D-3887

MODELING THE ESTIMATION OF PETROLEUM RESOURCES IN THE UNITED STATES

by

John D. Sterman, George P. Richardson, Pål I. Davidsen

WP-1900-87

D-3887

Modeling the Estimation of Petroleum Resources in the United States

John D. Sterman[1], George P. Richardson[2], Pål Davidsen[3]

May 1987

[1] Associate Professor, Sloan School of Management, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

[2] Associate Professor of Public Administration and Policy, The Rockefeller College, State University of New York, Albany NY 12222

[3] Associate Professor, Institute for Information Science, University of Bergen, 5014 Bergen, Norway

We gratefully acknowledge the helpful criticisms of M. K. Hubbert, Gordon Kaufman, and Robert Fildes, and the technical assistance of Becky Waring.

Modeling the Estimation of Petroleum Resources in the United States

ABSTRACT

Estimates of ultimate recoverable petroleum resources in the lower 48 states have increased since 1910, but peaked in the 1960s and have since declined by over fifty percent. The apparent tendency of the estimates to overshoot their target raises questions about the rationality and utility of estimation strategies. This paper describes a simulation-based study of the petroleum lifecycle in the United States undertaken to evaluate different resource estimation techniques. Protocols for the Hubbert life cycle and USGS geologic analogy methods are developed and applied to synthetic data generated by the simulation model. It is shown that the Hubbert method is quite accurate, with a tendency to underestimate the ultimate recoverable resource somewhat, while the simulated geologic analogy estimates overshoot the resource base quite dramatically. Analysis of the model pinpoints the sources of error and suggests ways to improve resource estimation strategies.

Accurate estimates of the petroleum resource base in the United States are of fundamental importance in the formulation of energy policy, regional and national economic policy, and even foreign policy. Yet estimating the resource base is difficult and methods for doing so controversial. There are fundamentally divergent views among forecasters regarding the nature of petroleum resources, the treatment of technological change, and the appropriate sources of information for estimating the resource base. The uncertainty, combined with the importance of oil, have spawned a minor industry which has witnessed a proliferation of forecasts, models, and estimation procedures (for surveys and reviews see Mathtech 1978, MIT 1982, Grenon 1975, and Meyer 1977). The effort devoted to resource estimation, however, has not reduced the uncertainty or settled the debate. Estimates of ultimate recoverable petroleum resources vary substantially, both across estimation methods and over time (Meyer 1977, Odell and Rosing 1980). Worse, the traditional approach to evaluating forecasting methods, repeated comparison of forecasts to actual outcomes, is of little use because the true resource base will not be known for decades.

The research reported here contributes to the development of methods for evaluating forecasting techniques before actual outcomes are known. The approach is based on the use of synthetic data generated by a model of the processes being forecasted (Richardson 1982, Sterman and Richardson 1985).

A wide variety of estimation techniques currently exist, including life cycle (Hubbert 1956, 1962, 1982; Ryan 1966; Wiorkowski 1981), geologic analogy (Zapp 1962, Hendricks 1965, USGS 1975, Jones 1975, Energy, Mines, and Resources Bureau (Canada) 1977, Semenovich et al. 1977), rate of effort (Hubbert 1974) econometric (Khazzoom 1971, MacAvoy and Pindyck 1973), and discovery process methods (Arps and Roberts 1958, Ryan 1973, Barouch and Kaufman 1977). The techniques range from the basin and play level to continental and global aggregation, from detailed structural and process models to curve-fitting.

Despite the differences, all estimation procedures can be thought of as information processing schemes which take certain data as input and produce an estimate of the resource base as the output. Previous appraisals of estimation methods (e.g. Mathtech 1978, MIT 1982) have focused on the logical structure, parameter estimation, and data requirements of the methods. But to compare the various methods it is necessary to apply them to a consistent set of data. Here, these data are generated by a detailed simulation model of the petroleum lifecycle in the United States.

We have chosen to investigate the Hubbert life-cycle method and the geologic analogy method. First, the estimation methods are formalized. The resulting protocols specify in a precise and reproducible manner the way in which information is processed to yield an estimate. Second, the protocols are applied to synthetic data generated by the model, and a dynamic path of estimates is generated. The evolution of resource estimates over time is then compared to the resource base assumed in the model, and the accuracy of the estimation protocols is evaluated.

The History of Petroleum Estimates

Estimates of petroleum resources in the United States date from the early 20th century. Exhibit 1 shows estimates of the ultimate recoverable petroleum resource (URR) in the United States, including adjacent offshore areas but exclusive of Alaska. The estimates rose substantially through the 1960s, reaching as high as 590 billion barrels. The growth in estimates reflects increasing geological sophistication, improvements (actual and anticipated) in recovery technology, and, of course, increases in discoveries and reserves. However, since 1960 the estimates have fallen by over fifty percent and currently range between 200 and 250 billion barrels.

Exhibit 1 also shows a dramatic conflict between two different methods of estimating the URR. The lowest estimates are those of M.K. Hubbert (1956, 1962, 1974, 1975, 1982). In 1956, Hubbert forecast that the URR for the lower 48 states and adjacent offshore areas would be between 150 and 200 billion barrels, and projected "the peak in production should probably occur within the interval 1966-1971" (Hubbert 1975, 371). At the same time, the USGS, using the geologic analogy method (Zapp 1962) projected ultimate recoverable resources of 590 billion barrels, and concluded that

> ...the size of the resource base would not limit domestic production capacity 'in the next 10 to 20 years at least, and probably [not] for a much longer time' (Gillette 1974 129).

Production actually peaked in 1970, and, as Renshaw and Renshaw (1980, 58) have pointed out, Hubbert's "projected values for cumulative discoveries and production have not yet been exceeded." Assuming 1970 was the true peak of production, Hubbert's 1956 forecast leads the peak by some fourteen years, an impressive forecast indeed.

In contrast, the huge decline after 1960 in estimates made by the geologic analogy method raises the question of whether the overshoot was mere accident or due to inherent, systematic flaws in the analogy method. The consequences of such overestimation are potentially serious. Overestimation may lead to inefficient allocation of exploration effort, overvalued lease tracts, and complacency in the development of oil substitutes. It is important, therefore, to identify possible sources of overshoot in the estimation methods currently in use. The simulation approach employed in this study provides a way of identifying such sources of information-processing bias.

Modeling The Estimation Process

The model described below is but one of many that could conceivably be used to generate the synthetic data for the investigation. Not all models of the oil supply process are appropriate, however. In addition to the obvious constraint that the model must generate data at an appropriate level of aggregation for the estimation protocols, the model should have the following characteristics.

First, it should be a structural model. It should attempt to represent the physical and causal structure of the processes modeled, as opposed to a model based on historical correlations. Nonlinearities and constraints may alter historical correlations in the future. Physical delays, such as the time required to develop an oil field or build a synfuel plant, should be represented explicitly.

Second, it should be a behavioral model, portraying the information available to actors and the procedures they use to process it and arrive at decisions. The petroleum system is characterized by imperfect information, uncertainty, and distributed decisionmaking. If the model is to respond to changes in the environment in the same way the real actors do, this bounded rationality should be incorporated (Simon 1979, Hogarth 1980, Morecroft 1983).

Third, the model should generate its behavior endogenously. The discovery and production process is tightly interconnected with energy price, demand, substitution, and technology. A change in one part of the system may have ramifications throughout. A model that relies heavily on exogenous variables is likely to produce inconsistent results as the feedback effects are ignored. A model that generates the petroleum life cycle endogenously constitutes an internally consistent theory that is subject to analysis, refutation, and revision (Bell and Senge 1980).

In addition to these general considerations, a model of petroleum resources to be used in forecast evaluation should include the following specific features as endogenous components:

1. Technology: The ultimate recoverable resource depends significantly on the recovery factor. Only 30 to 40 percent of oil-in-place can be recovered economically with current technology, but the fraction recoverable has been rising and may rise substantially in the future.
2. Economic incentives: Economic incentives (primarily determined by the price of oil) play a large role in determining proved reserves, exploration, and production. Oil that is subeconomic at $10 per barrel may be highly profitable at $30 per barrel. Regions that were not even considered for exploration may be prime candidates for test wells at a higher price.
3. Price: Because the price has a strong influence on the incentives for exploration and development, it must be modeled explicitly. The effects of production costs, supply and demand, and market imperfections should be incorporated.
4. Demand and Substitution: Petroleum demand is sensitive to price. As prices rise, the demand for oil will be depressed, and the production of substitutes ("backstops" [Nordhaus 1973]) such as synfuels will be stimulated. The pattern of demand and substitution will have a strong influence on production and investment in exploration. Delays in the response of demand and in the development of the backstop industry should be explicit.
5. Depletion: The total initial quantity of oil-in-place is finite. As it is consumed, the quantity remaining inevitably declines, and the marginal cost increases, ceteris paribus. Though improving technology may offset depletion and cause the real price of oil to decline, the finite resource base and its depletion must be treated explicitly.

The Data-Generating Model

The criteria proposed above impose strong constraints. Exhibit 2 provides an overview of the model developed in this study to meet them. The model described here is an outgrowth of the one described in Sterman and Richardson (1985) and used there to model the global petroleum lifecycle. The model employs the system dynamics approach to simulation (Forrester 1961, Richardson and Pugh 1981). Other applications of system dynamics to energy include Naill (1973, 1977), Backus et al. (1979), Choucri (1981), and Sterman (1983). The model is described in detail in Davidsen, Sterman and Richardson (1987), and a documented equation listing is available from the authors. As shown in exhibit 2, the model is divided into five basic sectors: (1) exploration; (2) production; (3) technology; (4) revenue and investment; and (5) demand and substitution.

1. Exploration: The model divides the total quantity of oil-in-place into three basic categories: as yet undiscovered oil, identified resources, and cumulative production. Within these broad categories, several finer divisions are portrayed. The disaggregation of the resource base follows the standard resource classification shown in the McKelvey box format (USGS 1976) in exhibit 3. The McKelvey box is a useful but static characterization of the resource base. Over time, exploration and production activity shift the boundaries in the box. Successful exploration shifts the boundary between identified and undiscovered resources to the right; improvements in technology or increases in the real price of oil shift the boundary between economic and subeconomic resources towards the bottom. Production shrinks the reserve base.

As an example of model structure, the determinants of the exploration rate are shown in exhibit 4. The rate at which undiscovered resources are identified is determined by investment in exploration and the productivity or yield of that investment. Note that additions to the identified resource include all oil-in-place identified through exploration and not just the economic, proven part that is immediately producible, which is often mis-labeled "discoveries" (MIT 1982, part II). Additions to the identified resource depend on investment expenditure and the desired discovery rate. To represent the time required to identify and explore a prospective oil-bearing region, the potential discovery rate is given by lagged investment expenditure. The rate of investment, in turn, depends on the desired discovery rate, modified by profitability. If the expected revenues from exploration activity do not justify the cost, or if the expected cost of developing new reserves exceeds the cost of oil substitutes, exploration is curtailed. Conversely, higher than normal return induces entry and expansion of exploration efforts. The desired discovery rate is the rate at which resources need to be identified to meet anticipated production and expected growth in production, and to provide the reserve levels required to meet anticipated production.

The cost of exploration activity is determined directly by the yield or productivity, which depends on technological and geological factors. At the dawn of the oil era, only a small fraction of oil-in-place was discoverable. As the ability to drill deeper wells was developed, a larger fraction of oil-in-place in a given region could be identified. As the ability to drill offshore and in increasingly hostile environments was developed, a larger fraction of potential oil-bearing areas could be economically explored. And as the sophistication of seismic detection technology grew, smaller and smaller oil deposits, for example, in stratigraphic traps, could be identified.

At the same time, however, depletion reduces the productivity of exploration efforts. Producers naturally explore those areas they believe most likely to yield oil first, drilling shallow wells and tapping giant oilfields when possible before moving on to less accessible and more expensive regions. To the extent producers are able to identify oil at a better than random rate, the productivity of future exploration activity is necessarily reduced (ceteris paribus), as future additions to the identified resource will involve more dry holes, deeper wells, and increasingly, drilling offshore or in distant and hostile locations. The evidence suggests exploration activity in the United States historically has been 2.75 times more effective than chance drilling (McCray 1975, 229). Hubbert and others (Hall and Clevelend 1981) have documented a significant decline in yield per foot drilled both as a function of time and as a function of cumulative exploratory effort for the United States:

> In fact, 'finding rates' had fallen sharply since the late 1930's as oilmen skimmed the cream off the prospects in Texas, Oklahoma, and California. From a high of 276 barrels per foot of exploratory drilling, discoveries have fallen to about 35 barrels per foot by 1965 and to 30 in 1972 (Gillette 1974, 129).

Though depletion causes yield to decline, close examination of the U.S. data show actual yields increased in the 1920s and again in recent years, illustrating the shifting dominance of technical, economic, and geological factors. These factors are represented in the model and consequently, as shown below, the simulated yield to exploration first rises with technology and then falls with depletion.

