we interpret the OLS, IV, and FIML results on finding costs? Are they meaningful relationships or merely spurious regressions à la Granger and Newbold (1974)?

Are there alternative ways to estimate the long-run relationship between finding costs, depletion, and technology?

To answer the first question, one must ask if the four variables in equation (1), [$\ln(C_t/P_tQ_t)$, $\mathbf{TECH_t}$, $\ln(Q_t)$, $\ln(Z_t)$], are cointegrated. If the variables are cointegrated, OLS and IV estimates of the β coefficients are 'superconsistent,'[12] although they may exhibit considerable small-sample bias. The estimated standard errors on the β coefficients, however, are inconsistent. This invalidates any hypothesis tests. If the four variables [$\ln(C_t/P_tQ_t)$, $\mathbf{TECH_t}$, $\ln(Q_t)$, $\ln(Z_t)$] are not cointegrated, on the other hand, OLS and IV estimates of (1) are spurious regressions.

III. AN ERROR CORRECTION MODEL OF FINDING COSTS

If I(1) variables are cointegrated, there are advantages to estimating the finding cost function by specifying a VECM, rather than relying solely on OLS or IV estimates. Suppose that one concludes, based on cointegration tests, that the variables under consideration are linked by a long-run equilibrium or cointegration relationship. Then one can obtain consistent coefficient estimates of that relationship – here, the average finding cost function -- *and* asymptotically-valid standard errors using an VECM. The VECM for the four variables in the finding cost function (1) is a four-equation system of the form:

$$\Delta\ln(C/PQ)_t = \gamma_{10}+\alpha_1[\ln(C/PQ)_{t-1}-\beta_0-\beta_1 TECH_{t-1}-\beta_2 lnQ_{t-1}-\beta_3\ln(Z_{t-1})]+\gamma_{11}(L)\Delta y_{t-i}+v_{1t} \quad (5)$$

$$\Delta TECH_t = \gamma_{20}+\alpha_2[\ln(C/PQ)_{t-1}-\beta_0-\beta_1 TECH_{t-1}-\beta_2 lnQ_{t-1}-\beta_3\ln(Z_{t-1})]+\gamma_{21}(L)\Delta y_{t-i}+v_{2t} \quad (6)$$

[12] That is, they converge to the true parameter values as the sample size T increase at a rate that is faster than the standard $T^{1/2}$.

$$\Delta\ln(Z_t) \quad = \quad \gamma_{40}+\alpha_4[\ln(C/PQ)_{t-1}-\beta_0-\beta_1 TECH_{t-1}-\beta_2 lnQ_{t-1}-\beta_3\ln(Z_{t-1})]+\gamma_{41}(L)\Delta y_{t-i}+v_{4t} \quad (8)$$

$$\Delta\ln(Q_t) \quad = \quad \gamma_{30}+\alpha_3[\ln(C/PQ)_{t-1}-\beta_0-\beta_1 TECH_{t-1}-\beta_2 lnQ_{t-1}-\beta_3\ln(Z_{t-1})]+\gamma_{31}(L)\Delta y_{t-i}+v_{3t} \quad (7)$$

The error correction term in square brackets appears in all four equations. It indicates that deviation of the four variables from their long-run equilibrium in the previous period. The speed-of-adjustment coefficients on the error correction term in each equation, α_i, indicate the extent to which each variable adjusts when there is a deviation from equilibrium. Letting y_t denote a vector with all four variables as elements, the lag polynomials $\gamma_{ij}(L)\Delta y_{t-i}$ is shorthand for the lagged first differences of each of the four variables in the vector autoregression. The lag length is chosen so as to mop up any potential serial correlation in the error terms in (5)-(8) and to capture short-run dynamic interactions among the variables. Given the highly uncertain nature and timing of the discovery and technology adoption processes, this flexibility in capturing dynamic adjustment toward the long-run equilibrium is especially valuable. Not only does the VECM allow for complicated short-run dynamics, it treats all variables as potentially endogenous.[13] Finally, as already mentioned, the estimated VECM yields consistent estimates of standard errors, thereby permitting hypothesis testing when the underlying data are I(1) series and cointegrated.

---

[13] When all variables are stationary, simultaneity bias is addressed by using instrumental variables estimation in place of OLS. With I(1) variables, on the other hand, OLS estimates are consistent in spite of potential simultaneity bias. Nevertheless, "OLS estimation...is proposed only as a quick way to obtain an initial estimate of the cointegrating vector. (See Hamilton (1994, p. 589.))

What are the pros and cons of alternative methods – VECMs vs. OLS and IV estimation – for estimating finding cost equations like those in (1) in the presence of nonstationary variables. The VECM-based estimates clearly dominate if one is interested in doing hypothesis tests based on asymptotic distribution theory. What can we say about the relative merits of the coefficient estimates obtained from OLS, IV, and VECMs in small sample contexts? These issues have been addressed using Monte Carlo simulations analysis by Stock and Watson (SW)(1993, Section 6) and others (see the references in SW). SW examine the small sample performance of six different estimators of long-run cointegration relationships using a variety of Monte Carlo experiments. The estimators considered include static OLS (SOLS), Stock and Watson's dynamic OLS (DOLS) and dynamic GLS (DGLS) estimators, and Johansen's VECM maximum likelihood estimator ("JOH"), among others. Their conclusion is that:

> "...each estimator (except the correctly-specified JOH) has substantial bias in at least some of the simulations, although the bias is in each case less than for SOLS: no single estimator is a panacea." (SW, 798)... A lesson suggested by the empirical investigation...and by the Monte Carlo results is that, when estimating cointegration vectors, it can be valuable to use more than one of the currently available asymptotically efficient estimators." (SW, 811)

IV. ESTIMATING FINDING COST FUNCTIONS FOR GAS AND OIL

This Section applies time series methods to the dataset used by CM (2001) in estimating finding cost functions for natural gas and oil. The variables are defined above. See CM (2001, Appendix for details regarding the variable definitions, and a description of how the TECH variable was constructed.) Taking Stock and Watson's advice, we compare our estimates based on the VECM to the OLS and IV estimates from CM (2001) in Tables 2 and 3 below. E&D for nonassociated natural gas and oil are considered separately. In each case, we estimate a specification where the TECH variable enters in level form and a second specification where

Ln(TECH) enters. These two specifications correspond to the *quality ladders* (QL) model and *technology varieties* (TV) model, respectively, derived in CM. Their OLS and IV results are reproduced in columns 2 and 3 (TV model) and columns 5 and 6 (QL model) of Tables 2 and 3 for comparison with those from our VECM estimation to follow.

Unit Root Tests

Given the theoretical presumption that some of the variables are nonstationary, we first carry out augmented Dickey-Fuller (ADF) tests for the presence of a unit root for each series. These are summarized in Table 1; I(1) indicates integrated of order one; none of the series had more than one unit root. The ADF tests suggest that the logs of real average finding costs for both gas and oil, $\ln(C_t / P_t Q_t)_{gas}$ and $\ln(C_t / P_t Q_t)_{oil}$ respectively, are nonstationary, as suggested in Section I above. For natural gas, the results on reserve additions (either in levels of logs) and cumulative reserve additions are also consistent with the theoretical predictions above. The finding that NEWTECH is I(0) is consistent with the finding that the cumulative technology variable, $TECH_t$, is I(1). For oil, in contrast, there are mutually inconsistent results: *both* the log of reserve additions, $\ln(Q_{oil})$, and the log of cumulative reserve additions appear to be I(1).[14]

[14] See fn.9, however, for a possible explanation.

