Journal of Quantitative Economics, Vol. 13, No. 2 (July 1997) 29 - 44

# AN ECONOMETRIC MODEL OF EXPLORATION AND EXPLOITATION OF HYDROCARBON

M. PAL AND P. MAITI[1]

*Economic Research Unit, Indian Statistical Institute*
*203 B. T. Road, Calcutta 700 035, West Bengal*

An Econometric Model of exploration and exploitation of hydrocarbon for estimation of Discovery and Production Costs has been presented in this paper. This is one of the two approaches developed for the purpose and is based on certain econometric relationships estimated on the basis of time series data on cumulated values of relevant variables like Reserve, Costs of Exploration and Development Drillings etc. The estimated model would enable us to work out and predict the Marginal Costs (MC) of discovery as well as the Exploratory Discovery Index (EDI) representing the expected discovery per meter of drilling, besides the discovery and production costs.

## 1. INTRODUCTION

One of the most essential commodities in the modern industrial world is petroleum. Over the last hundred years the importance of petroleum as a source of energy has increased tremendously and today it is the principal source of energy all over the world. With the OPEC price hike in 1973 the attention of the world was promptly focussed on exploration of new hydrocarbon reserves and management of reserves keeping in mind the essentially nonrenewable nature of this principal source of energy. As a result evolving appropriate policies for exploitation of hydrocarbon has been given much more importance than before. A major ingredient of any such policy is the knowledge on the costs of hydrocarbon. It is well known that the total cost of hydrocarbon comprises a number of components - viz; the cost of exploration and discovery of hydrocarbon, the cost of development of the fields and the cost of production or exploitation from the oil-fields.

[1] The paper is based on ISI-ONGC project entitled "Estimation of Discovery and Production Costs of Hydrocarbon With Some Applications to Indian Data" Completed in 1990. The authors gratefully acknowledge the ONGC for providing the data necessary for the study reported in this paper. The name of the basin, for which this exercise has been carried out, and other particulars of different data can not be disclosed. The ISI interdisciplinary team in charge for the work consisted of Professors D. Coondoo, A.C. Mukhopadhyay, T.S. Arthanari and R. Mukherjee apart from the present authors. Professor J.K. Ghosh acted as a consultant for the team. The authors also thank the referee for valuable comments made to improve the paper to the present form.

Further, given the uncertainty involved in the discovery of hydrocarbon reserves, estimating the cost of hydrocarbon becomes a challanging issue to energy researchers. In this paper, an econometric model and method of estimation and forecasting of the cost of hydrocarbon is presented.

For the purpose of developing our econometric method of estimation and prediction of cost of hydrocarbon, the hydrocarbon production has been viewed here as an integrated process starting with the geological and geophysical surveys followed by exploration and discovery and leading ultimately to actual extraction of oil and gas. Briefly, the process of hydrocarbon production can be classified into two broad stages-viz; exploration and exploitation. Each of these in turn is subdivided into a number of distinct steps. Thus, exploration starts with the geological and geophysical surveys aimed at primary detection/identification of locations of exploratory drilling. Once exploratory drilling is undertaken, a discovery is achieved if such a drilling strikes oil/gas. A discovery in turn is followed by a number of delineation drillings/wells in the vicinity of the discovery well in order to estimate the volume of hydrocarbon in the discovered pool. If the pool discovered is found to be commercially viable, depending on the demand for oil and the policy of exploitation, development production wells are dug from which actual production of oil/gas takes place. Thus, the cost of production have two broad components - the discovery cost comprising and the cost of exploitation.

The whole process of exploration and exploitation of hydrocarbon is shown in the accompanying diagram 1.1.

In what follows, we present our proposed econometric model for estimation and prediction of cost of hydrocarbon along with some quantitative results obtained on the basis of data for one of the explored basins in India. Section 2 also includes the statistical model for the sake of comparison and completeness. Section 3 deals with the data and measurement of variables. Section 4 discusses the results and their policy implications.

## 2. THE TWO MODELS

The main difficulty in estimating discovery cost and hence cost of production arises due to the uncertainty of outcomes of the exploration activity. It is not known with certainty if any exploration activity will result in a success or a failure and even if an exploration leads to a discovery, the amount of reserve to be discovered is unknown apriori. Keeping this in mind, we have developed two approaches for estimation of cost of exploration and exploitation of hydrocarbon. It should be mentioned at the outset that the approaches proposed in this study are applicable for areas where some explorations have been made. In the cases of virgin areas *i.e.*, where no exploration has been made, our models will not be applicable directly unless one is prepared to use the information available from other partially explored ares considered comparable on grounds of geological and geophysical similarities.