2. Production: Production in the model is determined by three major factors: the quantity of identified resource remaining, recovery technology, and investment in production facilities. Investment in production facilities depends on anticipated demand for natural pretroleum, modified by profitability. As in the exploration decision, higher than normal returns cause expansion of production. An insufficient return causes a cutback in production as existing wells are shut down and plans for new wells cancelled. Investment in production capacity is also constrained by the technically recoverable quantity of oil. Limitations on the rate of flow and on the density of producing wells constrain useful investment in producing wells, though it is assumed in the model that production/reserve ratios can be increased somewhat above normal levels in a situation of high demand or profit.

3. Technology: Technology in the model is endogenously generated. As shown in exhibit 5, the model distinguishes between the fraction of oil-in-place that is discoverable with current technology and the fraction of the identified resource recoverable with current technology. The fraction discoverable represents the feasible depth of wells, the ability to drill offshore and in hostile environments, and the effectiveness of geologic survey and identification technology. The fraction recoverable represents the effectiveness of secondary and tertiary recovery techniques.

Each type of technology improves as the result of research effort. Improvements in technology take time, and an average delay of six years is assumed between an increase in expenditures on research and development and the resulting improvement in technology. Expenditures on R&D are assumed to be a fixed fraction of industry revenues. The effectiveness of investment in technology is variable. As the level of technology improves, the marginal improvement in technology per dollar of research effort declines. The total R&D effort is allocated between discovery and recovery technology on the basis of the perceived marginal benefit to each. Initially, the majority of research is devoted to improved exploration technology designed to increase the fraction discoverable. As the fraction discoverable rises towards 100 percent, research effort gradually shifts to improving recovery from developed fields.

4. Revenues and Price: Revenues are given by the price and production of natural petroleum. The price of natural petroleum is determined by production and exploration costs and by supply and demand. When supply and demand are in balance, the price equilibrates at a level sufficient to cover exploration and production costs and to provide the required return on investment. Imbalances in the market cause price to adjust; such disequilibria may persist until supply and demand respond to the change in prices.

Investment expenditures are allocated among exploration, production, and R&D on the basis of the relative need for funds and expected profitability of each acitivity.

5. Demand and Substitution: The demand for petroleum is endogenously portrayed in the model (exhibit 6). The total demand for oil is determined by real GNP and the oil intensity of the economy. The oil intensity of the economy is determined by the average price of oil. The average price is given by the prices and market shares of domestic, imported, and synthetic petroleum. An average lag of fifteen years is assumed between a change in the price of oil and its full effect on demand. The fifteen-year lag is somewhat shorter than the twenty year average life of energy consuming capital (Coen 1975), to represent the potential for retrofitting existing capital.

The market share of natural petroleum is determined by its price relative to the price of imports and synthetic substitutes. Investment in production capacity for synthetics responds to changes in the relative prices of natural and synthetic oil. The acquisition delay for synthetic production capacity is assumed to be eight years. The real price of synthetic substitutes is taken to be constant.

Model Calibration

The model has been calibrated to represent the United States. The major quantitative assumptions are listed in exhibit 7. The simulation results reproduce the US experience accurately (exhibit 8). Exhibit 9 presents the mean-squared error between simulated and actual oil production, demand, imports, and price. Theil's inequality statistics (Theil 1966, Sterman 1984) decompose the mean-squared error between simulated and historic values into the fraction due to bias, unequal variance, and unequal covaration. Small mean-squared error and low bias indicate a good correspondence between simulated and historic data. The statistics show virtually no bias in the tracking of physical variables: the bulk of the mean-squared error is concentrated in the unequal covariation term, showing that the model captures the trends and turning points in the data and only differs point-by-point. The root-mean-squared error between simulated and historical prices is $4.43 per barrel. Forty percent of the MSE arises from bias: the simulated price is too high, particularly before 1930. However, the model captures the high volatility and general decline in real oil prices through 1973 (note the variances of simulated and actual price are virtually equal). The fit between model and history is particularly strong considering the fact that the model endogenously generates the complete lifecycle of the resource beginning in 1870, and, as will be shown, also replicates patterns of resource estimates.

However, the key aspect of the simulation is not the specific values of parameters or variables. The evolution of the petroleum system to date is but one draw from a large number of possibilities: the initial endowment of petroleum could have been different, discoveries could have occurred earlier or later, recovery technology could have developed at a different pace, and so on. A good estimation procedure should be able to produce accurate estimates for any consistent resource development scenario, and must not depend on the realization of a particular scenario. Thus, the results presented here are not contingent on the precision with which the model reproduces the past history of oil discovery and use. Our focus is the relationship between estimates of the resource base and the assumed resource base.

The total quantity of oil-in-place is assumed to be 550 billion barrels. It is assumed technology can improve so that all oil-in-place is potentially discoverable and that the recovery factor can rise to as high as 50 percent. The *maximum* ultimate recoverable resource is therefore 275 billion barrels, consistent with contemporary estimates. Note that the *actual* values of the discovery and recovery factors are endogenous and may not attain their maxima. Likewise, the ultimate quantity produced may be less than the potential due to the substitution of backstop technologies before exhaustion of the resource.

Results

Simulation results are shown in exhibit 10. The simulation starts in 1870 and runs until 2050. With the exception of the exogenous growth rate of real GNP and the price of imported oil, the behavior is endogenously generated over the nearly two-century life cycle of the resource.

In the late 19th and early 20th centuries, simulated demand and production grow rapidly (exhibit 10b). Growth of the industry stimulates R&D, and the fraction discoverable rises rapidly, particularly after 1920 (exhibit 10d). Between 1900 and 1925, improving technology causes the yield to exploration effort to rise from about 280 barrels per foot drilled to over 330 barrels per foot (exhibit 10d), comparable to the historic rise in yield clearly documented in the data (Hubbert 1969). As a result of rising yields, the rate at which resources are identified greatly exceeds production (exhibit 10a), causing recoverable resources (exhibit 10c) to rise faster than production.

Transient variations in the demand for oil induced by the uneven growth of real GNP causes the simulated real price to fluctuate, as did the historic price. Nevertheless, clear trends in price emerge in both the simulated and actual data. The improvement in technology and yield causes the real price of oil to decline by over fifty percent between 1900 and 1950 (exhibit 10e). The reduction in the real price of oil causes demand to grow faster than the economy, and the average oil intensity rises, corresponding to the transition from a coal- to oil-powered economy.

After 1930, simulated yield begins to drop. Though discovery technology is still improving rapidly, the very effectiveness of exploration in locating oil implies future efforts will be less successful. As the giant oilfields and shallow deposits are found, additional exploration yields more dry holes and smaller finds. By the late 1940s the rate of addition to identified resources reaches its maximum. But though new finds are declining, they remain well above production, and reserves continue to grow.

Initially, R&D activity was focused on discovery technology, and the fraction recoverable grows only slightly (exhibit 10d). But as discovery technology becomes more effective, R&D effort is shifted towards enhancing recovery factors. After 1940, the fraction recoverable begins to rise rapidly, corresponding to the development of secondary and tertiary recovery techniques.

Despite the substantial improvements in the technology of discovery and recovery, by the late 1960s the depletion of the resource begins to be felt. Exponential growth of production, stimulated by low prices and historic import controls, outstrips additions to recoverable reserves by the early 1960s. The ratio of technically recoverable reserves to production starts to drop, and simulated production peaks in 1968 (compared to 1970 in reality). Throughout the 1970s, despite high prices after 1973, production (actual and simulated) continues to decline.

By the 1970s, the industry has reached a turning point. Declining yield has caused real prices to rise, and higher prices begin to suppress demand and awaken interest in oil substitutes. Though natural petroleum still dominates the market, over three quarters of the total oil-in-place has been identified, and additions to identified resources are falling.

Stimulated by the high prices following the second oil shock in 1979 there is a burst of exploration effort in the early 80s, and additions to identified resources rise. The decline of production slows, and imports fall from their peak in the mid 70s. The precipitous decline of world oil prices in 1986 however causes simulated exploration activity to collapse. In fact, the number of rotary drill rigs in operation in 1987 is a post-war low. Production immediately begins to drop, and imports soar.

By the year 2000 the dominance of depletion over technology and economic incentives is complete. Improving technology boosts the fraction discoverable to over 85 percent and the fraction recoverable to over 44 percent by the year 2000. The real price, following the assumed path of import prices, reaches $30 per barrel by 2000. Nevertheless, finding rates continue to drop. Production declines from over 2.8 billion barrels in 1987 to less than 1.7 billion barrels in 2000. Significant investment in substitutes is undertaken, but due to long development lags natural petroleum loses market share only slowly. In consequence imports surge, capturing two-thirds of the market by 2000. Such large imports may in fact be infeasible due to political, economic, and technical factors surrounding Middle Eastern production. In all likelihood, the pressures created by such large imports would cause the price of imported petroleum to rise above the level assumed in the model. Due to the tremendous uncertainty surrounding such unpredictable price movements we have not attempted to capture the world price endogenously. Nevertheless, it is worth noting that depletion of conventional resources and lags in substitute development means vulnerability to future price hikes increases as imports grow (Hirsch 1987).

After 2000 the transition to substitutes accelerates. The market share of synthetics rises to 10 percent by 2000 and exceeds 35 percent by 2020. Because the world price is assumed to be exogenous, imports continue to satisfy a major portion of the demand. It is unlikely that imports will be available at the assumed price for such an extended period. Several studies point to the likelihood of a period of price overshoot in which the delays in development of substitutes cause prices to overshoot the long run equilibrium cost of backstops (Sterman 1983, Energy Modeling Forum 1981, DOE 1979).

By 2020 the petroleum era in the United States is largely over. Production is about 740 million barrels per year and falling. There is virtually no exploration activity, and virtually no oil left to find if there were. Substitutes and imports are less costly than marginal exploration effort or further investment in recovery technology. Reviewing the entire life cycle highlights the following points:

1. The life cycle of production follows a roughly bell-shaped path, though it is definitely asymmetrical, with periods of slowdown and acceleration. Note also that production falls off more gradually than it grows (exhibit 10a).
2. Consistent with the United States experience, the yield to exploration first rises, as a consequence of improving technology, and then falls as a consequence of depletion (exhibit 10d).
3. Likewise, improvements in technology first cause the real price to decline, but eventually depletion dominates technology and the real price rises (exhibit 10e).
4. Delays in the development of substitutes for natural petroleum cause an extended period of import dependency. Should the availability of imports be curtailed, the price of energy would

overshoot its long run level until sufficient substitution and conservation could be implemented (exhibit 10b).

5. Though the ultimate recoverable resource could have reached as high as 275 billion barrels, the actual resource recovered by 2050 is approximately 228 billion barrels or about 83% of the maximum. Substitution to the backstop causes production and investment in technology to stop before the ultimate limits are reached (exhibit 10c).

The Estimation Protocols

We have evaluated two estimation procedures, the Hubbert life cycle approach and the geologic analogy approach used by the USGS and others. Each of these techniques can be applied to the aggregate data generated by the model.

The Hubbert Method: Hubbert has actually developed two methods to estimate ultimate recoverable resources, the original life cycle approach and a later rate-of-effort approach. We consider here the life cycle approach. It was the first method he developed, the most controversial, and also the most accurate to date in projecting production and reserves in the United States.

Hubbert's method has been extensively described, criticized, and analyzed elsewhere (MIT 1982, Mathtech 1978). To apply the method to the model-generated data, we developed the following protocol:

1. Define cumulative proved discoveries as cumulative production plus technically recoverable reserves.
2. Assume cumulative proved discoveries follow a logistic path given by:

$$Q_t = \frac{Q^{\infty}}{1 + a^*\exp[b(t-t_0)]} \tag{H1}$$

where

Q^{∞} = ultimate recoverable resource

Q_t = cumulative proved discoveries at time t

a, b = parameters to be estimated

t_0 = an arbitrary initial time

3. Rearrange equation (H1) as

$$\ln[(Q^{\infty}/Q_t)-1] = \ln(a) + b(t-t_o) \qquad \text{(H2)}$$

4. Estimate the parameters of equation (H2) by ordinary least squares regression for various values of Q^{∞}, and select Q^{∞} from the regression that yields the highest R^2.

In Hubbert's original work, Q^{∞} was estimated by "a trial and error graphical method" in which he plotted the data on semi-log paper and, judging by eye, chose the Q^{∞} that best fit the data (MIT 1982, III-2-10). We have used regression so that our results are reproducible. Hubbert's graphical method is equivalent to the regression technique if one is willing to assume that the "best" fit judging by eye is roughly equivalent to the least squares estimates of the parameters in equation (H2). No measurement error is introduced, as we are primarily concerned with the tendency of estimation methods to overshoot even when perfect information is assumed. The robustness of the protocols in the face of process noise and measurement error is left as a topic for future research.

Values of Q^{∞} were estimated by the protocol above using the model-generated data from 1900 to 1935, 1940, 1945, and so on. The results are shown in exhibit 11, compared against the "true" ultimate recoverable resource. The Hubbert method eventually provides an unbiased estimate of URR, settling within 28 percent of the true value by 1955, within 10 percent by 1990, and within 5 percent by 2005. Before the year 1940, the best fit to the logistic curve actually yields an infinite value for Q^{∞}.