April 25, 00 source file: tables(4.25.00).wb3

Table 1
Augmented Dickey-Fuller Unit Root Tests

Variable Name	ADF t-stat.	Chosen † Specification: T,C,N	Lags †	Order of Integration
$\ln(C/PQ)_{gas}$	-1.88	C	4	I(1)
$\ln(C/PQ)_{oil}$	-1.72	C,D70,DD70	0	I(1)
$\ln(Q_{gas})$	**-3.16** *	C	4	I(0)
Z_{gas}	0.82	C	5	I(1)
$\ln(Z_{gas})$	-1.91	C	2	I(1)
$\ln(Q_{oil})$	-2.57	C,D70,DD70	0	I(1)
$\ln(Z_{oil})$	-2.02	C,D70,DD70	2	I(1)
$NEWTECH_t$	**-3.12** *	C	0	I(0)
$TECH_t$	-1.24	C	0	I(1)
$\ln(TECH_t)$	-1.26	C	3	I(1)

* (**) indicates significantly different from zero at the 5% (1%) level.

† In all ADF regressions, the general-to-specific (GTS) methodology was used to determine (i) the number of lags of the dependent variable included in the ADF regressions and (ii) whether to include a deterministic time trend, a constant, or both (neither). See Ng and Perron (1995) and Hall (1994) for Monte Carlo evidence supporting the use of the GTS procedure (rather than the Schwarz or Akaike criteria) for choosing lag length. In two cases, $\ln(Z_{gas})$ and $\ln(Z_{oil})$ the GTS method apparently left too few lags, as the correlogram of the residuals from the ADF regression indicated remaining serial correlation. Hence, the lag length was increased in these two cases until the Ljung-Box Q statistic indicated a failure to reject the null hypothesis of white noise residuals.

<u>Cointegration Tests and Alternative Estimates of the Long-run Cointegrating Equation</u>

Having found that many of the relevant variables in finding cost equations are nonstationary process due to the presence of unit roots, we test for cointegration among the variables. We carried out Johansen trace and max-eigenvalue tests for cointegration using the vectors of four variables in the TV and QL models, i.e., $y' \equiv [\ln(C_t/P_tQ_t)$, $\mathbf{\ln(TECH_t)}$, $\ln(Q_t)$, $\ln(Z_t)]$ or $y' \equiv [\ln(C_t/P_tQ_t)$, $\mathbf{TECH_t}$, $\ln(Q_t)$, $\ln(Z_t)]$, respectively. This was done separately for both natural gas and crude oil. In all four cases (i.e. TV model for Gas, QL-Gas, TV-Oil, QL-

Oil), the null hypothesis of no cointegration is easily rejected at the 1 percent significance level.[15]

Having strongly rejected the null hypothesis of no cointegration, an VECM is estimated to obtain consistent estimates of the long-run equilibrium relationship (the cointegrating equation (CE)) and, at the same time, capture short-run dynamic interactions among the variables.

The estimation results for the long-run equilibrium are shown in columns 1 and 3 in Tables 2 and 3 for gas and oil, respectively. IN both the TV and QL specifications, the estimated coefficients on the technology variable are highly significant and have the expected negative sign in both the natural gas and crude oil finding cost equations.

The depletion effect captured by the coefficient on ln(Z), on the other hand, is statistically insignificant in the TV models for both gas and oil. In the QL models, in contrast, depletion has a significant positive effect on finding costs, as theory would predict. The estimated coefficients on the log of reserve additions, LnQ, suggest constant returns to scale in the case of crude oil exploration. The results regarding economies of scale for natural gas are inconclusive, however, due to considerable variation in the coefficients on LnQ across the two

---

[15] In several cases, there was more than one significant cointegrating vector (as determined by the number of eigenvalues in the dynamic system that are statistically different from zero). It is now known that adding a stationary variable, like lnQ_t, to a collection of cointegrated I(1) variables will increase the number of cointegrating vectors by one. The added cointegrating vector is a trivial one (0,0,1,0), say, indicating that the third variable alone is stationary. In all cases, therefore, the VECM was estimated assuming a single cointegrating vector, so that the estimated long-run finding cost function will include the stationary variable, lnQ_t say, as well as the three I(1) variables.

In developing the error correction specification, Hansen and Juselius (1995) state: "Z_t is a px1 vector of stochastic variables... [Each element of] Z_t is ***at most*** I(1) and the statistical procedures discussed later are derived under this assumption. (My emphasis) However, not all the individual variables included in Z_t need to be I(1), as is often incorrectly assumed. To find cointegration between nonstationary variables, only two of the variables have to be I(1). Often a stationary variable might a priori play an important role in a hypothetical cointegration relation...*Note that, for every stationary variable included, the cointegration rank will increase accordingly.* (P.1, my emphasis) Section 2.8.8 entitled "Stationary Variables" provides more detail. It begins: "The system variables [in] Z_t should *always* be chosen because of their economic relevance, *not* because of their time series properties. Thus, *there are often both I(1) and I(0) variables in Z_t."* (p.11, their italics).

specifications.

[insert Table 2]

Table 2:
Estimated Average Finding Cost Functions for Nonassociated Natural Gas

Dependent Variable is Real Average Cost (in logs)

	Technology Varieties Model			Quality Ladders Model		
Estimation Method >>	**C.E. from ECM**	OLS estimates from CM (1999)	IV estimates from CM (1999)	**C.E. from ECM**	OLS estimates from CM (1999)	IV estimates from CM (1999)
constant	**-30.494**	**0.128** *	**1.163** *	**-43.387**	**-52.000**	**-62.653**
		(0.01)	(0.066)		(6.567) **	(6.044) **
ln(TECH)	**-4.094** **	**-0.207**	**-1.404**			
	(8.586)	(0.323)	(0.914)			
TECH				**-0.061** **	**-0.046**	**-0.054**
				(13.588)	(4.870) **	(3.448) **
lnQ	**2.180** **	**-0.482** *	**-0.805** *	**-0.449** **	**-0.482** #	**-0.222**
	(14.695)	(2.144)	(2.069)	(2.747)	(1.888)	(0.355)
lnZ	**-0.005**	**-0.194**	**0.314**	**2.284** **	**2.719**	**3.094**
	(0.028)	(0.288)	(0.285)	(8.609)	(5.287) **	(3.621) **
Serial correlation correction:	2 lagged differences in ECM++	AR(1) error process		2 lagged differences in ECM++	AR(1) error process	
		0.825 **	**0.818** **		**0.327**	**0.352**
		(13.707)	(19.089)		(3.899) **	(3.792) **
Goodness of Fit/Data Sample:	1970-1990	1968-1990	1968-1990	1970-1990	1968-1990	1968-1990
Standard Error of Regression	0.130 @	0.165	0.187	0.119	0.159 @	0.172
Standard Dev. of Dep. Variable	0.176 @	0.345	0.345	0.176	0.345 @	0.345
Adj. R2	**0.451** @	**0.772**	**0.706**	**0.545**	**0.787** @	**0.752**

Notes:

t-statistics in parentheses. ** (*,#) = significant at the 1% (5%, 10%) level.

++ lag length for the ECM was chosen by examining the Schwartz criteria for an unconstrained VAR in the (log) levels of the 4 variables: Ln(AC), N or LnN, LnQ, and Ln(sumQ). Likelihood ratio tests for lag length yielded the same choice.

@ note that goodness of fit statistics for the ECM columns pertain to the DLn(C/PQ) equation, which is in first differences, unlike the OLS and IV regression in CM (1999).