### 2.1 *The Econometric Model*

The econometric model which we have specified and estimated here has two modules - viz; (i) a discovery cost model which establishes the link between the volume of reserve

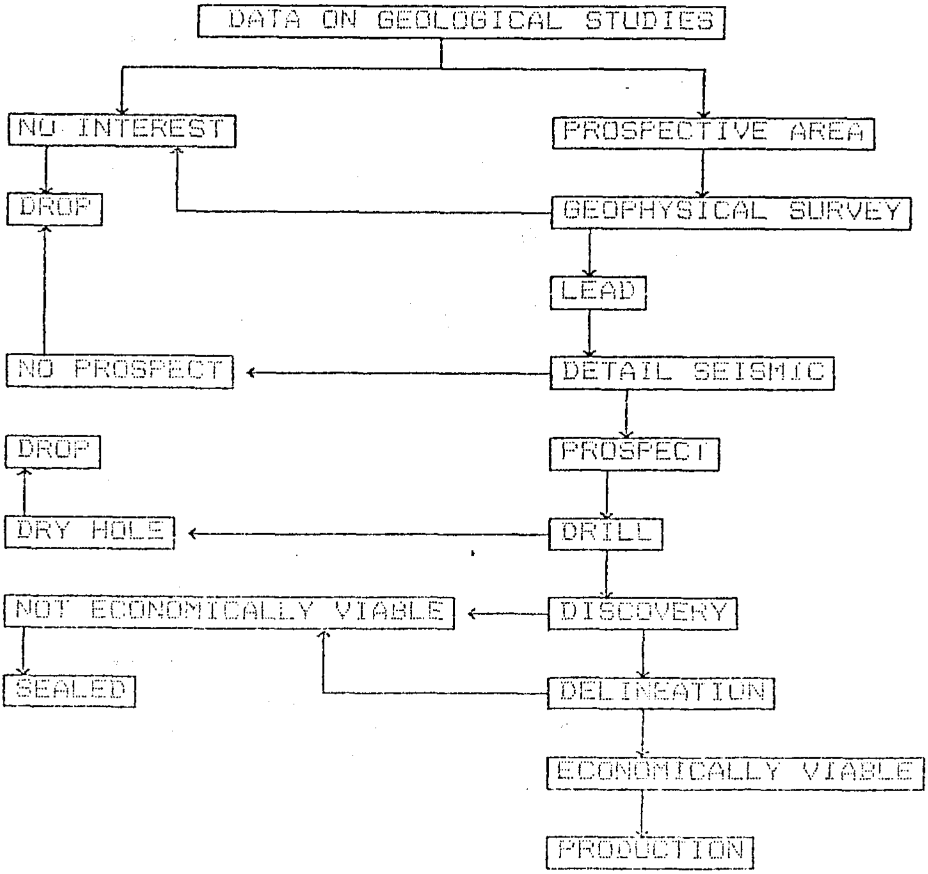

Diagram 1.1 : A Diagrammatic Presentation of the Process of Exploration and Exploitation

discovered and the corresponding costs of surveys and drilling of exploratory and delineation wells, and (ii) a production cost model which links up production costs to production and reserve discovered.

The model specifies functional relationships for each activity. More specifically it was assumed that cumulated survey cost (CS) is proportional to cumulated exploratory cost and cumulated reserve (R) is determined by the cumulated cost of completed exploratory wells (CE), and this R, in turn, determines the cumulated cost of completed delineation wells (CDL). The specification of the last two relationships was based on the postulate that exploratory drillings help discover reserves of hydrocarbon and once a discovery has been made the number of delineation drillings necessary for ascertaining the size of the discovered reserve depends on the size of the reserve itself.

To estimate the production cost as defined earlier, one must take into account the production activities subsequent to exploration and delineation. The proposed methodo-logy specifies functional relationships among the variables which are closely related to the volume of hydrocarbon in a basin, viz; relationship between (i) cumulated volume of reserve (R) and cumulated development cost (CDV), (ii) cumulated volume of production (Q) and cumulated development cost and (iii) cumulated volume of production and operating/production maintenance cost (CMC). Given these estimated relationships, the profile of production cost corresponding to a future plan of production can be estimated.

The cumulated total exploratory cost may be defined to be

$$C = CS + CE + CDL \tag{1}$$

where, CS = Total Cumulated Survey Cost

CE = Cumulated Cost of Completed Exploratory Wells

CDL = Cumulated Cost of Completed Delineation Wells

The cost of reserve relations defined earlier may be thought to relate $C$ to $R$, the cumulated reserve. In view of the logical causal relations described in the previous paragraph. It is $CS + CE$ which precedes discovery. Thus, the relations $CS + CE$ versus R may be thought to be more meaningful.