After 1940 the estimated Q^{∞} falls rapidly, dropping below the "true" URR of 228 billion barrels by 1945. By 1955, the best estimate of Q^{∞} produced by regression of the simulated data is 165 billion barrels, accurately reproducing Hubbert's 1956 estimate. By 1970 the best estimate has risen to 170 billion barrels. As the life cycle unfolds, the estimate rises toward the true value, coming within 10% of the final true URR by 1990.

In 1956 Hubbert forecast that the ultimate recoverable resource for the lower 48 states and adjacent offshore areas would be between 150 and 200 billion barrels. Stemming from his graphical technique for estimating Q^{∞}, Hubbert's range of uncertainty was intuitive, not based on a formal confidence interval. In our simulated world, that range in the estimated Q^{∞} happens to correspond to values which produce the same R^2 to three-decimal accuracy. Because it matches Hubbert's range of estimates, we have used the three-decimal criterion to produce the ranges for Q^{∞} shown in exhibit 12. It is interesting to compare this three-decimal criterion to a formal confidence interval computed from the regression on the simulated data. For data through 1955, with the optimal $Q^{\infty} = 165$ billion barrels, a 99% confidence interval for URR is 152 to 181 billion barrels. The lower bound matches Hubbert's intuitive choice of 150 billion barrels. But Hubbert accounted for uncertainty stemming not only from noise in the data, but also from the possibility of improvements in recovery technology. Thus the upper range of his intuitive confidence bound for the 1956 estimate (200 billion barrels) is substantially higher than than the 99% bound derived from the simulated data.

The life cycle approach relies on the fact that the finite nature of the resource necessarily implies a roughly S-shaped path for cumulative production and discoveries. The logistic model satisfies this requirement, but imposes the constraint that the fractional rate of growth declines continuously and linearly throughout the life cycle. In order to estimate the logistic successfully, therefore, the data must continuously reflect the decline in the growth rate caused by depletion. As demonstrated by the simulation, the growth rate may not decline monotonically, much less linearly, even when depletion of the resource is in fact strictly monotonic (exhibit 13). Improvements in discovery technology cause the reserve/production ratio to rise in the early years of the lifecycle, while Hubbert's model presumes a constant reserve/production ratio. Moreover, technological progress causes the real price of petroleum to fall, increasing the growth rate of demand above that of the economy. Thus increasing demand and improving technology dominate the depletion effect in the first third of the life cycle. Consequently depletion, though occurring continuously, is masked in the aggregate data, causing Hubbert's method to predict an essentially infinite Q^{∞} before 1940.

The life cycle approach, therefore, is only likely to give accurate estimates after the depletion effect dominates over other forces that may conspire to cause the fractional rate of production or discovery to rise. In the simulation, that shift in dominance occurs between 1950 and 1970. Note, however, that the simulated Hubbert estimate undershoots the true URR and then gradually rises after 1950, converging by 1990 to within 10% of the true value (exhibits 11 and 12). There seem to be two reasons for the underestimation of URR. First, fluctuations in the rate of economic growth may masquerade as the depletion effect. In particular, the growth of cumulative discoveries fell significantly in the wake of the great depression (exhibit 10a), depressing the estimate of Q^{∞} in the late 1940s. As oil production accelerated in the boom years of the 50s and 60s, the optimal Q^{∞} rose. Second, Hubbert's model does not allow for future expansion of the fraction recoverable. As technology improves, however, the best estimate of Q^{∞} gradually rises. Hubbert himself notes (1969, 183):

> The figure of 165×10^9 bbls is accordingly the best present estimate of the value of Q^{∞} for the conterminous United States, although it is admitted that a somewhat higher figure resulting from further improvement in recovery efficiency is a physical possibility.

The Geologic Analogy Method: Geologic analogy or volumetric methods are the prime alternative to Hubbert's approach to estimating ultimate recoverable resources. In essence, the method consists of

> ...projecting average yield factors (barrels of oil per cubic mile of sedimentary rock or per square mile of surface area) uniformly over a sedimentary rock stratum (Mathtech 1978, III-297).

The USGS estimates of 1975 present one of the most comprehensive and detailed uses of these techniques to date. The essence of the method was described by the Survey as follows:

> Estimates of recoverable oil and gas resources are based upon a series of resource appraisal techniques....The techniques used include: (1) an extrapolation of known producibility into untested sediments of similar geology for a well-developed area; (2) volumetric techniques using geologic analogs and setting upper and

> lower yield limits through comparisons with a number of known areas; (3) volumetric estimates with an arbitrary general yield factor applied when direct analogs were unknown; (4) Hendricks' (1965) potential area categories; and (5) comprehensive comparisons of all known published estimates for each area to all estimates generated by the above methods (USGS 1975, cited in MIT 1982, III-5-13).

Despite the apparent rigor, the USGS study actually involved a high degree of subjective judgment and discussion, and the protocols used to reach consensus have been criticized as "mismanaged" (MIT 1982, III-5-19). Our representation of the process abstracts from the subjective and political nature of the process to focus on the sources of information for the economic, technical, and geologic assumptions made in the study.

The survey divided the resource base into the standard classifications of the McKelvey box, and assumed

> ...that undiscovered recoverable resources will be found in the future under conditions represented by a continuation of price/cost relationships and technological trends generally prevailing in the recent years prior to 1974. Price/cost relationships since 1974 were not taken into account because of the yet undetermined effect these may have on resource estimates....
>
> These assumed conditions permit the appraisal of recoverable oil and gas resources to be made on the basis of: (1) relevant past history and experience concerning recovery factors; (2) the geology favorable to the occurrence of producible hydrocarbons; and (3) the size and type of reservoirs which have geen found, developed, and produced....
>
> The economic recovery factor used was based on a current national average of approximately 32 percent....Sub-economic identified resources of crude oil were calculated on the following assumptions: (1) that on the average, 32 percent of original oil-in-place is recoverable if there are no substantial changes in present economic relationships and known production technology, and (2) that ultimately the recovery factor could be as large as 60 percent....
>
> It is extremely optimistic to assume that 60 percent of the oil-in-place will eventually be recovered. If [this] becomes a reality, it is likely to occur only through gradual development over an extended time period. The remaining 40 percent of oil-in-place is not included as it is considered to be nonrecoverable.... (USGS 1975, cited in MIT 1982, III-5-9, 10).

The protocol used to test the geologic analogy method appears in the appendix. The protocol assumes far better information than is actually available to real estimators. Cumulative production, technically recoverable reserves, cumulative identified oil-in-place, the current recovery factor, and the area explored are assumed to be known exactly. The only potential sources of error are in the estimation of the future recovery fraction and in the expected yield of oil-in-place in unexplored areas. In both of these cases, the model, like the Survey, assumes "a continuation of price/cost relationships and technological trends generally prevailing in...recent years...."

Applying the geologic analogy protocol to the data generated by the model yields the path of estimates summarized in exhibit 14. The components of the estimates are shown in exhibit 15. The estimates start low, rise rapidly, overshoot the ultimate quantity recovered, and settle at the ultimate quantity recovered.

Exhibit 15 tells the following tale. In 1900 the estimates are very low--only a small fraction of the sedimentary basins of the country have been surveyed, both discovery and recovery technology are primitive, and little of the resource has been identified. With increasing exploration experience, improving exploration technology, and growing knowledge of sedimentary basins, the estimates steadily rise, reaching almost 200 billion barrels by 1940. Most of the estimated URR consists of probable recovery from unidentified resources--the quantity expected with current recovery technology from known sedimentary basins that are as yet unexplored, assuming historic yields to exploration.

Between 1940 and 1960 the simulated estimate jumps by 260 percent, reaching a peak in excess of 500 billion barrels in 1959. Historic estimates from the same period ranged from 400 to 600 billion barrels. Though all components of estimated ultimate recovery are growing, the bulk of the estimate in 1960 (58 percent) is still due to probable future discoveries. By 1960, the accelerating growth of recovery technology has caused the expected recovery fraction to exceed the current fraction. Expectations of technical improvement for both the identified and estimated unidentified resource are still quite cautious, however, accounting for less than 20 percent of the estimate.

The estimate peaks around 1960 in excess of 220 percent of the true value. The sources of the overshoot are first the overestimation of the yield to be expected per foot drilled in unexplored areas and second an overly optimistic assessment of future recovery technology. Actual yield per foot drilled falls beginning in the 1930s. Nevertheless, as a consequence of the lags in recognizing and adapting to lower yields, the expected yield is four times higher than the actual yield on the margin in 1960 (exhibit 16). Similarly, the rapid progress in recovery technology that begins after 1940 leads forecasters to project continued improvement. By 1960, the expected fraction recoverable has risen to 55 percent; by 1970 it has risen to about 60 percent. The actual fraction recoverable in 1960 and 1970 is about 30 and 35 percent, respectively; it reaches only 46 percent by 2020. This compares to the survey's mid-1970s estimate that the fraction recoverable might rise to as high as 60 percent when the actual recovery fraction was 32 percent. Thus extrapolation of technical progress leads to overoptimistic assessments of production from identified oil-in-place. When the expected future recovery fraction is applied to the estimate of future oil-in-place, the overshoot is compounded.

After 1960 evidence rapidly mounts that prior estimates were overly optimistic. The estimate falls rapidly, reaching 324 billion barrels by 1980 and 290 by 2000. The estimated quantity of oil-in-place remaining to be identified is falling, partly because the area unexplored is shrinking, but primarily because of significantly declining yield to exploration. The decline in yield that began in the 1930s is finally recognized, and the future yield is discounted below historic levels (exhibit 16).

One may wonder how reasonable it is to assume such long time lags in the recognition of the yield to exploration. Expected yield per well is based in the model on historic yields, discounted according to past trends in the yield. The delay is assumed to be 15 years, reflecting the time required to compile yield data, to separate a systematic change in the yield from the noise, and for the revised yield estimates to become accepted throughout the geologic community. For example, the largest decline in yield per foot drilled in the United States occurred between 1940 and 1950. Hubbert pointed out the declining trend in yield per foot in the United States in 1962. Until 1965, the USGS continued to use the so-called "Zapp hypothesis" of constant future yield. Even then the Survey assumed a value that exceeded more recent yields (Gillette 1974, 129):

> That year, the USGS noted a 'definite decline' in discoveries and postulated now that oil would, on the average, prove to be only half--not equally--as abundant in unexplored rock as in explored rock. Now this number is in contention, with Hubbert claiming that it's at least five times too large for onshore terrain. [USGS director] McKelvey acknowledges that the figure of one-half was largely a 'subjective judgment' and another official describes it as 'mostly a guess.'

Another perspective on the overshoot appears in exhibit 17 which compares the simulated estimates of the recoverable resource remaining to the simulated true recoverable resource remaining. The history of resource estimation divides into two distinct phases. At first, estimates are too low and rise steeply as more knowledge is gained. The estimated recoverable resource remaining overtakes the true quantity remaining in the 1940s. The estimates continue to rise for a decade or so, then reverse and fall towards the true quantity remaining. The estimates lag substantially behind the true quantity remaining due to expectations of continued technical progress and near-historical yields.

Note, however, that though there is no change in the estimation process between 1950 and 1970, there is a dramatic shift in perspective. Within twenty years, the historic trend of growing estimates reverses. The result of this shift was conflicting estimates and method-ological disagreements (Warman 1972, Odell 1973, Seidl 1977), disagreements still echoing today (Odell and Rosing 1980, Adelman 1986, Wilkinson 1986, Scanlan 1986, Gall 1986, Hirsch 1987).

Conclusions and Implications

Most previous appraisals of resource estimation techniques have focused on the sources of information, the statistical procedures, and the analytic framework used by the various estimators. This work suggests that a complementary approach based on simulation of the various methods offers important insights into the dynamics of resource estimation. By formalizing estimation protocols and applying them to synthetic data, it is possible to assess the accuracy of an estimation technique before the true resource base is known. Further, it appears to be possible, as in the Hubbert case, to identify time frames in which the method is accurate. The results of this study explain the striking divergence between historic estimates from the Hubbert lifecycle method and the geologic analogy method (exhibit 18).

Hubbert's method has been criticized as merely an exercise in fitting data to an arbitrary curve. Yet these results show the life cycle approach can yield an accurate estimate of the ultimate recoverable resource, provided the resource is far enough into its life cycle so that the depletion effect begins to dominate other factors and depress the growth rate of cumulative discoveries. In the case of the United States, this point was reached approximately twenty years before the peak in production. Before then, the life cycle method overestimates the ultimate recoverable resource. The results explain the impressive accuracy of Hubbert's projections for the United States. It is worthwhile noting that Hubbert presumed a logistic curve and has been criticized for not using a more flexible functional form that allows the data to dictate the presence of asymmetries (MIT 1982, III-2). The model used to generate the synthetic data does not presume a logistic curve, nor does it generate one, but Hubbert's approach produces accurate estimates nonetheless. These caveats aside, the simulation clearly demonstrates that the astounding accuracy of Hubbert's method was no fluke. There are fundamental reasons, deeply rooted in the physics of exploration and production, for the accuracy and long lead time of Hubbert's resource and production predictions.