Table 3:
Estimated Average Finding Cost Functions for Oil

Dependent Variable is Real Average Cost (in logs)

	Technology Varieties Model			Quality Ladders Model		
Estimation Method >>	**C.E. from ECM**	OLS estimates from CM (1999)	IV estimates from CM (1999)	**C.E. from ECM**	OLS estimates from CM (1999)	IV estimates from CM (1999)
constant	**-34.018**	**2.897** (0.341)	**2.395** (0.238)	**-46.528**	**-28.546** ** (4.600)	**-39.187** ** (4.793)
DUM1970	na	**-1.816** ** (3.562)	**-2.213** ** (3.621)	na	**-3.363** ** (5.244)	**-4.369** ** (5.383)
ln(TECH)	**-0.453** ** (4.336)	**-0.592** (0.987)	**-1.073** (0.990)			
TECH				**-0.008** ** (4.843)	**-0.003** (0.847)	**-0.003** (0.865)
lnQ	**1.000** ** (23.979)	**-0.409** * (2.339)	**-0.285** (1.405)	**1.125** ** (16.534)	**0.311** (1.312)	**0.734** * (2.399)
lnZ	**-0.061** (0.708)	**-0.354** (1.205)	**-0.347** (1.140)	**0.304** ** (3.199)	**0.241** # (2.048)	**0.328** # (1.924)
Serial correlation correction:	2 lagged differences in ECM++	AR(1) error process **0.788** ** (18.710)	AR(1) error process **0.788** ** (24.623)	2 lagged differences in ECM++	na	na
Goodness of Fit/Data Sample:	1970-1990	1968-1990	1968-1990	1970-1990	1968-1990	1968-1990
Standard Error of Regression	0.240 @	0.169	0.174	0.225 @	0.246	0.272
Standard Dev. of Dep. Variable	0.840 @	0.631	0.631	0.840 @	0.619	0.631
Adj. R2	**0.918** @	**0.928**	**0.924**	**0.928** @	**0.842**	**0.815**

Notes:

t-statistics in parentheses. ** (*,#) = significant at the 1% (5%, 10%) level.

++ lag length for the ECM was chosen by examining the Schwartz criteria for an unconstrained VAR in the (log) levels of the 4 variables: Ln(AC), N or LnN, LnQ, and Ln(sumQ). Likelihood ratio tests for lag length yielded the same choice.

@ note that goodness of fit statistics for the ECM columns pertain to the DLn(C/PQ) equation, which is in first differences, unlike the OLS and IV regression in CM (1999).

V. ASSESSING THE EFFECTS OF TECHNOLOGICAL CHANGE

The use of VECMs permits an innovative approach to forecasting future reserve additions, average finding costs, and the arrival of new technologies. Moreover, the question: "to what extent has ongoing technological change offset increasing resource scarcity in the U.S. petroleum industry?" can be addressed. We do this by considering a counterfactual scenario with 'no technological improvement' after a specified date. This is done in a way that allows average finding costs and reserve additions in each period to adjust endogenously to changes in the reduced rate of technological improvement. Thus, one can simulate the extent to which annual reserve additions and average finding costs would have been higher or lower in the absence of ongoing technological change.

In contrast to the simulations developed here, the 'no technological change' simulations in CM (2001) make the strong assumption that the time path of annual reserve additions remains equal its realized historical time path with or without ongoing technological change. This is a strong assumption. CM emphasize that their:

> 'no technological change' scenario must be interpreted carefully. Without technological improvements, reserve additions might have followed a very different time path than the historical series... To avoid these difficulties due to the endogeneity of the production of reserve additions, our simulations keep this 'output' (lnQ) on its historical path and ask: *what would the average finding cost have been to prove up reserves in a way that matched the historical time path, but without any technological improvements?*

In theory, reserve additions could be higher or lower under a scenario with no technological progress. There are opposing supply and demand-side effects. Without technological improvements, the average finding cost curve would undoubtedly shift upward, tending to cause the profit-maximizing 'supply' of reserve additions each period to fall. On the other hand, without advanced technology the probability of actually finding additional reserves when E&D

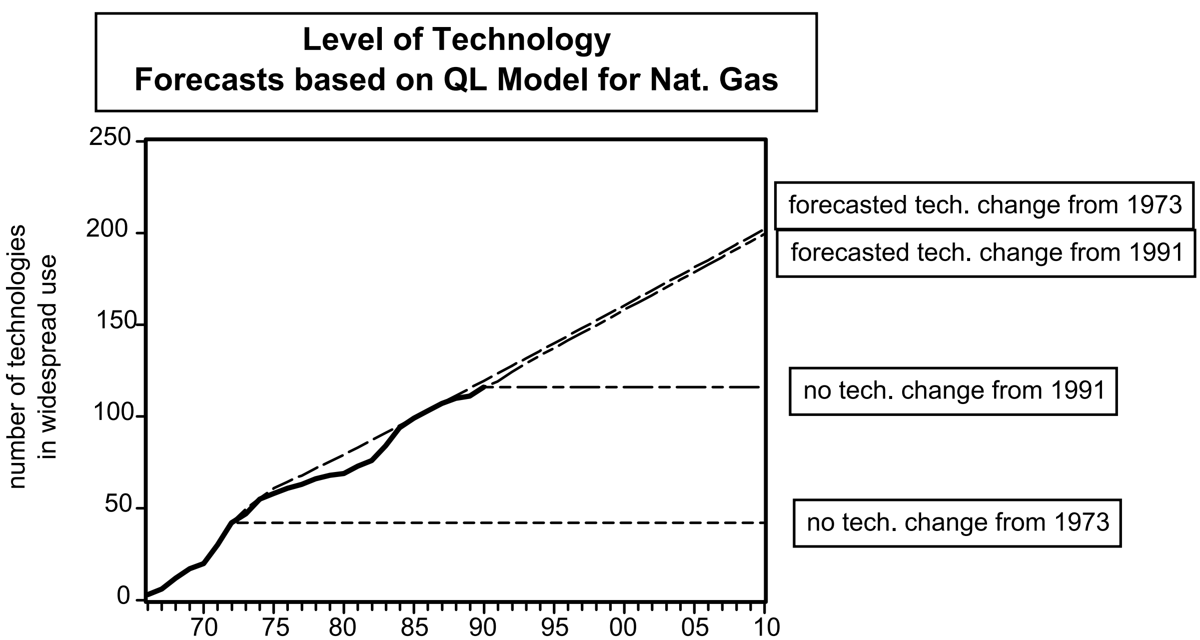
activity is undertaken will be lower. Thus, without technological improvements the 'demand curve' for reserves (by firms downstream in the petroleum industry, say), will shift outward. *A priori*, either effect could dominate. Thus, actual reserve additions may rise or fall in the counterfactual 'no technological change' scenario relative to historical experience. Our simulations based on VECMs allow either possibility. The estimation and simulation results suggest that annual reserve additions in the U.S. would have been considerably *higher* in the absence of ongoing technological change.

How are the simulations carried out? Recall that the VECM in (5)-(8) is a dynamic system that includes separate dynamic error-correction equations for first-differences of each of the four variables: (the log of) real average finding cost, the level of technology, annual reserve additions, and a depletion proxy. As mentioned above, the VECM treats all variables as potentially endogenous. When using the VECM to forecast, it provides time paths for all four variables. We focus below on the forecasts for the technology variable (panel A), log of real average finding costs (panel B), and annual reserve additions (panel C). For brevity, only the results based on the QL model are shown. Those based on the TV model are available for the author. [Appendix attached, but not for publication.]