The treatment of CS in our model is, however, different from what is thought to be logically the best. CS has been treated as an endogenous variable rather than an exogenous variable. This is because the amount of survey cost is not much if not neglible compared to other costs. Also, the time series data of survey costs were found for a region which does not exactly correspond to the basin. We had to extrapolate the regional data to get the basin-wise data. Again, the time series data on survey costs supplied to us do not cover the whole region under consideration. If we assume the exploratory dirilling to depend on the survey costs, the estimate will be very poor and subsequent analysis and forecasts on the basis of these results will be more unreliable. Thus the present model is based on the postulate that CE (and not CS+CE) determines R and R, in turn, determines CDL.

Many form of the Reserve Relationsip have been proposed so far. Amongst these, the logistic model by Hubbert (1962, 1974), the three parameter Gompertz curve by Moore (1970), Wiorkoski's (1981) model combining exponential, logistic and Gompertz curves: Arps, Mortada and Smith's model (1971), etc. deserve special mention. We have in our analysis taken the following form.

$$R = K - ab^{CE} \tag{2}$$

which is a considerably versatile form and can accommodate a saturation value of $R$ equal to $K$ as $CE$ is increased indefinitely. In other words it gives an estimate of maximum possible reserve which can be discovered in the basin. This form also conforms with the property that larger pools of reserves are likely to be discovered first.

It should be reported here that many alternative forms of the reserve relationship were tried. These include linear, semilog, double-log and exponential. Most of the forms do not possess the desirable properties a reserve function should have. Moreover, the modified exponential form turned out to be the best in terms of the correlation coefficient measured between observed and corresponding estimated $R$-values (i.e.; $r(R, \hat{R})$).

Among the alternative forms for the delineation cost relation (i.e; linear, semilog, exponential and doublelog), linear was found to give best fit. Thus,

$$CDL = c + dR \tag{3}$$

The survey cost relation is

$$CS = \alpha CE \tag{4}$$

It was observed that the year wise survey cost as a proportion of the corresponding yearwise cost of exploratory driling was fairly stable over years (the median value of this proportion being approximately 0.21). We, therefore, have assumed CS = .21 CE in our empirical analysis. This simplifying assumption about CS was in a sense unavoidable, since the exact nature of relationship between CS and R could not be specified apriori.

It should be pointed out that since the time variable does not appear explicitly in the present formulation of the discovery cost model, there is virtually no scope for introducing time lags in the reserve and cost functions specified above. However, as soon as the model and the corresponding reality as revealed by the observed historical data on costs and reserve, are juxtaposed, the question of introducing such time lags come in.

To elucidate the issue, let us examine the situation in the light of time series data observed at each time point of discovery. Consider the time interval between any two successive discoveries. A number of exploratory wells are completed in this time interval. Some of these prospects will be declared as dry before the next discovery and others will continue to be explored. The prospect which is declared dry must have been started approximately $\lambda$ years back where $\lambda$ is the average time lag for exploration. Strictly speaking, the exploratory cost to be associated with the volume of next discovered reserve should be the cost of all exploratory wells in prospects for which the decision regarding absence of reserve is taken within this time interval plus the exploratory cost of the prospect for which next discovery is made. Thus the cumulated exploratory cost that should correspond to the cumulated reserve $R(t)$ at discovery point $t$ should be $CE(t - \lambda/2)$, *i.e.*, the cumulated cost upto time point $t - \lambda/2$ and not $CE(t)$. The question of time lag in the relationship between cumulated delineation cost and reserve can also be discussed in a similar manner.

In the present exercise we have not taken into account such time lags in the reserve/cost relation primarily on two considerations. First, here we visualize the process of discovery of reserve in a basin as a continuous process. Thus, so long as discovery effort in a basin is not going to be suspended in the immediate future, it should not matter much if the exploratory cost assignable to one discovery is associated with the ones immediately preceding it. This apart, even if we opted for the introduction of time lags in the analysis, it would be difficult to ascertain the fixed lengths of this lag, since nothing is known as to how the costs of abortive explorations could be apportioned between successive discoveries.

Various plausible causal relationships may be thought of and estimated which depict the nature of dependence between relevant variables to highlight various aspects of the process of production. Amongst these are Development cost vs. reserve, development cost vs. delineation cost, production vs. balance of recoverable reserve, production vs. development cost and delineation cost, production vs development cost, production maintenance cost vs. development cost, components of production maintenance cost vs. total production maintenance cost etc. deserve special mention. Many of these relationships do not appear to be convenient for use in forecasting exercises, either because they lack enough explanatory power (as reflected by the respective goodness of fit measure $R^2$), or they portray a fairly involved mechanism of the determination of production. We therefore sought to establish simpler, yet statistically stable, relationships between the relevant cumulated variables. Three causal relationships were ultimately used in the production model. These are development cost relation which gives the cumulated development cost given a level of cumulated reserve. Maintenance cost relation which gives the cumulated maintenance cost as a function of cumulated development cost and production relation giving production as a function of cumulated development cost.