Examination of the geologic analogy approach, in contrast, shows that the historic overshoot and collapse of resource estimates can be explained in terms of the information sources available to resource estimators and the estimation procedures used. Though ostensibly superior to the Hubbert method because it involves the use of disaggregate, primary geologic data, the analogy method actually involves a high degree of judgment, extrapolation of past trends, and educated guessing. The simulation results suggest the substantial overshoot of the estimates was a consequence of systematic biases intrinsic to the method, biases that persist even when a high degree of perfect information is assumed.

In a sense, the superior performance of Hubbert's method can be attributed to its 'mechanical' or formal character. Precisely because it relies on only a small number of data inputs and processes these inputs in a straightforward manner, it does not create the opportunity for subjective judgmental biases to creep in to the forecast. This result is consistent with much research in behavioral decision theory (Hogarth and Makridakis 1981, Armstrong 1985). The implications for resource forecasting are clear: the expense and effort required to create sophisticated forecasts may be, at best, ineffective when compared to simple and apparently 'naive' methods.

APPENDIX: PROTOCOL FOR THE GEOLOGIC ANALOGY METHOD

In the equations, the prefix 'G' denotes a quantity estimated by the geologic analogy protocol; other variables denote the true values generated by the model. A '<P>' denotes the assumption of perfect information.

GEURRt = GCUMPRt + GTRRRt + GEATRRt + GEFDt (G1)
<P> GCUMPRt = CUMPRt (G2)
<P> GTRRRt = TRRRt (G3)

where

GEURR = Estimated ultimate recoverable resource (bbls)
GCUMPR = Estimated cumulative production (bbls)
GTRRR = Estimated technically recoverable resource remaining (bbls)
GEATRR = Estimated additions to technically recoverable resource (bbls)
GEFD = Estimated future discoveries (bbls)
CUMPR = Cumulative production (bbls)
TRRR = Technically recoverable resource remaining (bbls)

The expected ultimate recoverable resource is divided into four basic categories: cumulative production, technically recoverable reserves, expected additions to technically recoverable reserves, and expected future discoveries. The estimated values of cumulative production and technically recoverable reserves are assumed to be equal to the true values.

GEATRRt = GETRRRt - GTRRRt (G4)
GETRRRt = GCUMAIRt*GEFRt - GCUMPRt (G5)
<P> GCUMAIRt = CUMAIRt (G6)
GEFRt = GFRt + GEIFRt (G7)
<P> GFRt = FRt (G8)
GEIFRt = (1-GFRt)*f_1(GEGFRt) $f_1(0)=0$, $f_1'>0$ (G9)
GEGFRt = TREND(GFRt) (G10)

where

GEATRR = Estimated additions to technically recoverable resource (bbls)
GETRRR = Estimated expected technically recoverable resources remaining (bbls)
GTRRR = Estimated technically recoverable resource remaining (bbls)
GCUMAIR = Estimated cumulative additions to identified resource (bbls)
CUMAIR = Cumulative additions to identifed resource (bbls)
GEFR = Expected fraction recoverable (dimensionless)

GCUMPR = Estimated cumulative production (bbls)
GFR = Estimated current fraction recoverable (dimensionless)
FR = Fraction recoverable (dimensionless)
GEIFR = Expected increase in fraction recoverable (dimensionless)
GEGFR = Expected growth in fraction recoverable (1/years)
TREND = Function to estimate growth rate of a variable (Sterman 1987)

Technically recoverable reserves include all the known resource that can be recovered with current technology, whether it is currently economic to do so or not. Expected additions to technically recoverable reserves represents the additional recovery from currently identified resources due to anticipated advances in recovery technology. The expected addition is given by the difference between what could be recovered at anticipated levels of technology and what is currently recoverable. We assume perfect knowledge of the quantity of identified resource and of the cumulative original oil-in-place identified. Similarly, the current fraction recoverable is assumed known.

The expected increase in the fraction recoverable is based on the expected rate of technical progress. The expected rate of technical improvement is based on the trend in the recovery fraction over the past ten years. We assume changes in the trend in recovery factors are incorporated in the forecast after an average lag of ten years. The lag stems from the time required to become aware of new recovery techniques, to evaluate and build confidence in their effectiveness, and for that information to diffuse through the geological community and become enough a part of "conventional wisdom" to be included in government projections.

The maximum possible addition to the fraction recoverable is, of course, the fraction unrecoverable. The fraction of this maximum improvement that is expected is nonlinearly related to the expected rate of technical improvement. When the recovery fraction is not growing, no improvement in technology is expected and the anticipated increase in the recovery fraction is zero. When the growth rate is higher than 1.5 percent per year, the expected increment in the fraction recoverable reaches a maximum, assumed to be 40 percent of the fraction unrecoverable. (The USGS assumed a maximum potential recovery factor of 60 percent compared to an average of 32 percent in 1975. Thus the anticipated improvement was expected to be 28 percentage points out of a maximum of 68, or .41 of the maximum.)

$$\mathrm{GEFDt} = \mathrm{GPFDt} + \mathrm{GSFDt} \qquad (G11)$$

$$\mathrm{GPFDt} = \mathrm{GFRt} * \mathrm{GEURt} \qquad (G12)$$

$$\mathrm{GSFDt} = \mathrm{GEIFRt} * \mathrm{GEURt} \qquad (G13)$$

where

GEFD = Expected future discoveries (bbls)
GPFD = Probable future discoveries (bbls)
GSFD = Speculative future discoveries (bbls)
GFR = Estimated current fraction recoverable (dimensionless)
GEUR = Estimated undiscovered resource (bbls)
GEIFR = Expected increase in fraction recoverable (dimensionless)

Expected future recovery from unexplored areas is the least certain component of any resource estimate. We have disaggregated the total into two components: (1) the quantity of currently unidentified oil expected to be recovered at current recovery factors (GPFD) and (2) the additional quantity expected to be recovered at anticipated recovery levels (GSFD). Both of these quantities depend directly on the estimate of unidentified oil-in-place (GEUR).

$$\mathrm{GEURt} = \mathrm{GAUt} * \mathrm{GFDt} * \mathrm{GEYUAt} \qquad (G14)$$

$$\mathrm{GAUt} = \mathrm{GASt} - \mathrm{GAEt} \qquad (G15)$$

<P> $\mathrm{GAEt} = \mathrm{AEt}$ (G16)

$$\mathrm{GASt} = f_2(t) \quad f_2' \geq 0 \qquad (G17)$$

<P> $\mathrm{GFDt} = \mathrm{FDt}$ (G18)

where

GEUR = Estimated undiscovered resource (bbls)
GAU = Estimated area unexplored (sq. mi.)
GFD = Estimated fraction discoverable (dimensionless)
GEYUA = Expected yield from unexplored area (bbls/sq. mi.)
GAS = Surveyed area of sedimentary basins (sq. mi)
GAE = Estimated area explored (sq. mi.)
AE = Area explored (sq. mi.)
FD = Fraction discoverable (dimensionless)

Estimated unidentified oil-in-place is the product of the area unexplored, the fraction of that area in which exploration is feasible given current technology, and the expected yield in that area. The fraction of oil-in-place that is currently discoverable is assumed to be known exactly.

The area unexplored is given by the total global area in which sedimentary basins are known to exist less the area already explored. The area in which sedimentary basins are known to exist is specified exogenously. Assumed to be quite small in 1900, knowledge of sedimentary basins expands to 2.6 million square miles by 1970. The area actually explored is endogenously generated by the model and is related to the cumulative resource identified. If oil were distributed uniformly over the total area of sedimentary basins, and if exploration activity were no better than random, the relationship between area explored and identified oil-in-place would be linear. However, oil is distributed very unevenly, and exploration activity is better than random. Giant and supergiant fields account for one percent of known fields but 75 percent of known reserves and 65 to 70 percent of production (Klemme 1977). The assumed curve is therefore highly nonlinear.

GEYUAt = GWD*GAWD*GEYEt (G19)
GWD = .5 (G19.1)
GAWD = 6000 (G19.2)
GEYEt = GHYEt*GFHYEt (G20)
GHYEt = DLINF3(YEt,GTAEY) (G21)
GTAEY = 15 (G21.1)
GFHYEt = DLINF3(f3(GTYt),GTAYE) f3(0)=1, f3'>0 (G22)
GTAYE = 10 (G22.1)
GTYt = TREND(YEt) (G23)

where

GEYUA = Expected yield from unexplored area (bbls/sq. mi.)
GWD = Estimated well density (wells/sq. mi.)
GAWD = Estimated average well depth (ft/well)
GEYE = Expected yield to exploration (bbls/ft)
GHYE = Estimated historical yield to exploration (bbls/ft)
GFHYE = Fraction of historical yield expected (dimensionless)
DLINF3 = Third order exponential information smoothing
YE = Yield to exploration (bbls/ft)

GTAEY = Time to adjust estimates of historical yield (years)
GTY = Estimated trend in yield to exploration (1/years)
TREND = Function to estimate growth rate of a variable

The expected yield of oil-in-place per square mile of unexplored area is based on the density of wells, the average well depth required to explore a region fully, and the expected yield per foot drilled. We assume average well density and depth to be one well every two square miles and 6000 feet per well, respectively (Gillette 1974). Expected yield per well is based on historic yields, as discussed above, discounted according to past trends in the yield.

It is assumed in the model that the expected yield in unexplored areas is discounted below the historic yield when the yield is perceived to be falling. In the USGS study, the choice of the discount factor was highly subjective. The Survey acknowledged that

> ...the proper [discount factor] is open to conjecture. The fraction can range from one (or greater) to zero. Precedents exist for both 1.0 and 0.5. Qualitatively, 1.0 seems optimistic but not unreasonable; 0.5 seems conservative; less than 0.5 seems pessimistic (Mallory, 'Synopsis of Procedure' cited in MIT 1982, III-3 -12).

The assumed discount becomes progressively larger as the decline rate grows.

REFERENCES

Adelman, M. (1986) "The Competitive Floor to World Oil Prices," The Energy Journal, 7(4), 9-31.

Armstrong, J. S. (1985) Long Range Forecasting. New York: Wiley.

Arps, J. and Roberts T. (1958) "Economics of Drilling for Cretaceous Oil Production on the East Flank of the Denver-Julesburg Basin," AAPG Bulletin, 42(11), 2549-66.

Backus, G. et al. (1979) FOSSIL79: Documentation. (3 vols.) DSD 166. Hanover NH: Resource Policy Center, Dartmouth College, 1979.

Barouch, E. and Kaufman, G. M. (1977) "Estimation of Undiscovered Oil and Gas," Proceedings of the Symposia in Applied Mathematics, 21. Providence: American Mathematical Society, 77-91.

Bell, J. and P. Senge (1980) "Methods for Enhancing Refutability in System Dynamics Modeling," Legasto et al. (ed.) System Dynamics, TIMS Studies in the Management Sciences, 14, 61-73.

Choucri, N. (1981) International Energy Futures: Petroleum Prices, Power, and Payments. Cambridge: The MIT Press.

Coen, R. (1975) "Investment Behavior, the Measurement of Depreciation, and Tax Policy," American Economic Review, 65, 59-74.

Davidsen, P., J. Sterman, and G. Richardson (1987) "A Petroleum Lifecycle Model for the United States with Endogenous Technology, Exploration, Recovery, and Demand," Proceedings of the 1987 International System Dynamics Conference, Shanghai, China.

Department of Energy (1979) National Energy Plan II, DOE/TIC-10203 and Appendix A: World Oil Prices, DOE/TIC 10203 (APPA). Washington DC Department of Energy, April 1979.

Energy, Mines, and Resources Bureau (Canada) (1977) Oil and Gas Resources of Canada, 1976. EMS Report EP77-1.

Energy Modeling Forum (1981) World Oil: EMF6 Summary Report (Draft 3).Stanford University: Energy Modeling Forum, January 1981.

Forrester, J. W. (1961) Industrial Dynamics. Cambridge: The MIT Press.

Gall, N. (1986) "We Are Living off our Capital," Forbes, 22 Sept., 62-66.

Gillette, R. (1974) "Oil and Gas Resources: Did USGS Gush Too High?" Science, 185, 12 July, 127-30.

Grenon, M. (ed.) (1975) Methods and Models for Assessing Energy Resources. Oxford: Pergamon Press.

Hall, C. A. and Cleveland, C. J. (1981) "Petroleum Drilling and Production in the United States: Yield per Effort and Net Energy Analysis," Science, 211, 4482, 6 February, 576-579.

Hendricks, T. A. (1965) "Resources of Oil, Gas, and Natural Gas Liquids in the United States and in the World," USGS Circular 522.

Hirsch, R. (1987) "Impending United States Energy Crisis," Science. 235 (20 March), 1467-73.

Hogarth, R.M. (1980) Judgement and Choice. New York: Wiley.

Hogarth, R. and S. Makridakis (1981) The Value of Decisionmaking in a Complex Environment: An Experimental Approach," Management Science. 27, 93-107.

Hubbert, M. K. (1956) "Nuclear Energy and the Fossil Fuels," Drilling and Production Practice. Washington: American Petroleum Institute, 7-25.

Hubbert, M. K. (1962) Energy Resources. Washington: National Academy of Sciences--National Research Council 1000-D (Reprinted as National Technological Information Service PB-222401, 1973).