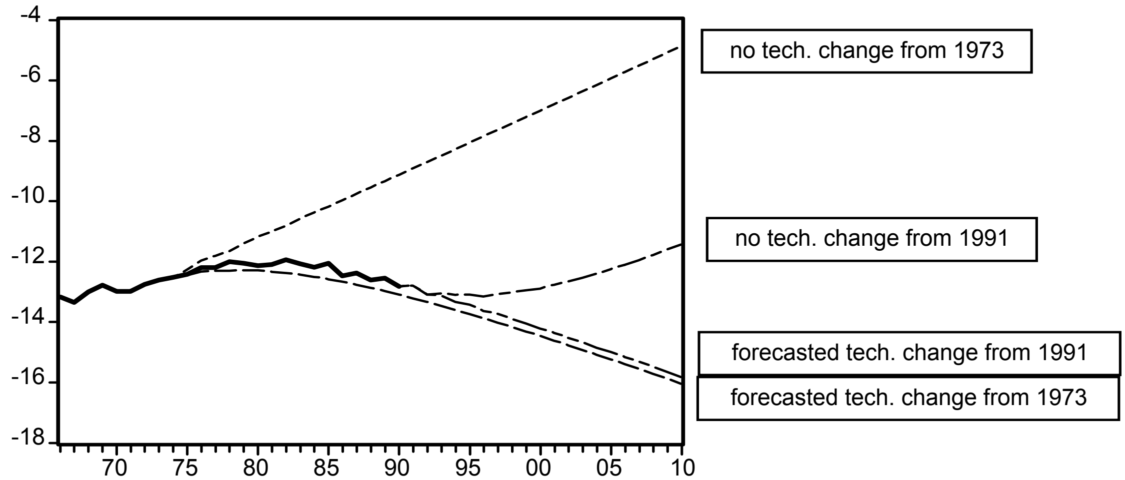
Fig.1 shows in-sample forecasts of these variables for natural gas from 1973, as well as out-of-sample forecasts from 1991 through 2010. (Actual average cost over the 1967-90 period is also shown for comparison.) The model predicts an ongoing decline in average finding cost of natural gas from the mid-1980s through the year 2010. Technological advance continues at roughly its historical rate, while annual reserve additions trend slightly downward.

To examine the impact of technological change on average finding costs, the baseline forecasts just described can be compared to those from scenarios where technological improvements do not occur after 1973, or alternatively 1991. The "no tech. change from 1973 [1991]" scenarios is generated by replacing the estimated change in technology equation from the VECM ((6) above) with: $\Delta TECH_t = 0$ as of 1973 [1991]. This change in specification for the technology equation then affects the forecasted time path for average finding costs, both directly and indirectly through the change in reserve additions and cumulative reserve additions (depletion) in the VECM. Thus, using the VECM for counterfactual simulations, we can avoid making the assumption used in CM (2001) that the annual level of discoveries remains unchanged regardless of what happens to the rate of technological change.

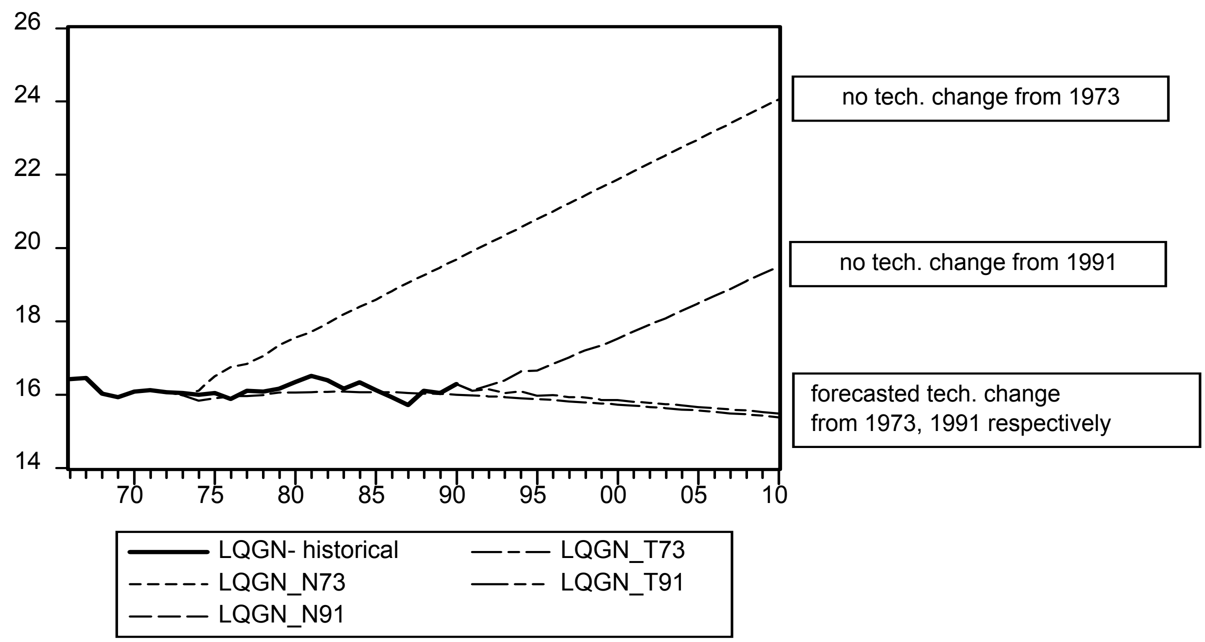
Figures 1A, 1B and 1C
Nonassociated Natural Gas: Quality Ladders Model



Average Finding Cost of Nonassociated Nat. Gas Forecasts based on QL Model



Annual Reserve Additions for Nonassociated Nat. Gas Forecasts based on QL Model



First consider natural gas. Comparison of the corresponding with and without technological change scenarios shows the extent of technology's impact on finding costs. Our 'with technological improvement' forecast from 1973 implies a gradual decline in average finding costs of 2.9% per year averaged over the 1973-90 period. (In contrast, the actual historical series shows an average decline of 1.2% per annum.) In contrast, the 'no-tech change after 1973' scenario predicts that average finding costs would have had to rise at an average rate of 20.4% per year. In part, this reflects a rather substantial increase in the rate at which new reserves are proven up (as Fig. 1C shows). Clearly, technological change has been instrumental in ameliorating the effects of increasing resource scarcity over time.

Fig.2 reports analogous "with tech" and "no-tech" simulations for crude oil.[16] Forecasting forward from 1973 assuming ongoing technological progress, the model predicts an increase in average finding costs of modest 0.7% per annum, averaged over the 1973-90 period. See fig. 2B (For the same period, actual findings costs fell by 1.8%, albeit with considerable year to year fluctuation.) If we specify that there are no technological improvements after 1973, on the other hand, the model predicts that average finding costs would have risen at an average annual rate of 15.4% over the 1973-90 period. Again, technological advance again has a significant effect in offsetting increasing scarcity over time for U.S. crude oil reserves.

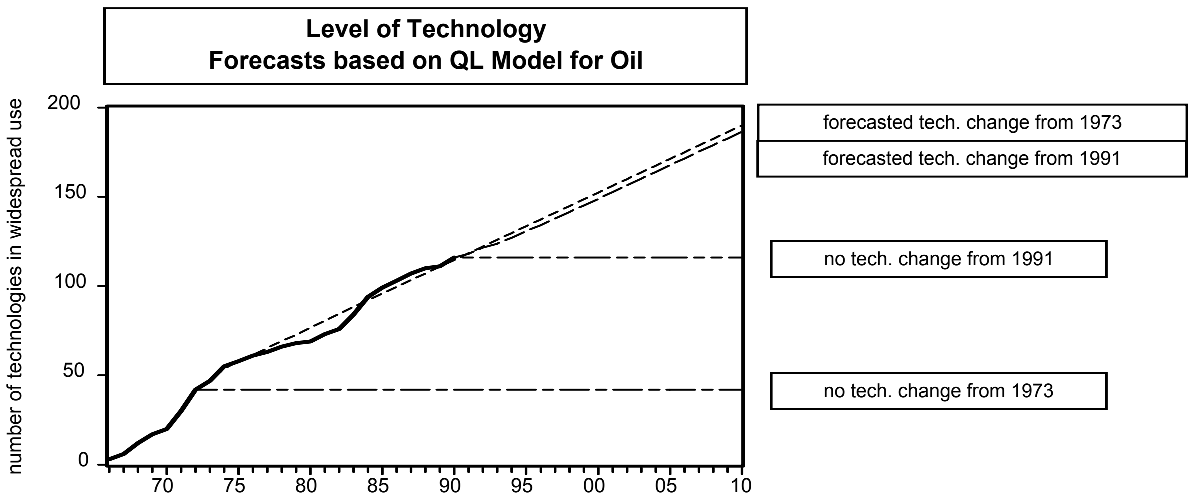
Comparing the growth rate figures reported above, or inspecting Figs. 1B and 2B, for natural gas and crude oil, an interesting conclusion emerges: technological change appears to have had a more significant impact reducing the E&D costs for natural gas than it has had for crude oil. Upon reflection, this is not surprising. Industry experts agree that the oil resource