A number of alternative two parameter functional forms for each the three relations were tried out. Among these, the linear form turned out to be the best fitting form in all the cases. The intercept in the maintenance cost relation, moreover, was not found to be significant. Hence the equation was reestimated (by least squares method) assuming the proportional form. Below we present the three forms for the sake of completeness.

Development Cost Relation : $CDV = \alpha + \beta R$ (5)

Maintenance Cost Relation : $CMC = \gamma CDV$ (6)

Production Relation : $Q = \Theta + \delta\, CDV.$ (7)

For a better comprehension of the complete econometric model, let us examine briefly the causal relations involved in linking exploratory, discovery and production and the corresponding costs. This is presented in a nut shell by the following Diagram 2.1.

### 2.2. *The Statistical Model*

In contrast to the econometric approach, the statistical model explicitly takes into account probabilistic nature of the outcome of the exploratory activity. Specific forms of distribution of the relevant random variables have been assumed for the purpose of the statistical modeling of the problem and a simulation method is proposed for its solution. The essential idea behind the statistical formulation can be explained as follows.

It has been assumed that the basin consists of a set of $N$ fields which constitutes a random sample from the super-population. $N$ may be treated as known, an unknown parameter or a random variable following a suitable probability distribution.

The chronological list of the results of exploratory wells drilled give rise to a 0-1 sequence, "0" denothing a failure and "1" denoting a success. Considering only the success points we have a new sequence $\{y_i\}$, $i \geq 1$, where $y_i$ represent the size of the $i$-th discovered field. $y_1, y_2, \ldots, y_n$ are thus the sizes of discovered fields. The fields with higher sizes had the higher probability of being drawn. In other words, the sampling

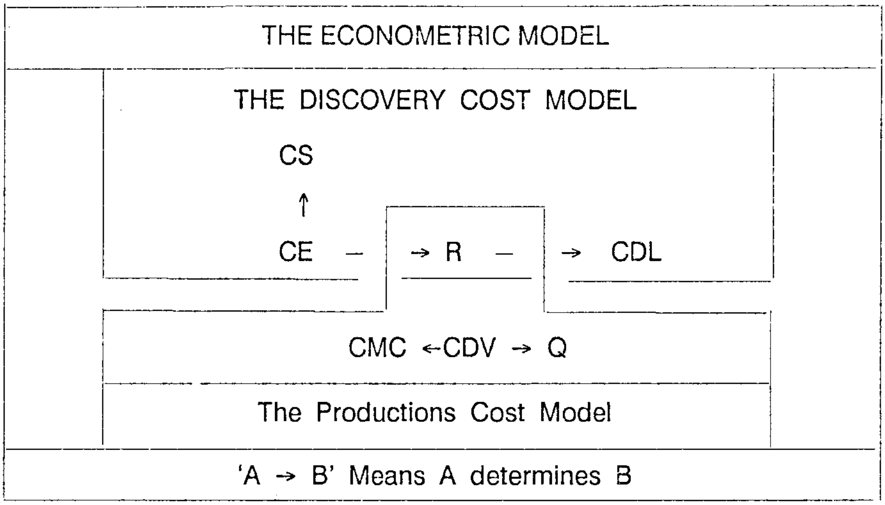

Diagram 2.1 The Causal Relationships in the Econometric Model

scheme is probability proportional to size without replacement (ppswor). The production plan except for the starting point can be determined quite accurately and it more or less depends only on the size.

Form the above assumptions joint probability distribution of $y_1, y_2, \ldots, y_n$ in the given order can be obtained as

$$K \prod_{i=1}^{n} y_i \, f(y_i \mid \theta) \sum_{i=1}^{n} \int_0^\infty A_i \exp(-\lambda b_i) \{L(\lambda \mid \theta)\}^{N-n} \, d\lambda$$

where $K = N!/(N-n)!$, $A_i = \left\{ \prod_{j=1(\neq i)}^{n} (b_i - b_j) \right\}^{-1}$ and

$$L(\lambda \mid \theta) = \int_\theta^\infty \exp(-\lambda y) \, f(y \mid \theta) \, dy.$$

Also $b_i = \sum_{i=1}^{n} y_i$ and $f(y_i \mid \theta)$ is the probability density function of field size distribution in the super population.

Regarding the 0-1 sequence, let $w_n$ represent the number of wells drilled following the $n$-th success and upto and including the next success. The appropriate distribution of $w_n$ is taken to be the geometric as data suggested. After ignoring the small and marginal fields, $N$ is estimated as

$$\hat{N} = (\text{Number of Prospects}) / \hat{p}$$

where $p$ is the probability of success.