Hubbert, M.K. (1969) "Energy Resources," in Committee on Resources and Man (eds), Resources and Man. San Francisco: W.H. Freeman, 157-242.

Hubbert, M. K. (1974) "U.S. Energy Resources: A Review as of 1972," A National Fuel and Energy Policy Study, Part I. Washington: Committee on Interior and Insular Affairs, U.S. Senate, Serial No. 93-40.

Hubbert, M. K. (1975) "Hubbert Estimates from 1956 to 1974 of U.S. Oil and Gas," Grenon, M. (ed.) Methods and Models for Assessing Energy Resoures. Oxford: Pergamon Press, 370-83.

Hubbert, M. K. (1982) "Techniques of Prediction as Applied to the Production of Oil and Gas," Gass, S. (ed.) Oil and Gas Supply Modeling, Washington, D.C.: National Bureau of Standards Special Publication 631, 16-141.

Jones, R. W. (1975) "A Quantitative Geologic Approach to Prediction of Petroleum Resources," Haus, J. D. (ed.) Methods of Estimating the Volume of Undiscovered Oil and Gas Resources. Tulsa: American Association of Petroleum Geologists, 186-95.

Khazzoom, J. D. (1971) "The FPC Staff's Econometric Model of Natural Gas Supply in the United States," Bell Journal of Economics and Management Science, 2(1), 51-93.

Klemme, H. D. (1977) "World Oil and Gas Reserves from Analysis of Giant Fields and Basins," Meyer, R. F. (ed.) The Future Supply of Nature-Made Petroleum and Natural Gas. New York: Pergamon Press, 217-60.

MacAvoy, P. and Pindyck, R. S. (1973) "Alternative Regulatory Policies for Dealing with the Natural Gas Shortage," Bell Journal of Economics and Management Science, 4(2), 454-98.

Mathtech (1978) Comparative State-of-the-Art Assessment of Oil Supply Modeling. Palo Alto, Electric Power Research Institute.

McCray, A. W. (1975) Petroleum Evaluation and Economic Decisions. Englewood Cliffs NJ: Prentice-Hall.

Meyer, R. F. (ed.) (1977) The Future Supply of Nature Made Petroleum and Gas. New York: Pergamon Press.

MIT (1982) Estimation of Resources and Reserves. Cambridge MA: MIT Energy Lab, MIT-EL 82-010.

Morecroft, J. (1983) "System Dynamics: Portraying Bounded Rationality,"Omega, 11, 131-42.

Naill, R. F. (1973) "The Discovery Lifecycle of a Finite Resource: A Case Study of U.S. Natural Gas," Meadows, D. L. and Meadows, D. H. (eds.),Toward Global Equilibrium. Cambridge: The MIT Press, 1973, 213-56.

Naill, R. F. (1977) Managing the Energy Transition. Cambridge MA, Ballinger, 1977.

Nordhaus, W. D. (1973) "The Allocation of Energy Resources," Brookings Papers on Economic Activity, vol. 3, 529-570.

Odell, P. R. (1973) "The Future of Oil: A Rejoinder," Geographical Journal, 139, 436-54.

Odell, P. R. and Rosing, K. E. (1980) The Future of Oil. London: Kogan Page.

Renshaw, E. and Renshaw, P. F. (1980) "U.S. Oil Discovery and Production," Futures, February, 58-66.

Richardson, G. P. (1982) "Dynamics of Estimates of Petroleum Resources," Dynamica, 8(1), 54 -56.

Richardson, G. P. and Pugh, A. L. III (1981) Introduction to System Dynamics Modeling With DYNAMO. Cambridge: The MIT Press.

Ryan, J. M. (1966) "Limitations of Statistical Methods for Predicting Petroleum and Natural Gas Reserves and Availability," Journal of Petroleum Technology, 281-86.

Ryan, J. M. (1973) "An Analysis of Crude-Oil Discovery Rate in Alberta," Bulletin of Canadian Petroleum Geology, 21(2), 219-235.

Scanlan, T. (1986) "A Comment on 'The Competitive Floor to World Oil Prices'," The Energy Journal, 7(4), 34-35.

Seidl, R. F. (1977) "Implications of Changing Oil Prices on Resource Evaluations," Meyer, R. F. (ed.) The Future Supply of Nature Made Petroleum and Gas. New York: Pergamon Press, 113 -37.

Semenovich, V. V. et al. (1977) "Methods Used in the U.S.S.R. for Estimating Potential Petroleum Resources," Meyer, R. F. (ed.), The Future Supply of Nature Made Petroleum and Gas. New York: Pergamon Press, 139-53.

Simon, H. (1979) "Rational Decisionmaking in Business Organizations," American Economic Review, 69, 493-513.

Sterman, J. D. (1983) "Economic Vulnerability and the Energy Transition," Energy Systems and Policy, 7(4), 259-301.

Sterman, J. D. (1984) "Appropriate Summary Statistics for Evaluating the Historical Fit of System Dynamics Models," *Dynamica,* 10: 51-66.

Sterman, J. D. (1987) "Expectation Formation in Behavioral Simulation Models," Behavioral Science. Forthcoming.

Sterman, J.D. and G.P. Richardson (1985) "An Experiment to Evaluate Methods for Estimating Fossil Fuel Resources," Journal of Forecasting, 4 (1985) 197-226.

Theil, H. (1966), Applied Economic Forecasting, Amsterdam: North Holland.

USGS (1975) "Geological Estimates of Undiscovered and Recoverable Oil and Gas Resources in the United States," Circular 725, Washington, D.C.

USGS (1976) "Principles of the Mineral Resource Classification System of the U.S. Bureau of Mines and U.S. Geological Survey," USGS Bulletin 1450-A, Washington, D.C.

Warman, H. R. (1972) "The Future of Oil," Geographical Journal 138, 287-97.

Wilkinson, J. (1986) "The Competitive Floor to World Oil Prices: A Comment," The Energy Journal, 7(4), 32-33.

Wiorkowski, J. (1981) "Estimating Volumes of Remaining Fossil Fuel Resources: A Critical Review," Journal of the American Statistical Association 76(375), 534-47.

Zapp, A. D. (1962) "Future Petroleum Producing Capacity of the United States," USGS Bulletin 1142-H.

Exhibit 1: Estimates of ultimate recoverable petroleum in the coterminous United States and adjacent offshore area. Source: MIT 1982, II-1-3.

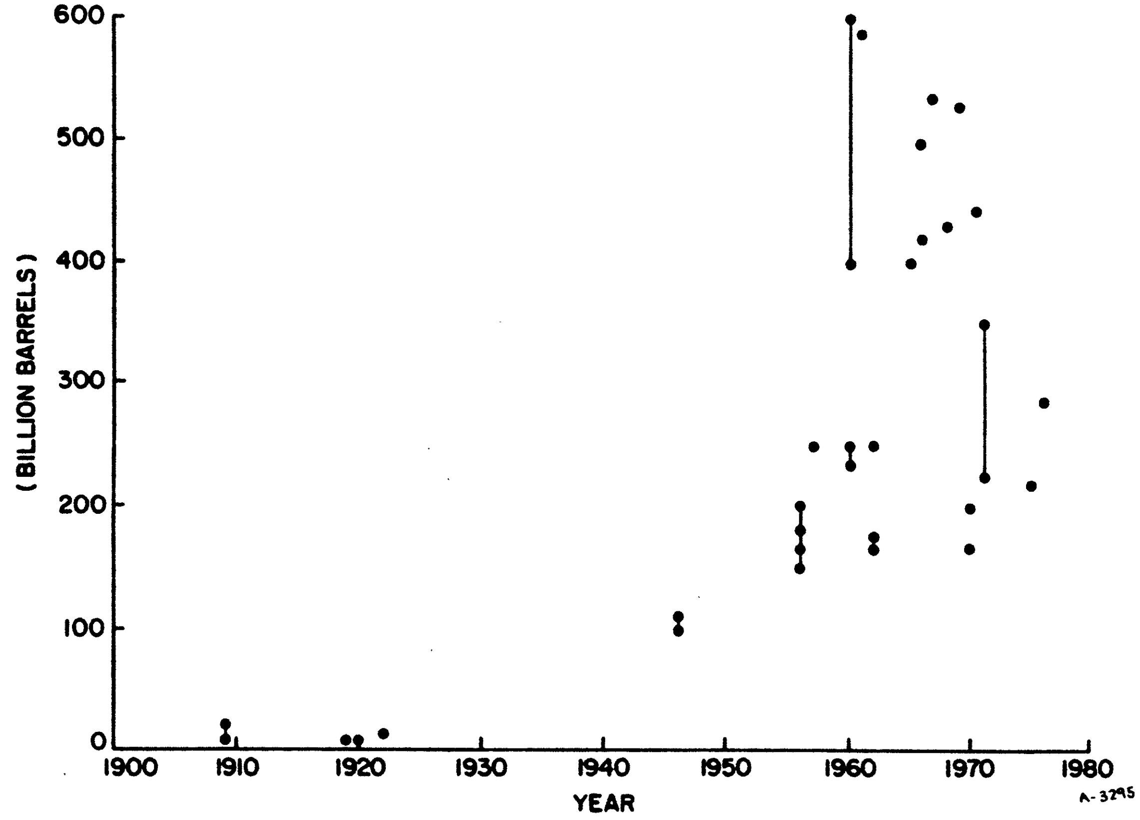

Exhibit 2: Model overview.

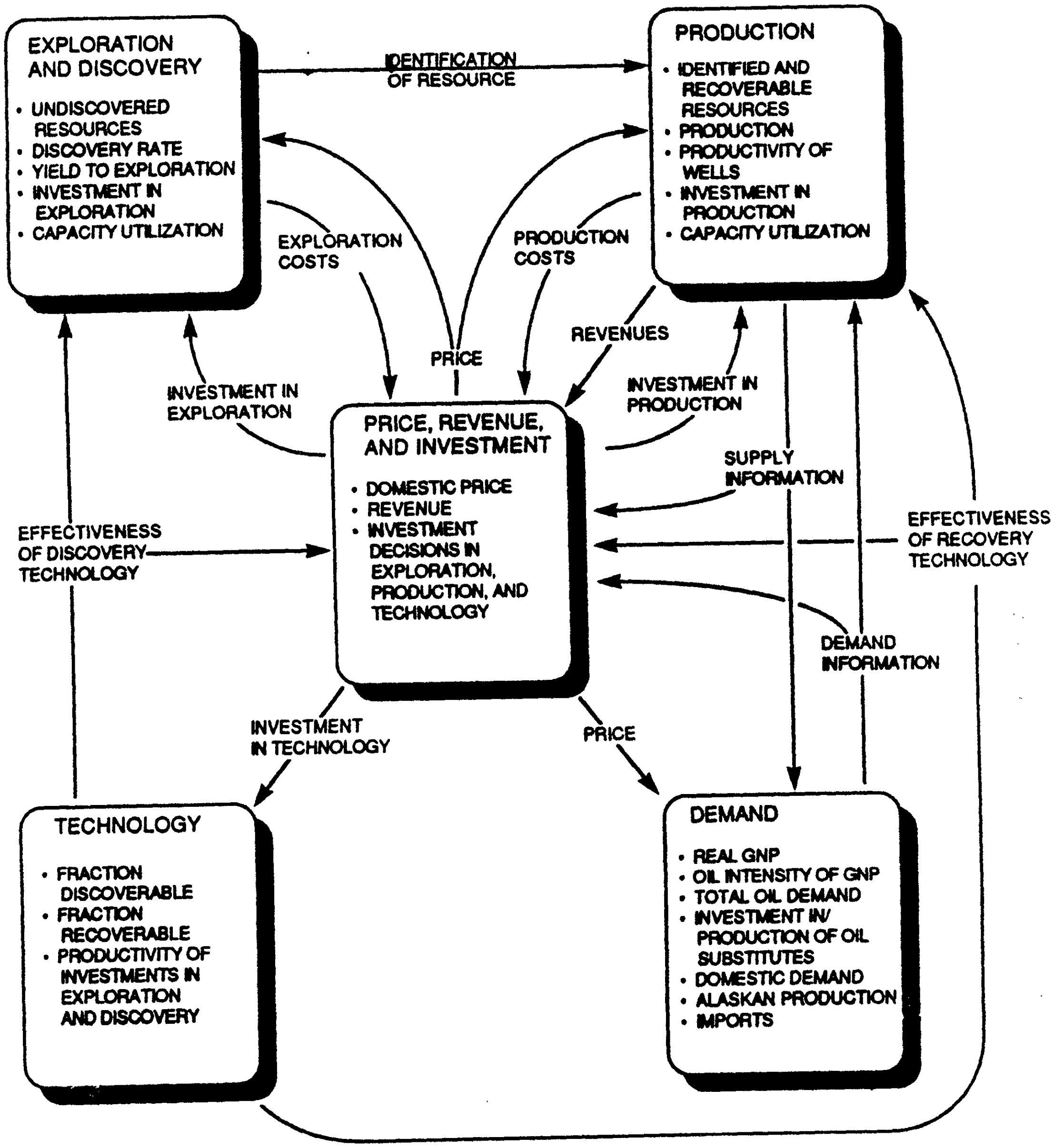

Exhibit 3: Classification of resources. Source: USGS 1976.