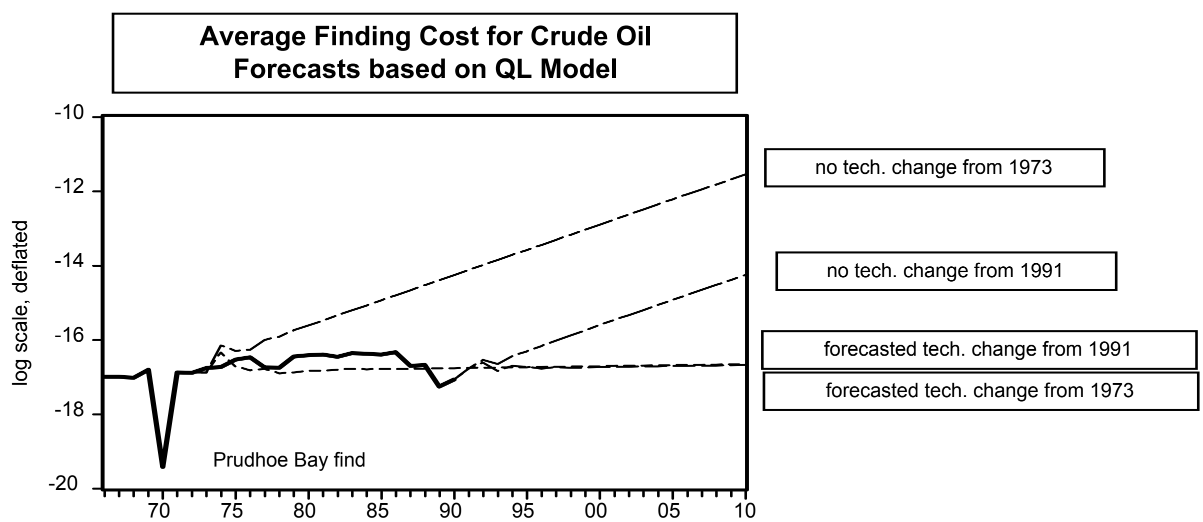
---

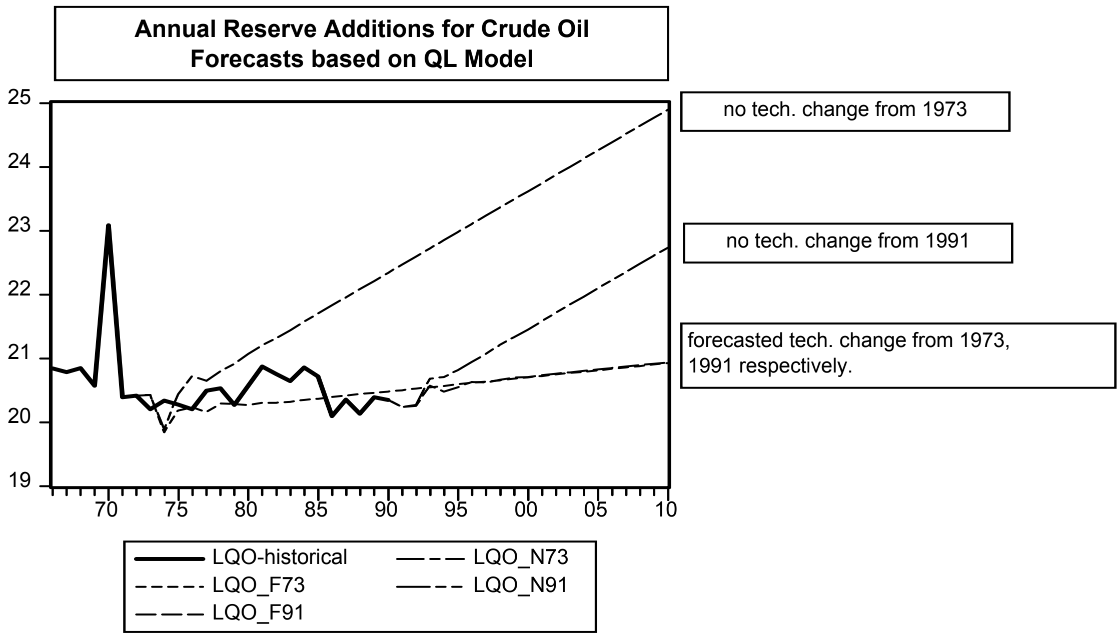
[16] It is reassuring that the forecasted technology paths are very quite similar regardless of whether they are obtained from the gas or oil VECMs. Compare the forecasts for $TECH_t$ in Fig. 1A (gas) to Fig. 2A (oil).

base has been more intensively explored than that for gas. The advanced stage of depletion of the oil resource base, therefore, may make it more difficult for technological improvements to offset increases in finding costs over time.

Figures 2A, 2B, and 2C:
Crude Oil: Quality Ladders Model







VI. CONCLUSION

This study contributes to the literature on estimating finding cost functions for nonrenewable resources by arguing that the previous literature ignores a key feature of the variables employed: theory suggests and unit root tests support the claim that they are (almost certainly) nonstationary. It argues, therefore, that appropriate time series methods must be employed to avoid spurious regressions problems that potentially plague the existing literature.

Using the Moss (1993) index of technological change in exploration and development (E&D) activities, the paper estimates a vector error correction model (VECM) to capture how average finding costs depend on technological change and depletion over time. The effects of depletion and technological advance on average finding costs are successfully isolated in the estimated VECMs. The VECM is then used to address a key, but under researched, empirical question: to what extent has ongoing technological change offset the effect of ongoing depletion on the costs of finding additional reserves of natural gas and crude oil in the U.S. petroleum industry? By comparing simulations with “no technological change” and historical rates of technological change, we show that technological change has played a key role in prevented what would otherwise have been a sharp rise in the finding cost of additional *natural gas* reserves. Moreover, it enabled firms to prove up additional reserves at a slower rate, presumably by reducing the uncertainty about whether additional reserves could indeed be discovered as needed. The impact of technological change on finding costs for crude oil, which is a much more mature sector in terms of U.S. exploration, has been more modest.

REFERENCES

Adelman, Morris A. 1991. "U.S. Oil/Gas Production Cost: Recent Changes," *Energy Economics*, October, 235-237.

American Gas Association. 1990. *Changes in Natural Gas Recovery Technology and Their Implications*. Arlington, Virginia.

American Petroleum Institute. 1970. *Organization and Definitions for the Estimation of Reserves and Productive Capacity of Crude Oil*, Technical Report Number 2, (June), p.19. Washington, D.C.

American Petroleum Institute. 1992. *Basic Petroleum Data Book*, XII, 3 (September), Section V: Tables 9 and 9a. Washington, D.C.