Given $w_n$ and $y_{n+1}$, the additional cost needed to make an additional discovery of size $y_{n+1}$ is given by

$$T_{n+1} = c_1 w_n + c_2 y_{n+1},$$

where $c_1$ is cost of an exploratory well and $c_2(x)$ is a linear function of x giving the cost of delineation of a field of size x. The simulation techniques were used because $y_{n+1}$ is a random variable and the conditional expectation is difficult to obtain. We assume that the conditional distribution of $y_{n+1}$ given $y_1, y_2, \ldots, y_n$ to be Gamma. The next step is to treat $y_{n+1}$ as the realized $(n+1)-th$ field size and repeat the above with $n$ replaced by $n+1$. The whole sequence of operations can be repeated as many times as one likes and the relevant expectations with associated standard deviations can be computed.

## 3. THE DATA AND THE MEASUREMENT OF VARIABLES

Apart from thorough understanding of the process (of exploration and exploitation of hydrocarbon), model building is constrained by the knowledge on the availability of types of relevant data. For the present study a chronological history of all the wells drilled in a basin were made available. To be precise, we have, for each well, the data on the starting date, date of completion, different costs and the total costs of drilling, type of well (*i.e.*, whether exploratory or development/production well), depth in metre, whether discovery made or dry in case of exploratory well, if discovery made then the size of the discovery, the prospect/field to which it belongs etc. In addition, data on the yearwise aggregate production in the basin and a future programme of exploration in terms of number of exploratory wells to be drilled in the basin in successive five year periods up to the year 2005 were also available to us.

Given the basic information on chronological wellwise data, we obtained the cumulated exploratory cost and the corresponding cumlated volumes of reserve discovered by adding the costs/reserves of all completed exploratory wells upto specified time points. As regards the time points to be used for cumulation of exploratory cost and reserve data, we have three natural choices - viz; (i) cumulation up to successive time points of discovery, (ii) yearwise cumulation and (iii) cumulation upto time points recording equal increments of exploratory cost at successive levels. Thus, the basic well-level data were arranged in three alternative cumulated forms. Superiority of one arrangement over the others may be examined by taking into account the specific features of the basic data. However, since production data are available yearwise, one can relate this to yearwise cost data only. Thus, we shall report in this paper the results based on yearwise data.

Since wellcost figures are available in nominal terms, for comparability of the yearwise cost figures and to obtain a meaningful cumulation of such costs over time, it was necessary to deflate the well costs for different years. For this purpose, we constructed a series of yearwise well-cost-index using the chain-base principle based on Fishers link index $(I_{t-1}, t)$ for successive pairs of years as defined below:

$$I_{t-1,t} = (L_{t-1,t} \times P_{t-1,t})$$

where $L_{t-1,t} = (\Sigma P_{it}\, q_{it-1})/(\Sigma p_{it-1}\, q_{it} - 1)$

and $P_{t-1,t} = (\Sigma p_{it}\, q_{it})/(\Sigma p_{i\,t-1}\, q_{it})$.

Here $p_{it}$ denotes average cost in year $t$ of wells in the $i$-th depth interval (there being a specified number of such depth intervals) and $q_{it}$ denotes the number of wells in the $i$-th depth interval in year $t$. Once the link indices as defined were computed the chain base index number series was constructed on the basis of them using the relation.

$$I_{0,t} = I_{0,1}\, I_{1,2} \cdots I_{t-1,t}, \qquad t = 1, 2, \ldots;$$

$I_{0,t}$ being a measure of drilling cost during year $t$ relative to that during the base year 0.

Thus we get a time series of yearwise chain base Fisher's index number of well drilling cost of the basin for the period 1963-64 to 1983-84. We then fitted a smooth trend curve to the time series, because we saw some unrealistic fluctuations. On trying alternative forms of trend curves, the exponential form was seen to give the most satisfactory fit.

Another major problem on data had to be resolved before we could build up any series. In order to obtain the two components of the discovery drilling cost, viz; the cost of exploratory drillings and the cost of delineation drillings, it is necessary to clasify the wells in individual fields into exploratory and delineation wells according to the definitions of these two types of wells used in the present analysis. The available data shows only whether a well is exploratory (which includes delineation wells also) or develolpment/production well. The delineation wells in a field could not be readily identified from the available data, because some of the exploratory wells were shown to be producing wells as well. To distinguish between the two types of wells we resorted to the following procedure: All exploratory wells up to the discovery well in a field and all wells completed in prospects where either no discovery could be made or the decision regarding discovery was not yet reached were treated as exploratory wells. All wells in a field completed after the discovery well and marked as exploratory wells were taken to be delineation wells. Also, any well drilled between the discovery well and the first development well and marked as producing well has been treated as a delineation well (under the assumption that the well has been converted later into a developlment well). Following this procedure, out of the 1625 wells completed in the basin before 1.1.1996, 228 were found to be exploratory wells of which 63 were discovery wells and 165 were identified as exploratory wells with no discovery.