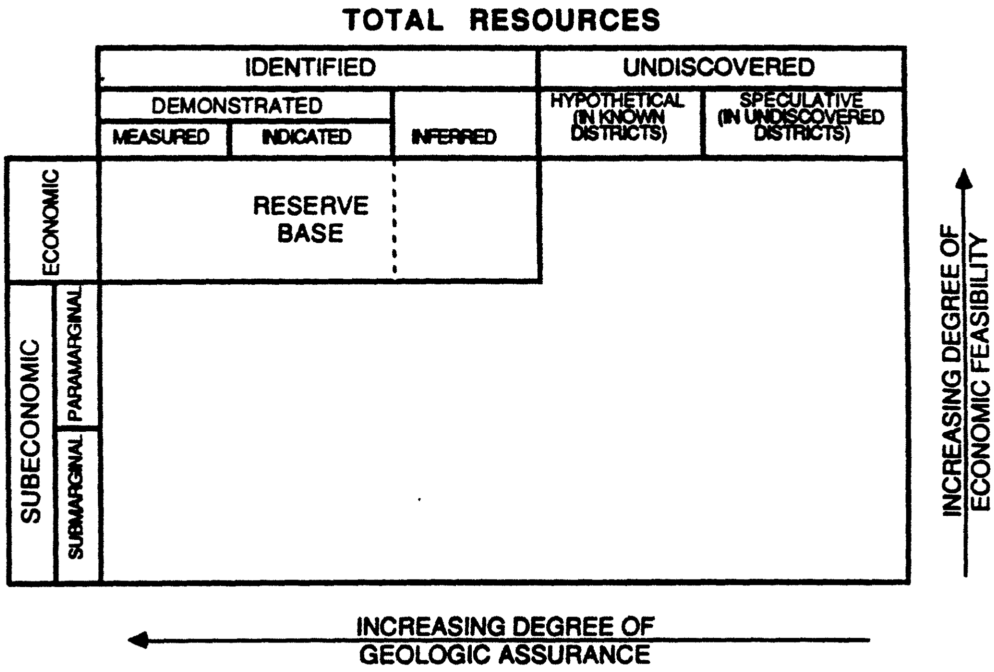

Exhibit 4: Determinants of exploration and discovery.

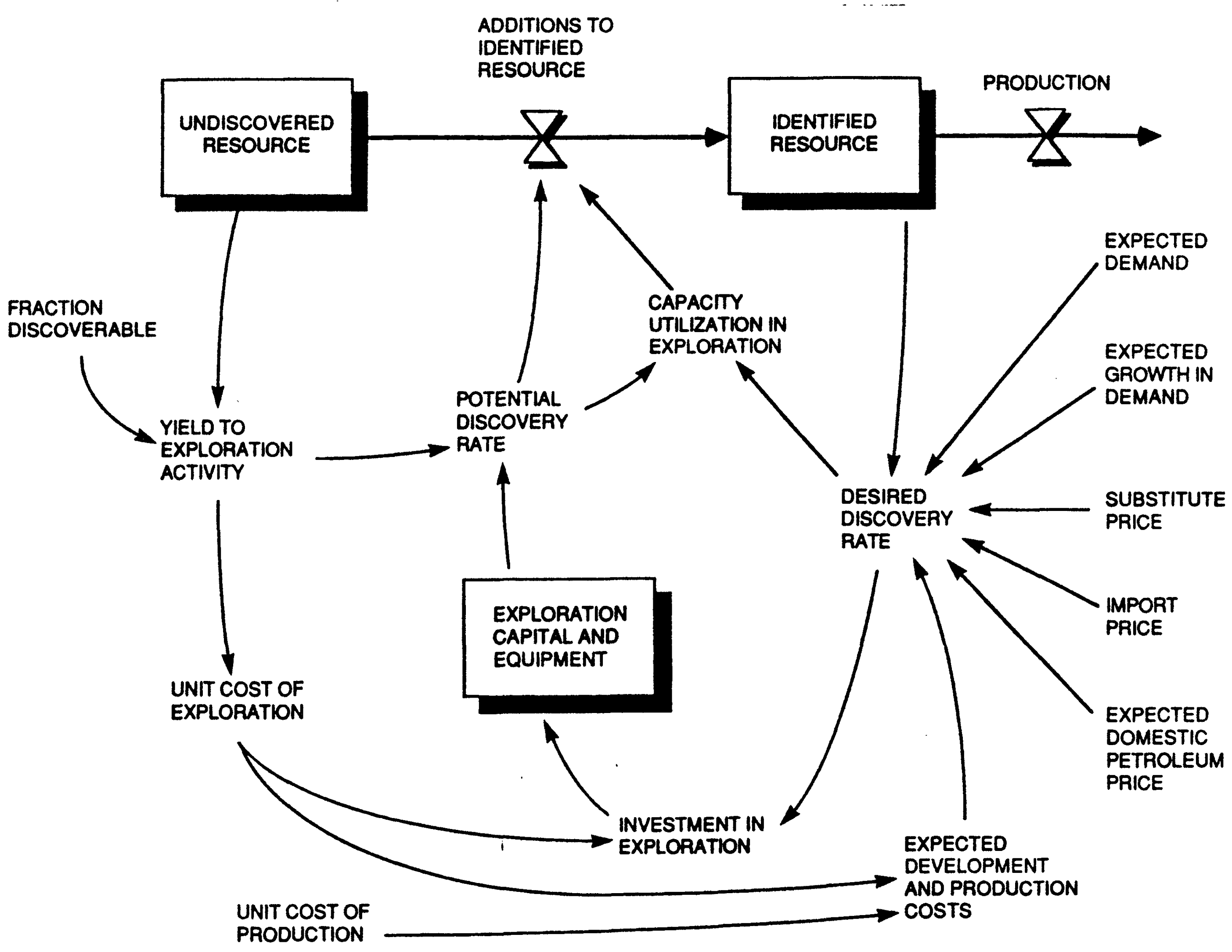

Exhibit 5: Technology sector.

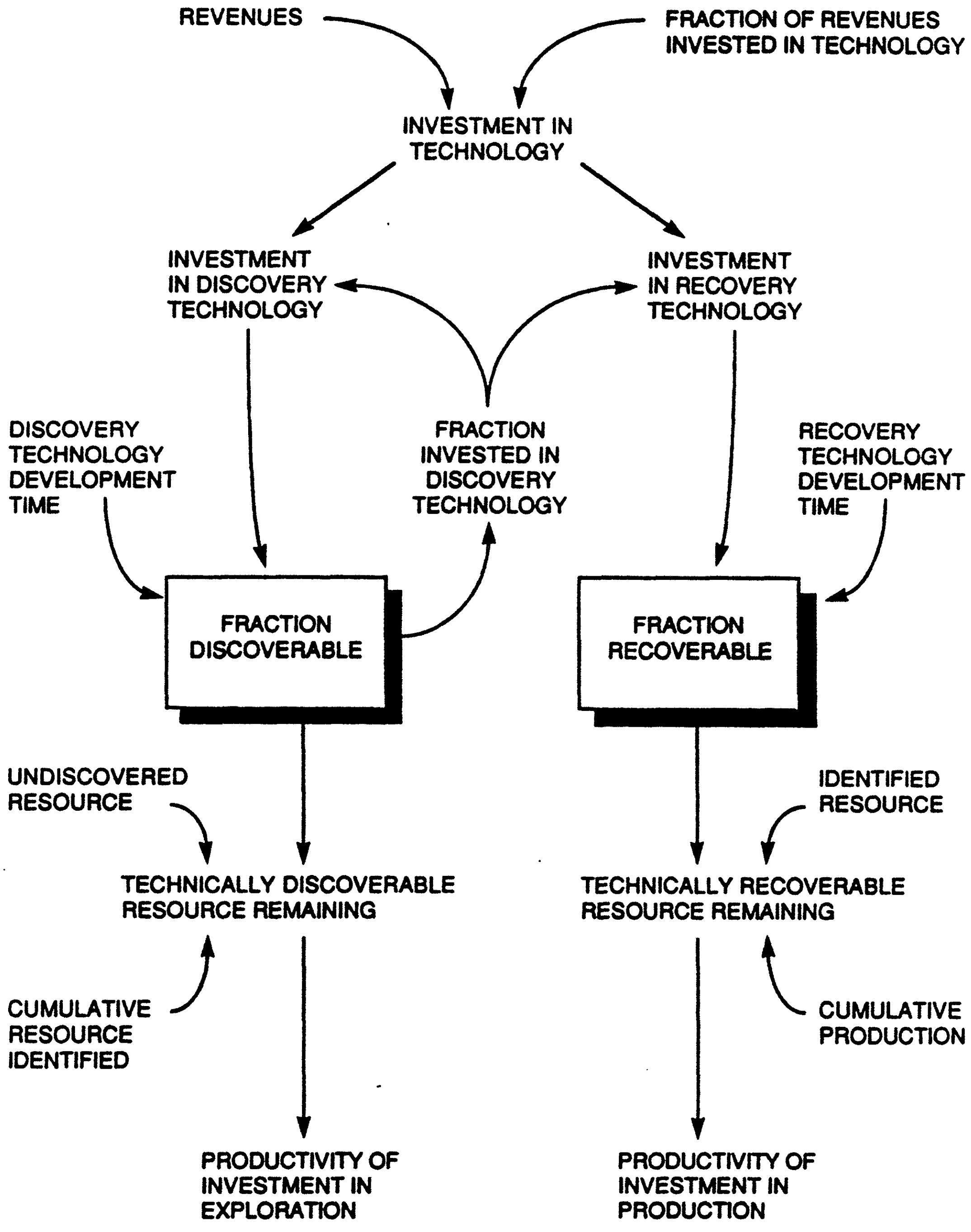

Exhibit 6: Demand sector

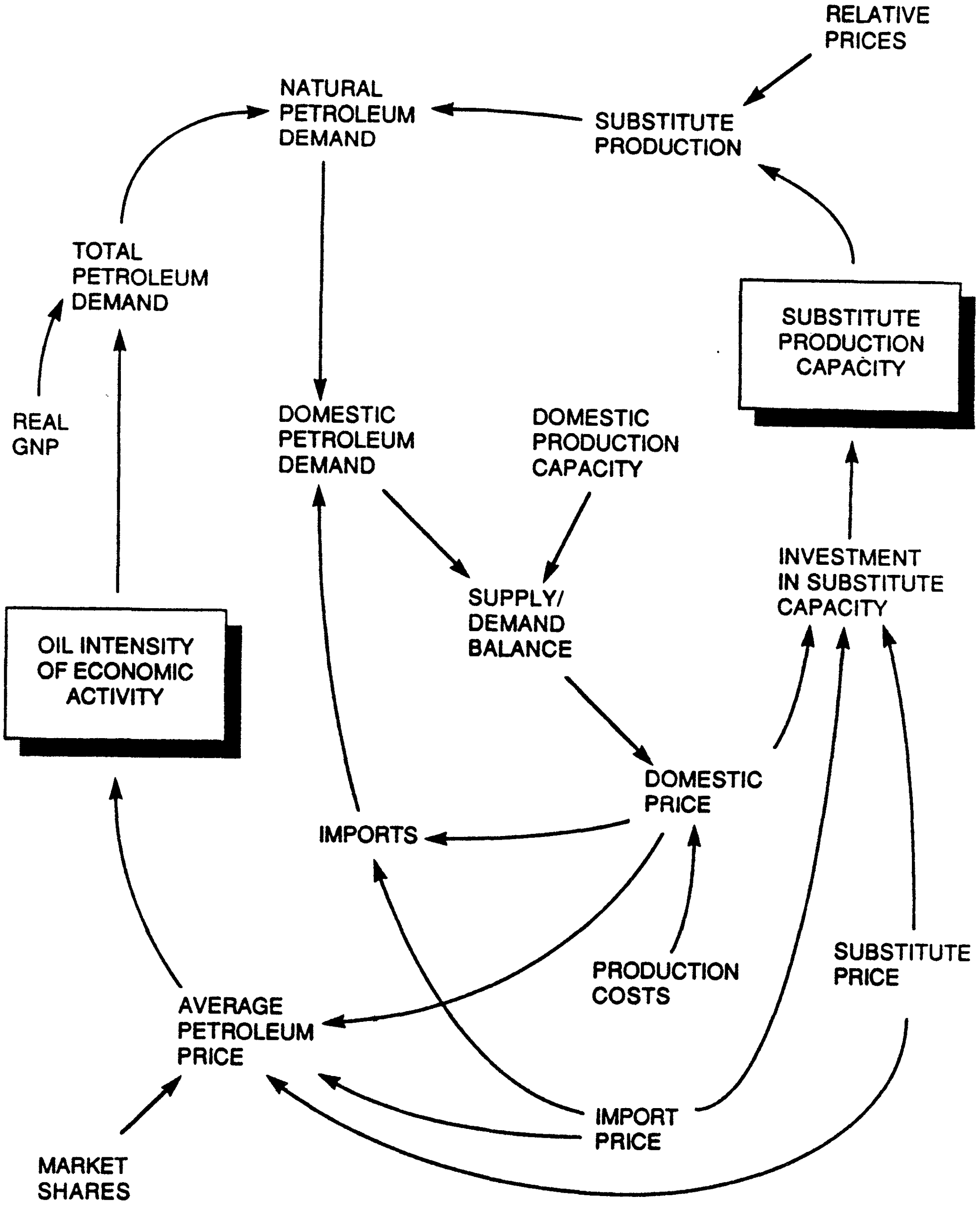

Exhibit 7. Major Parametric Assumptions

Parameters:

Quantity	Value
Total resource (billion bbls)	550
Exploration development delay (years)	4
Average technology development time (years)	6
Initial fraction discoverable (dimensionless)	0.2
Maximum fraction discoverable (dimensionless)	1.0
Initial fraction recoverable (dimensionless)	0.2
Maximum fraction recoverable (dimensionless)	0.5
Long run price elasticity of petroleum demand	0.95
Average lag in adjustments of petroleum demand to price (years)	15
Price of petroleum substitutes ($/bbl)	50
Average lag in development of petroleum substitutes (years)	8

Time series:†

Year	GNP (trillion 1982$)	Import price (1982$/bbl)
1990	3.94	20
2000	4.97	30
2010	6.00	40
2020	7.03	50
2030	8.06	60
2040	9.09	70
2050	10.12	80

† Linear interpolation between values.