Barnett, Harold J. and Chandler Morse. 1963. *Scarcity and Growth: The Economics of Natural Resource Availability*. Baltimore: Johns Hopkins University Press.

Berndt, E.R. 1991. *The Practice of Econometrics: Classic and Contemporary*. Ch.3: Costs, Learning Curves, and Scale Economies: From Simple to Multiple Regression," Reading, Mass: Addison-Wesley Publishing Co.

Bohi, Douglas R. and Michael A. Toman. 1984. *Analyzing Nonrenewable Resource Supply*, Washington, D.C.: Resources for the Future.

Bohi, Douglas R. 1999. "Changing Productivity in U.S. Petroleum Exploration and Development" in R. David Simpson (ed.), *Productivity in Natural Resource Industries: Improvement Through Innovation*, Washington, D.C.: Resources for the Future.

Cuddington, John T. and Diana L. Moss. 2001. "Technological Change, Depletion and the U.S. Petroleum Industry," *American Economic Review* 91 (4), 1135-1148.

Dasgupta, Partha S. and Geoffrey M. Heal. 1974. "The Optimal Depletion of Exhaustible Resources," *Review of Economic Studies*, 3-28.

Devarajan, Shantayanan and Anthony Fisher. 1982. "Exploration and Scarcity," *Journal of Political Economy* 90(6), 1279-1290.

Devarajan, Shantayanan and Anthony Fisher. 1981. "Hotelling's 'Economics of Exhaustible Resources': Fifty Years Later," *Journal of Economic Literature,* 19, 65-73.

Epple, Dennis N. 1975. *Petroleum Discoveries and Government Policy*, Cambridge, Mass: Ballinger Publishing Company.

Epple, Dennis N. and J. Londregan. 1993. "Strategies for Modeling Exhaustible Resource Supply," in A.V. Knees and J.L. Sweeney (eds.) *Handbook of Natural Resource and Energy*

Economics, Vol III. Amsterdam: Elsevier.

Franses, Philip Hans and Gary Koop. 1998. "On the Sensitivity of Unit Root Inference to Nonlinear Data Transformations," *Economic Letters* 59, 7-15.

Granger, C. W. J. and Paul Newbold. 1974. "Spurious Regressions in Econometrics," *J. Econometrics* 2, 111-120.

Hall, Alastair. 1994. "Testing for a Unit Root in Time Series with Pretest Data-Based Model Selection," *Journal of Business and Economic Statistics,* 12(4), 461-470.

Hansen, Henrik and Katarina Juselius. 1995. *CATS in RATS: Cointegration Analysis of Time Series*. Evanston, Illinois: Estima, Inc.

Hamilton, James D. 1994. *Time Series Analysis*. Princeton, N.J.: Princeton University Press.

Hotelling, Harold. 1931. "The Economics of Exhaustible Resources," *Journal of Political Economy,* 39, 137-175.

Kamien, Morton I., and Nancy L. Schwartz. 1978. "Optimal Exhaustible Resource Depletion with Endogenous Technical Change," *Review of Economic Studies,* 45, 179-196.

Kortum, S. and Saul Lach. 1995. "Patents and Productivity Growth in U.S. Manufacturing," working paper.

Livernois, John R. 1988. "Estimates of Marginal Discovery Costs for Oil and Gas," *Canadian Journal of Economics* 21(2), 379-393.

Livernois, John R. and Russell S. Uhler. 1987. "Extraction Costs and the Economics of Nonrenewable Resources," *Journal of Political Economy* 95(1), 195-203.

Moss, Diana L. 1993. "Measuring Technological Change in the Petroleum Industry: A New Approach to Assessing Its Effect on Exploration and Development," National Economic Research Associates, Washington, DC, (October) working paper #20.

National Petroleum Council. 1966. *Impact of New Technology on the U.S. Petroleum Industry: 1946-1965*, Washington, D.C.

National Petroleum Council. 1987. *Factors Affecting U.S. Oil and Gas Outlook*. Washington, D.C.

National Petroleum Council. 1992. *The Potential for Natural Gas in the United States: Executive Summary*. Washington, D.C.

Ng, Serena and Pierre Perron. 1995. "Unit Root Tests in ARMA Models with Data-Dependent Methods for the Selection of the Truncation Lag," *Journal of the American Statistical*

Association, 90, 429, 268-281.

Pindyck, Robert S. 1978. "The Optimal Extraction and Production of Nonrenewable Resources," *Journal of Political Economy* 86 (5), 841-861.

Slade, Margaret. 1982. "Trends in Natural-Resource Commodity Prices: An Analysis of the Time Domain," *Journal of Environmental Economics and Management,* 9, 122-137.

Solow, Robert M., 1974. "The Economics of Resources or the Resources of Economics," *American Economic Review* 64(2), 1-14.

Stock, James H. and Mark W. Watson. 1993. "A Simple Estimator of Cointegration Vectors in Higher Order Integrated Systems," *Econometrica* 61, 4 (July), 783-820.

Stoneman, Paul. 1983. *The Economic Analysis of Technological Change*, Oxford: Oxford University Press.

Tippee, Bob and Bob Beck. 1991. "Costs of Adding Reserves Sliding in the U.S." *Oil and Gas Journal Special* (April 8), 45-48.

Uhler, Russell S. 1979. *Oil and Gas Finding Costs*. Calgary: Canadian Energy Research Institute.

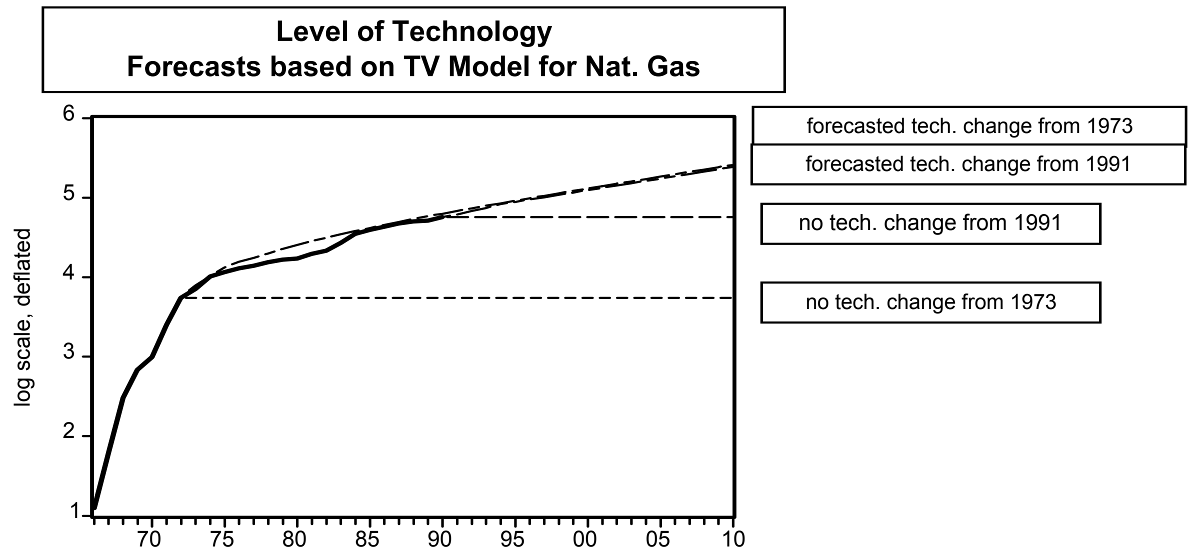
United States Congress, Office of Technology Assessment. 1985. *Oil and Gas Technologies for the Arctic and Deepwater*.

Walls, Margaret. 1992. "Modeling and Forecasting the Supply of Oil and Gas: A Survey of Existing Approaches," *Resources and Energy* 14, 3, 287-309.