Based on this classification of wells into exploratory and delineation wells, we next built up the series of cumulative exploratory drilling cost (CE), delineation cost (CDL) and reserve discovered (R).

Once the fieldwise operating costs were obtained, these were aggregated, over the fields for every year to build up a time series of yearwise maintenance cost, and the series thus obtained was deflated by the price index mentioned earlier to express the yearwise cost figures at 1984-85 prices.

The time series of cumulated production for the basin as a whole was constructed by the following procedure: First, for each producing field for which yearwise production figures were available, the total production of hydrocarbon in each of the years was calculated in units of oil by converting the figure of gas production, if any, into oil equivalent by using the conversion factor (1000 cublic metre of gas ≡ 1 mt of oil). A time series of yearwise production for the basin as a whole was then constructed by aggregating the production in different fields during each year. The series of yearwise production thus obtained was then cumulated over time to give the time series of cumulated production.

## 4. THE RESULTS AND THE POLICY IMPLICATIONS

Table 4.1 gives the estimated equations for the reserve relation.

The estimated equations based on the year-wise data are given below :

1. Reserve Relation : $R = 989.83 - 770.84\ (0.955)^{CE}$, (8)

   $r(R, \hat{R}) = 0.987$

2. Delineation Cost Relation : $CDL = -8.086 + 0.0655\ R$, (9)

   $r(CDL, \widehat{CDL}) = 0.970$

3. Survey Cost Relation : $CS = 0.21\ CE$, (10)

4. Development Cost Relation : $CDV = -10.02 + 0.0455\ R$, (11)

**Table 4.1**

Estimates of the Modified Exponential Form of Reserve Function

$R = K - ab^{CE}$ : 1958-59 to 1985-86

| Type of data | Estimate based on* | Estimated value | | | $r(R, \hat{R})$ |
|---|---|---|---|---|---|
| | | k | a | b | |
| (1) | (2) | (3) | (4) | (5) | (6) |
| Yearwise | Entire data set | 711 | 651 | 0.8728 | 0.960 |
| Yearwise | First three obs. deleted | 990 | 771 | 0.9550 | 0.987 |

* It was noticed that deleting first few observations improved the regression results to a great extent. This is because the first few observations were not very much reliable. So we had to delete first observation, first two observations and so on and fitted the remaining data in each case to get the best fitted curve. This had to be done, because modified exponetial for, especially the estimated value of K as can be seen from the above two estimated relations, is sensitive to each observation. And the whole future plan depends mainly on this value **K.**

$$r\,(\mathrm{CDV},\ \widehat{\mathrm{CDV}}) = 0.95$$

5. Maintenance Cost Relation : $\mathrm{CMC} = 10.79\ \mathrm{CDV}$, (12)

6. Production Relation : $Q = -12.786 + 4.56\ \mathrm{CDV}$, (13)

$$r\,(Q\,,\,\hat{Q}) = 0.99$$

The fitting of each of the equations is highly satisfactory ( $r\,(R\,,\hat{R}) \geq 0.96$ ) . This is quite expected given the cumulative nature of the variables considered. A comparison of the goodness of fit of the functions estimated from these types of data shows that for the reserve function, the magnitude of $r\,(R\,,\hat{R})$ is highest for the yearwise data (when the first three observations of the data set are deleted). The estimates for each of the reserve equations seem quite plausible as it has been verified by subsequent calculations of marginal and average costs of discovery, EDI etc., and by observing the trends of these values, although the results obtained seem to be sensitive to the type of cumulation used. We have however used estimates based on yearwise data for subsequent calculations mainly because the data on production, and future action plan etc. are given yearwise or in some cases five-year-plan-period wise.

Action plan of exploratory activities corresponding to the strategy envisaged by ONGC for the basin is given in the following Table 4.2

Once the estimates of number of exploratory wells (NE) and hence the number of delineation wells (NDL) for each of the five year plan period are known, the estimates of different cumulated costs, marginal costs and EDI etc. can easily be found out. Table 4.2 summarises the picture.

The results presented in the Table 4.3 clearly bring out the rising cost implication of further production effort in the basin. Not with standing the fact that there is an increase in the planned drilling of exploratory wells at a high pace, there is a steady decrease in the discovery of hydrocarbon. Since the basin is already fairly explored, a relatively large marginal effort is necessary for additional discovery, and hence the marginal cost of production is likely to follow a rising trend over time.