Exhibit 8: Comparison of historical and simulated values

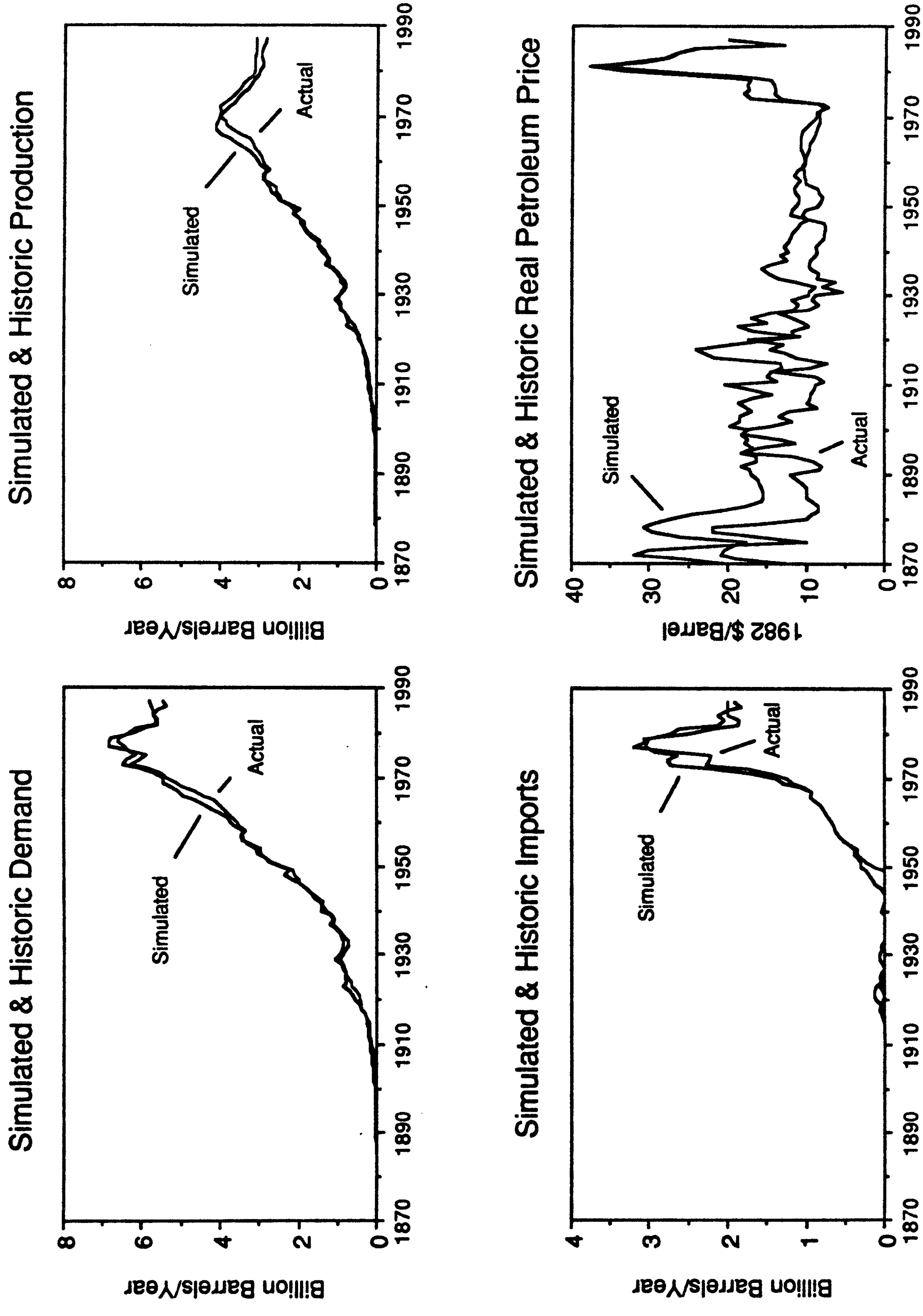

Exhibit 9: Mean-squared errors (MSE) and Theil statistics, 1900-1985

			Theil Inequality Statistics*		
	Error†	MSE	U^M	U^S	U^C
Production	11.8%	22.5E15	.000	.073	.926
Demand	14.3%	27.7E15	.015	.236	.749
Imports	160E6 bbls	24.9E15	.036	.383	.581
Price	$4.43/bbl	19.6	.407	.016	.577

† Error: Production and demand are root-mean-squared percent errors; Price and imports are root-mean-squared error.

*Theil statistics:
U^M: Fraction of MSE due to unequal means of simulated and historic values
U^S: Fraction of MSE due to unequal variance
U^C: Fraction of MSE due to unequal covariation