Williamson, H.F., and A.R. Daum. 1959. *The American Petroleum Industry, The Age of Illumination 1859-1899*, Evanston, Ill.: Northwestern University Press.

APPENDIX
(not for publication; available on request)

Figures 3A and 3B:
Nonassociated Natural Gas: Technology Varieties Mode l



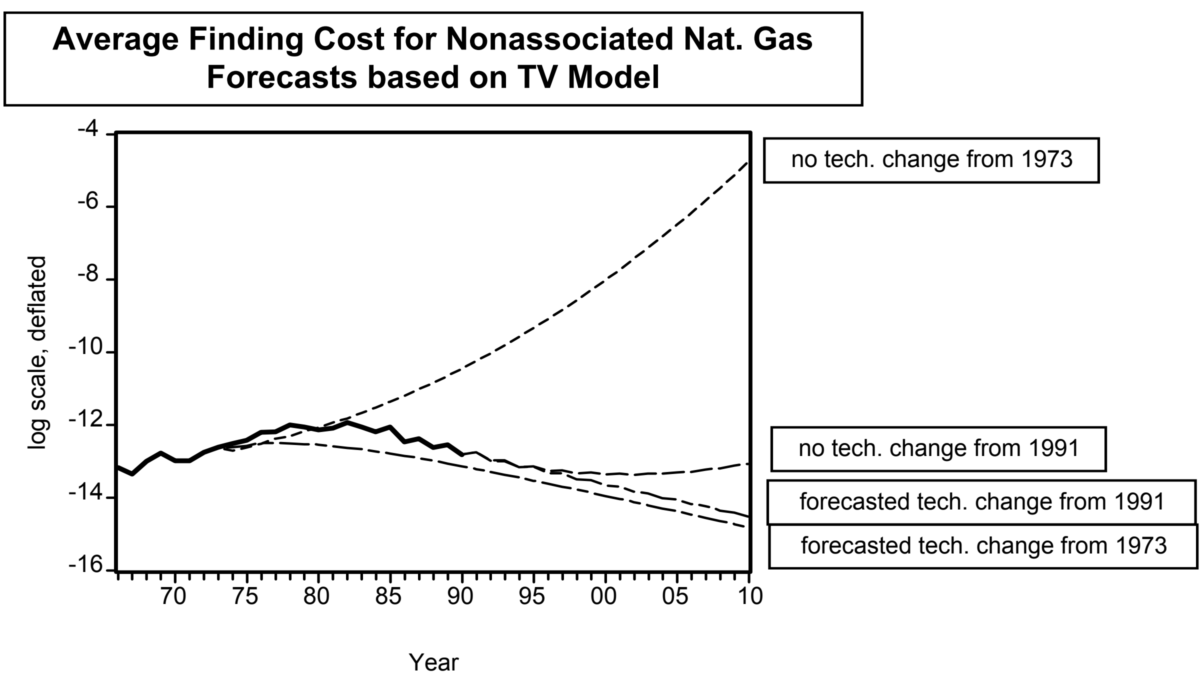
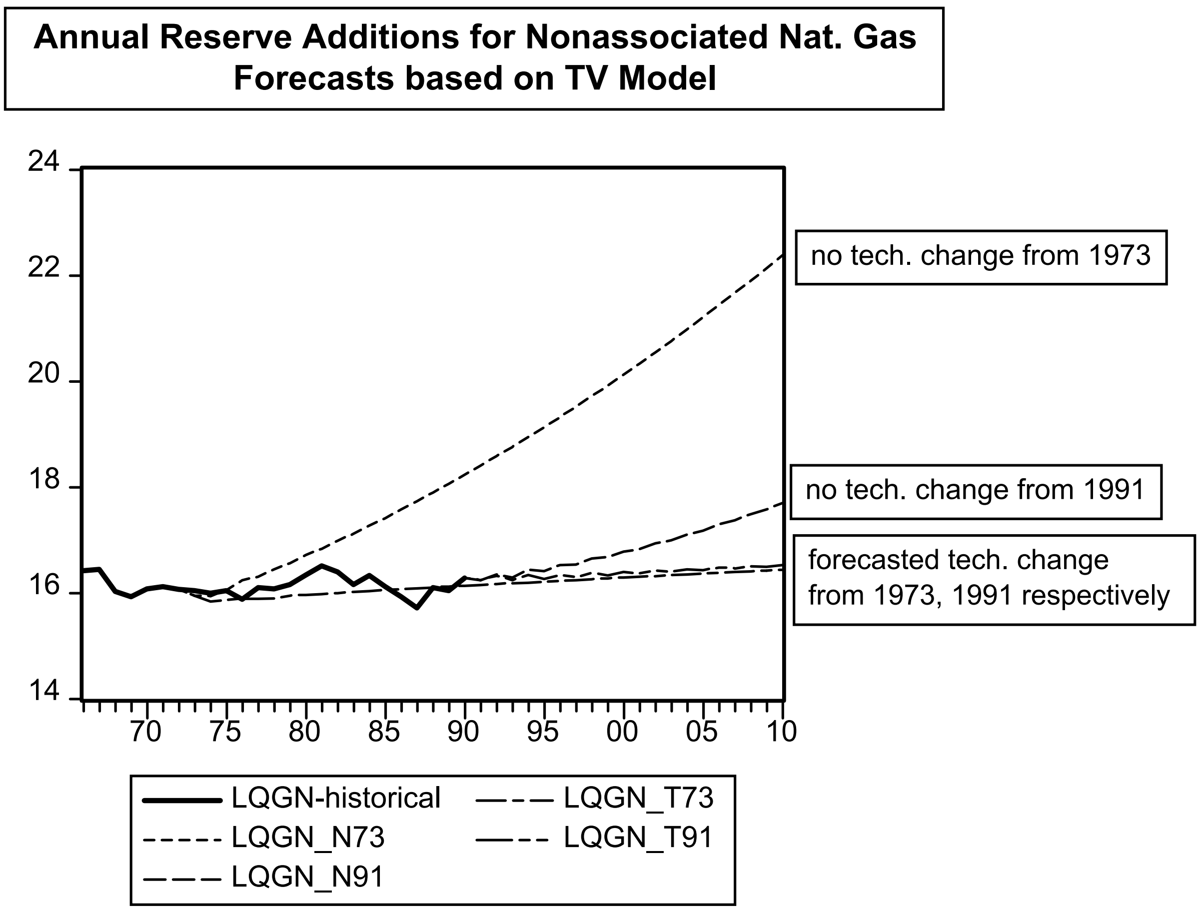
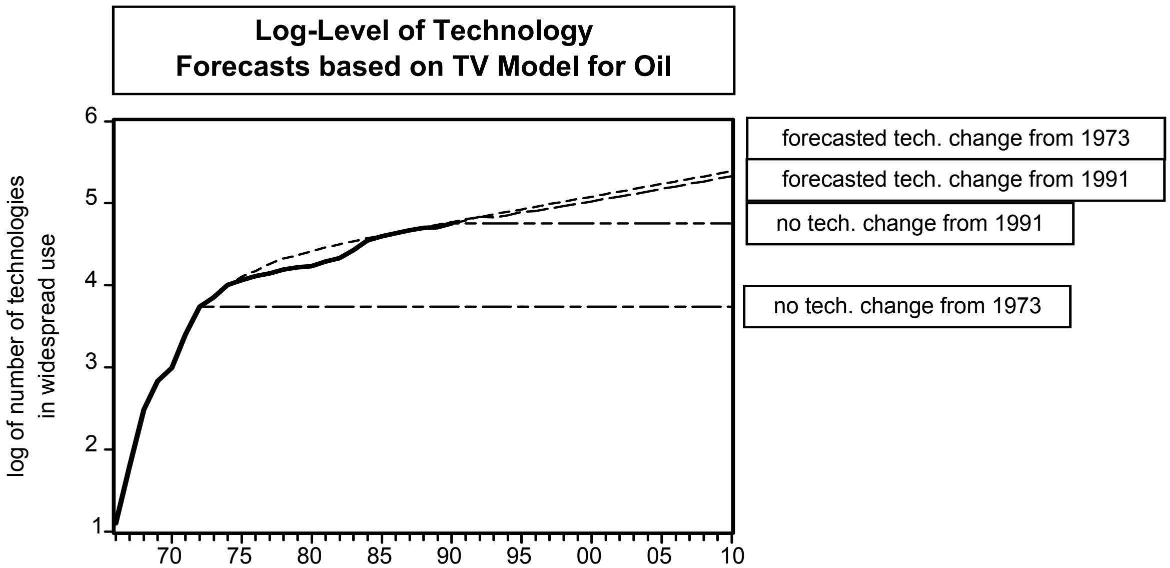