The empirical results which have been presented in this paper have been backed

**Table 4.2**

Periodwise Estimates of Number of Exploratory Drilling (NE) in the basin.

| Year | Estimated Number of Exploratory Drillings (NE) |
|---|---|
| 1985/86-1989/90 | 93 |
| 1990/91-1994/95 | 155 |
| 1995/96 - 1999/00 | 221 |
| 2000/01-2004/05 | 289 |
| Total | 758 |

by a model consisting of a set of relationships, mostly defined in terms of cumulated values of costs and other variables. To build up this model we have postulated various plausible causal relationships linking the relevant variables to explain the observed changes of these variables in the past. This is a fairly simple model. However, it seems to provide a reasonable good fit to the observed past data. The results of all the empirical exercises clearly bring out the phenomenon of rising marginal cost of production in the basin.

It should be mentioned here that the model developed here is suitable for a partially explored basin. The necessary data for the basin were collected from different offices of the ONGC because of the non-availability of the relevant information centrally. The period covered in this study extends upto 1984 and updating is now necessary.

Also, there were gaps and missing figures and it must be recognised that the quality of the results obtained from this model depends heavily on the quality of the input data.

The econometric model, presented in the paper, does not try to capture explicitly the random mechanism inherent in the process of exploration. It simply projects the past behaviour, observed at an aggregate level, into the future. Clearly, therefore, so long as the past pattern of movement remains unaltered, this approach is likely to

**Table 4.3**
Period-wise Discovery and Production Cost Implication of the Proposed Strategy of the Future Exploratory Activity in hte Basin*

| | Estimate for/at the end of terminal year | | | |
|---|---|---|---|---|
| Item | 1985/86-89/90 | 1990/91-94/95 | 1995]96-99/00 | 2000/01-04/05 |
| NE | 93 | 155 | 221 | 289 |
| NDL | 110 | 120 | 85 | 44 |
| CE | 2692 | 3952 | 5742 | 8080 |
| CDL | 4213 | 4857 | 5316 | 5553 |
| CS | 565 | 830 | 1206 | 1697 |
| R | 767 | 865 | 935 | 971 |
| CDV | 2484 | 2930 | 3248 | 3412 |
| Q | 101 | 121 | 135 | 143 |
| CMC | 26820 | 31615 | 35046 | 36815 |
| EDI | 200 | 133 | 67 | 25 |
| MCD | 18 | 28 | 54 | 147 |
| MCP | 332 | 371 | 483 | 866 |

*. All cost are in Rs. million, **R** and **Q** in MMt, marginal costs in Rs/tonne and EDI in tonnes/meter.

provide good results. As compared to the stochastic model, the results of the econometric model seems to be more optimistic. In the stochastic model it was found that not much reserve is left for further discovery. In other words, the amount of effort in terms of costs needed for a unit increase of discovery is much more in the stochastic model than in the econometric model. The values of the other costs are also inflated accordingly in the stochastic model. (See Ghosh *et al* (1997) for details).

The process of cumulation adopted in the econometric analysis smooth out, to a large extent, the fluctuations and thereby somewhat obscures the reflection of the stochastic behaviour of the process. To verify whether the estimated reserve equation suffered from such a deficiency, we did the validation exercise. Taking the cumulated data for the first $n$ years ($n = 10, 11, 12, \ldots, 23$) a modified exponential form of the reserve equation as described in the text was fitted. The estimates of the year-wise (non-cumulated) discovery for the first 5, 7, and 10 post sample years were obtained by the difference method. Finally the correlation coefficient between the observed and the estimated values $(r(R, \hat{R}))$ of year-wise discovery for the post sample years were calculated. We have got reasonably high values of $(r(R, \hat{R}))$ for most of the cases, indicating an interdependence of discovery and exploratory effort.

## APPENDIX

### Estimation of ACD, MCD, and EDI

Given the model described above, once the relationships of the model are estimated on the basis of actual data, the estimated model may be put to three major uses, viz; (i) finding our the average and the marginal cost of discovery (ACD and MCD), and (ii) calculation of the exploratory discovery index (EDI).

The average cost of discovery may be defined as

$$ACD = \frac{CE + CDL + CS}{R} \tag{A1}$$

which gives the total cost per unit of discovery at a time point $t$ (say). Thus, given $CE = CE_0$, we can estimate ACD by

$$\hat{ACD}(CE_0) = \frac{CE_0 + \hat{CDL} + \hat{CS}}{\hat{R}} \tag{A2}$$

$$= \frac{CE_0 + \left\{ C + d\left(k - b^{CE_0}\right) \right\} + \alpha\, CE_0}{K - b^{CE_0}} \tag{A3}$$

On the other hand, by fixing $R = R_0$, we can get

$$\hat{ACD}(R_0)\; \frac{\left\{\log(K-R_0) - \log a\right\}/\log b + (C+dR_0) + \alpha\left\{\log(K-R_0) - \log a\right\}/\log b}{R_0}.$$

A more appropriate way of finding $ACD(R_0)$ is to estimate the reverse equation

$$CE = a' + b' \log(K' - R) \tag{A4}$$

and then use this to get $ACD(R_0)$ as

$$\widehat{ACD}(R_0) = \frac{\{a' + b' \log(K - R_0)\} + (c + dR_0) + \alpha\{a' + b' \log(K - R_0)\}}{R_0} \tag{A5}$$

we have, however, applied (A3) to estimate ACD ($CE_0$) since all productions in our case were done through CE only.

Similarly, if ΔCE, ΔCS and ΔCDL are the marginal increase in cost corresponding to the (marginal) increase ΔR of reserve at time *t* the marginal cost of discovery is

$$\text{MCD} = \frac{\Delta CE + \Delta CDL + \Delta CS}{\Delta R} \tag{A6}$$

For prediction purposes, the observed values are to be substituted by the estimated values.

To predict marginal cost, first $\Delta \hat{R}$, $\Delta C\hat{D}L$ and $\Delta \hat{C}S$ given $CE_0$ are estimated from the following approximations

$$\Delta\hat{R} = \left(\frac{d\hat{R}}{d\hat{CE}}\right)_{CE=CE_0} \Delta CE_0 = \hat{R}'(CE_0)\,\Delta CE_0 \text{ (say)}, \tag{A7}$$

And in a similar way

$$\Delta C\hat{D}L = C\hat{D}L'(CE_0)\,\Delta CE_0 \tag{A8}$$

and

$$\Delta\hat{C}S = \hat{C}S'(CE_0)\,\Delta CE_0 \tag{A9}$$

Thus,

$$M\hat{C}D(CE_0) = \frac{1 + C\hat{D}L'(CE_0) + \hat{C}S'(CE_0)}{\hat{R}'(CE_0)} \tag{A10}$$

$$= \frac{1 + d\,(-ab^{CE_0} \log b) + \alpha}{(-ab^{CE_0} \log b) + \alpha} \qquad \text{(A11)}$$

Similar expressions can be found if we fix *R* instead of *CE*. Also, Marginal Cost of Production (MCP) can be callculated in a similar manner.

Recall that EDI is defined as the deiscovery per metre of drilling. As the data on metreage drilled are not available, we calculate EDI using a different approach desvribed below.

As a general framework, we define EDI as

$$EDI = MD \times \overline{C}\,\overline{M} \qquad \text{(A12)}$$

where, $MD$ is the reciprocal of the marginal cost giving the discovery per unit of marginal cost, and $\overline{C}\,\overline{M}$ is the average drilling cost per meter of drilling. These are calculated as

$$MD = \Delta R/(\Delta CE + \Delta CDL + \Delta CS) \qquad \text{(A13)}$$

and

$$\overline{C}\,\overline{M} = (CE + CDL + CS)/(\alpha_1 CE + \alpha_2 CDL) \qquad \text{(A14)}$$

where the cost of exploratory and delineation drillings are constants per meter of drilling and these costs are $\alpha_1$ and $\alpha_2$ meters per rupee respectively.

## REFERENCES

ARPS, J.J., MORTADA, M. AND SMITH, A.E. (1971) : **Relationship between Proved Reserves and Exploratory Effort,** Journal of Petroleum Technology, Vol. 23, pp. 671-675.

GHOSH, J.K., P. MAITI, A.C. MUKHOPADHYAY AND M. PAL (1997) : Stochastic Modeling and Forecasting of Discovery, Reserve and Production of Hydrocarbon - With an Application, Revised version submitted to Sankhya B.

HUBBERT, M.K. (1962) : **Energy Resources : A Report to the Committee on Natural Resources, National Academy of Sciences**, National Research Council Publication 1000-D, USA.

HUBBERT, M.K. (1974) : **U.S. Energy Resources, A Review as of 1972 : A Background Paper,** Sr. No. 93-40 (92-75), US Govt. Printing Office, Washington D.C., p.66.

JAIDA. ALI M. (1983) : **An Appraisal of OPEC Oil Policies,** Longman, London, NY.

MOORE, C.M. (1970) : **Future Petroleum Provinces of the US - A Summary**, Washington D.C; National Petroleum Council.

MUKHOPADHYAY. A.C. **et al** (1990) : **Estimation of Discovery and Production Costs of Hydrocarbon with some Applications to Indian Data,** Report of the Project undertaken in collaboration with ONGC, Dehradun; Indian Statistical Institute, Calcutta.

WIORKOWSKI, J.J. (1981) : **Estimating Volume of Remaining Fossil Fuel Resources** : A Critical Review, JASA, Vol. 76, pp. 534-547.