Fig. 3C



Annual Reserve Additions for Nonassociated Nat. Gas
Forecasts based on TV Model
24
22
20
18
16
14
70
75
80
85
90
95
00
05
10
no tech. change from 1973
no tech. change from 1991
forecasted tech. change
from 1973, 1991 respectively
LQGN-historical
LQGN_T73
LQGN_N73
LQGN_T91
LQGN_N91

Figures 4A and 4B:
Crude Oil: Technology Varieties Model



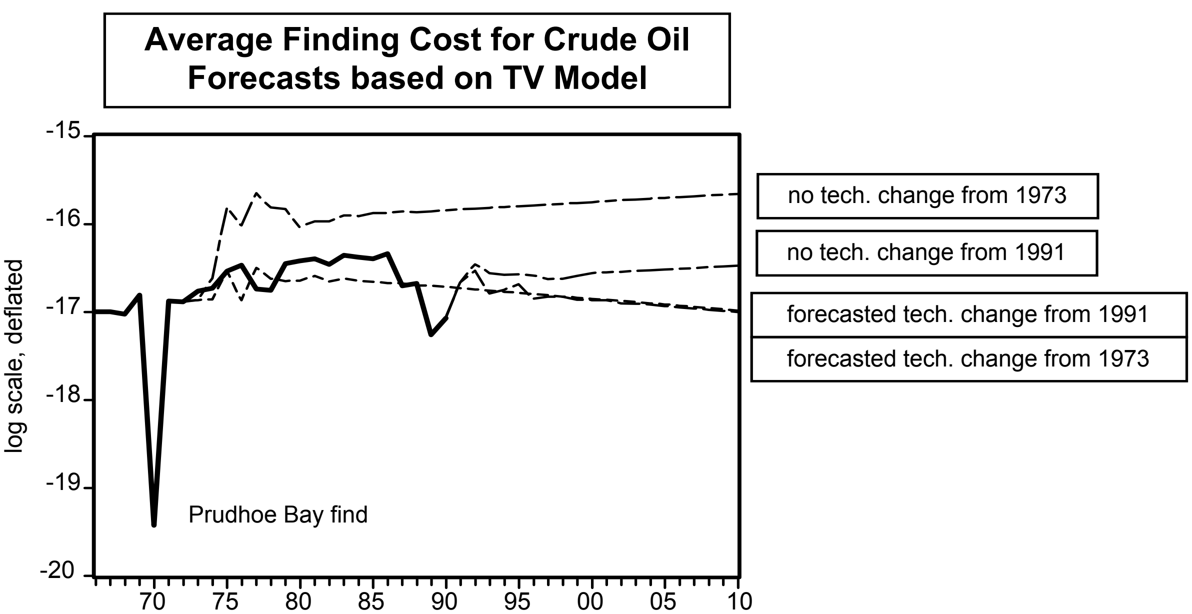
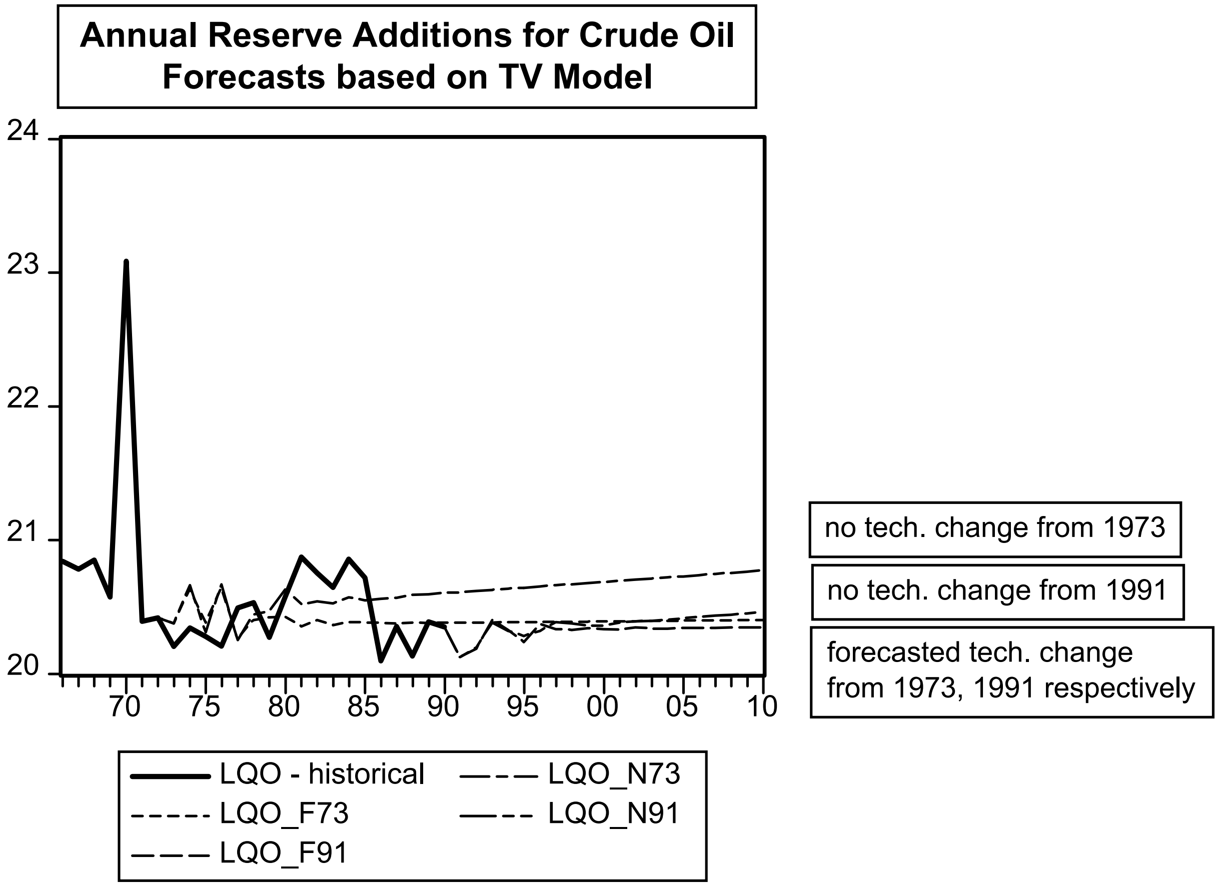


Fig. 4C


Annual Reserve Additions for Crude Oil
Forecasts based on TV Model
24
23
22
21
20
70
75
80
85
90
95
00
05
10
no tech. change from 1973
no tech. change from 1991
forecasted tech. change
from 1973, 1991 respectively
LQO - historical
LQO_N73
LQO_F73
LQO_N91
LQO_F91