Exhibit 10a: Base run - Production and discoveries

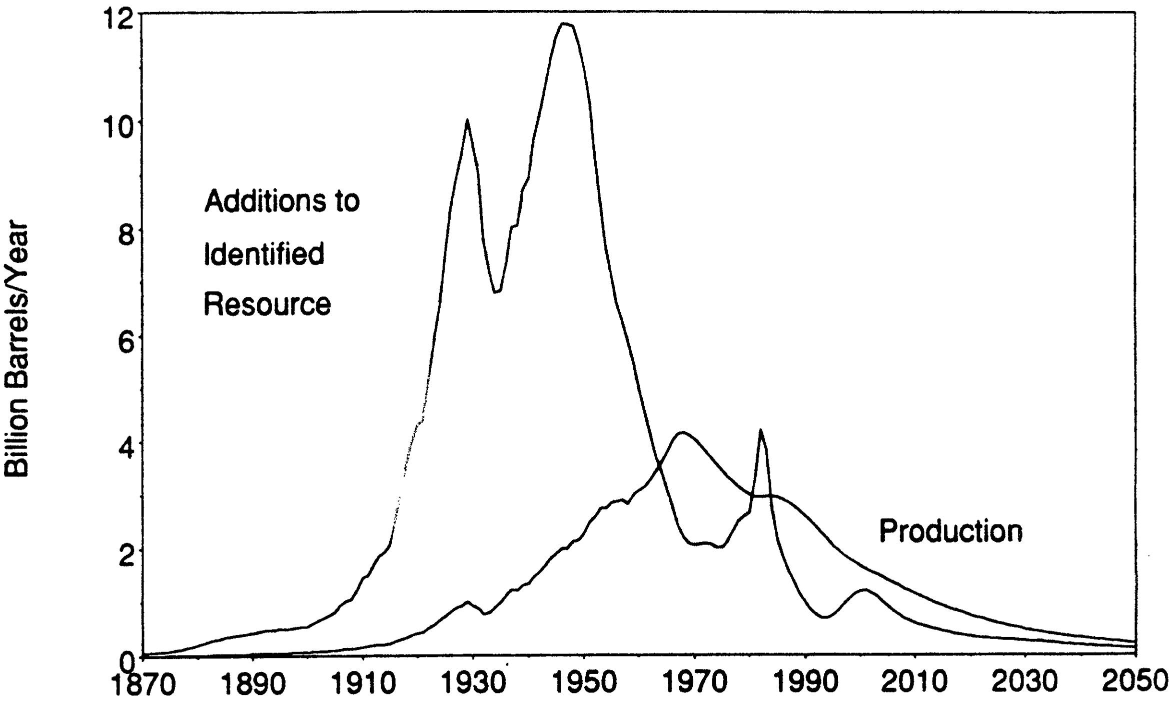

Exhibit 10b: Base run - Demand and production

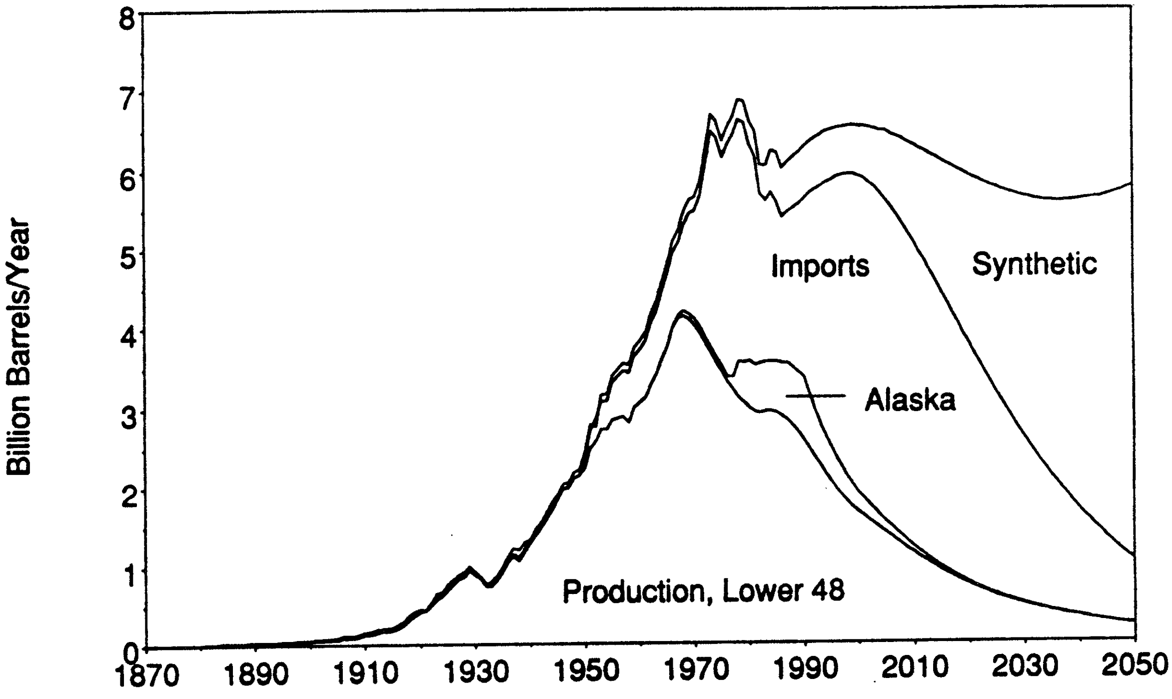

Exhibit 10c: Base run - Reserves and undiscovered resources

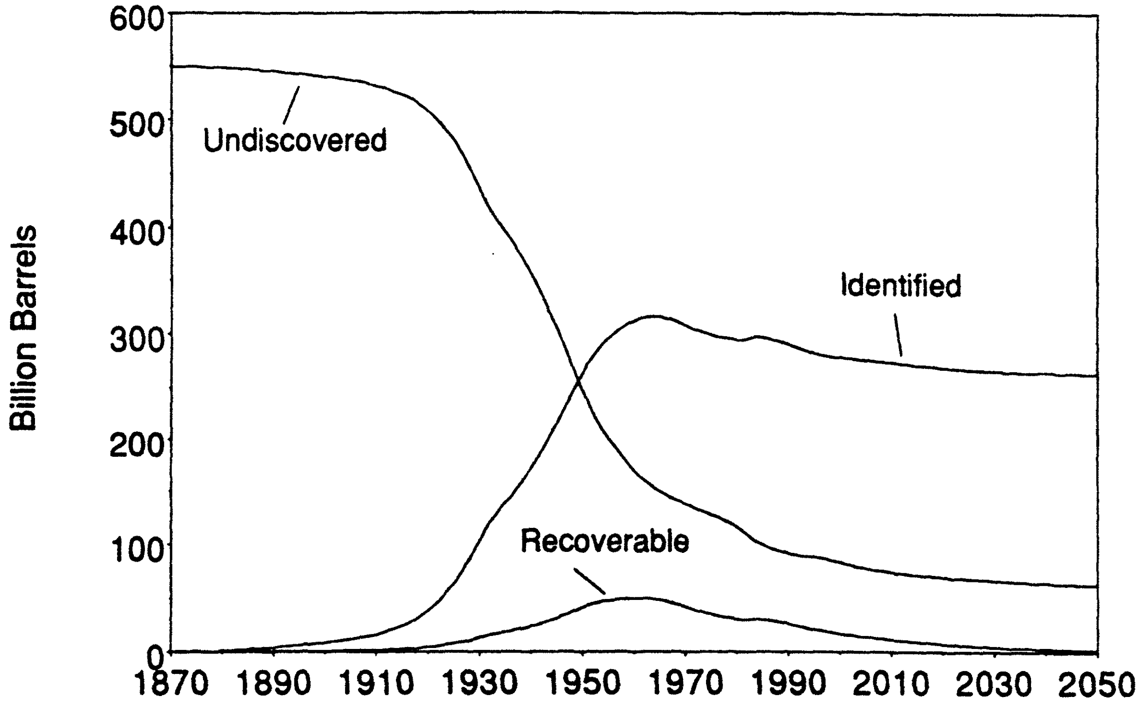

Exhibit 10d: Base run - Technology and yield

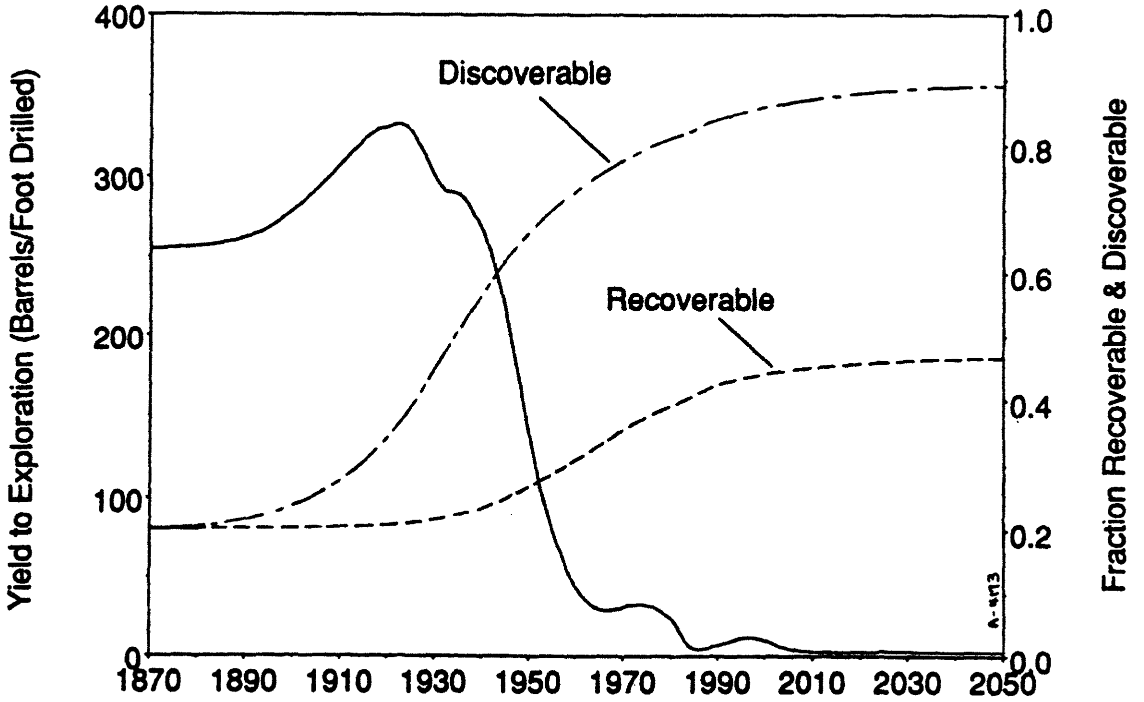

Exhibit 10e: Base run - Price (after 1990 price follows assumed world oil price; see exhibit 7)

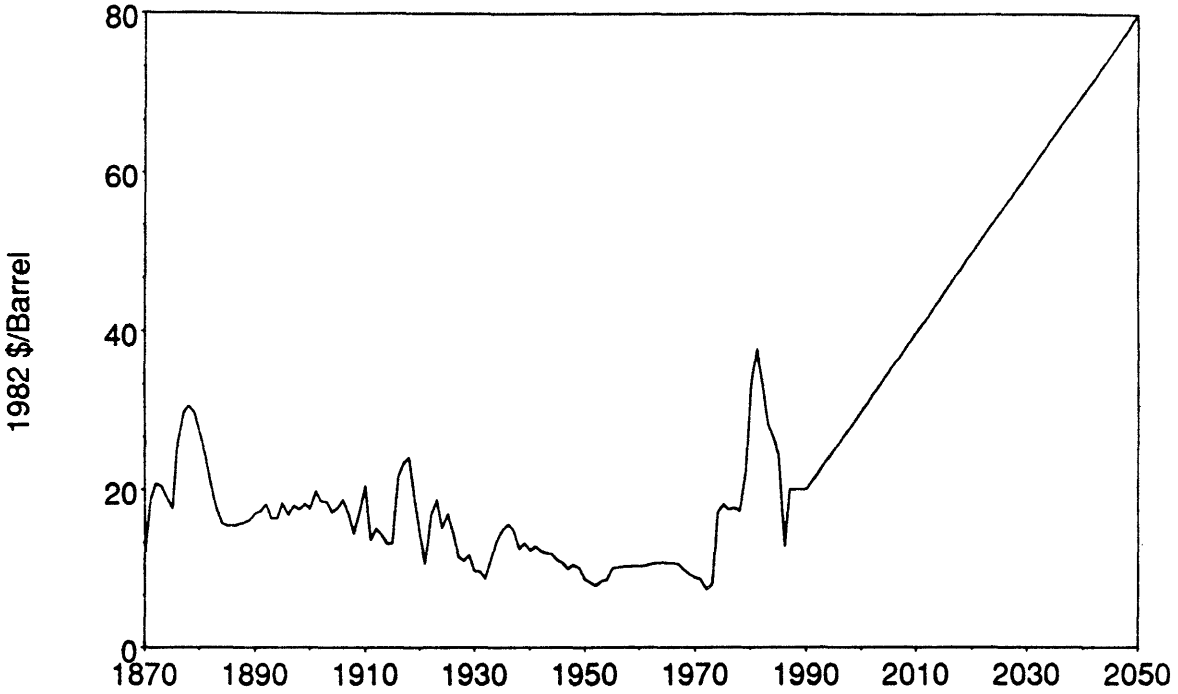

Exhibit 11. Estimates of Ultimate Recoverable Resource by the Hubbert Life Cycle Method

Range (1900 to ...)	Estimated Q^{∞} (billion bbls)	R^2	Percent Error†
1930	∞	-	∞
1935	350	.990	+53.5
1940	200	.994	-12.3
1945	165	.995	-27.6
1950	155	.996	-32.0
1955	165	.997	-27.6
1960	165	.998	-27.6
1965	165	.998	-27.6
1970	170	.998	-25.4
1975	175	.998	-23.2
1980	185	.998	-18.9
1985	197	.996	-13.6
1990	207	.995	-9.2
1995	210	.995	-7.9
2000	213	.994	-6.6
2005	217	.994	-4.8
2010	220	.989	-3.5
2015	223	.989	-2.2
2020	225	.990	-1.3

† 100*(Estimated Q^{∞} - True Q^{∞})/True Q^{∞}, where True Q^{∞} = 228 billion barrels.

Exhibit 12: 'Confidence bounds' for Hubbert estimates of the simulated URR

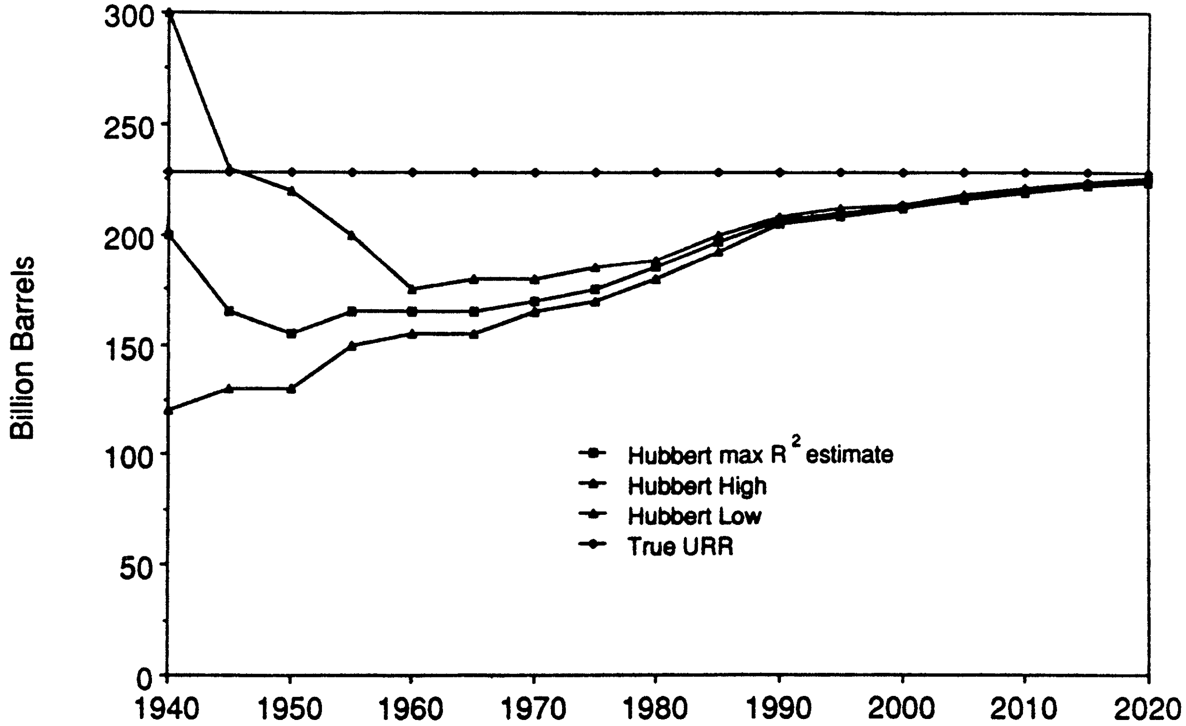

Exhibit 13: Fractional growth rate of simulated cumulative discoveries as a function of cumulative discoveries. Compare this nonlinear pattern to the linear fractional growth rate a - bQ assumed by the logistic model, dQ/dt = (a - bQ)*Q.

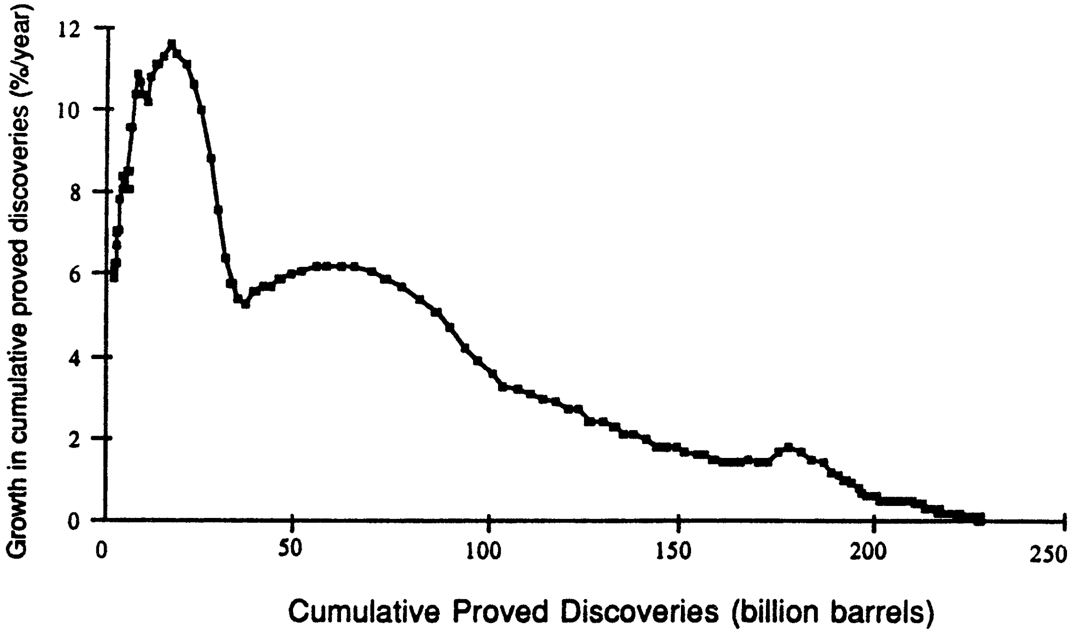

Exhibit 14. Estimates of Ultimate Recoverable Resource by the Geologic Analogy Method

Year	Geologic analogy estimate of URR (10^9 bbls)	Error† (%)
1900	7.6	-96.7
1910	13.1	-94.3
1920	28.4	-87.5
1930	71.2	-68.8
1940	194.0	-14.9
1950	417.0	82.9
1960	501.0	119.7
1970	384.0	68.4
1980	324.0	42.1
1990	315.0	38.2
2000	289.0	26.8
2010	264.0	15.8
2020	246.0	7.9

† 100*(Estimated Q^{∞} - True Q^{∞})/True Q^{∞}, where True Q^{∞} = 228 billion barrels.

Exhibit 15: Components of the geologic analogy estimate of the ultimate recoverable resource

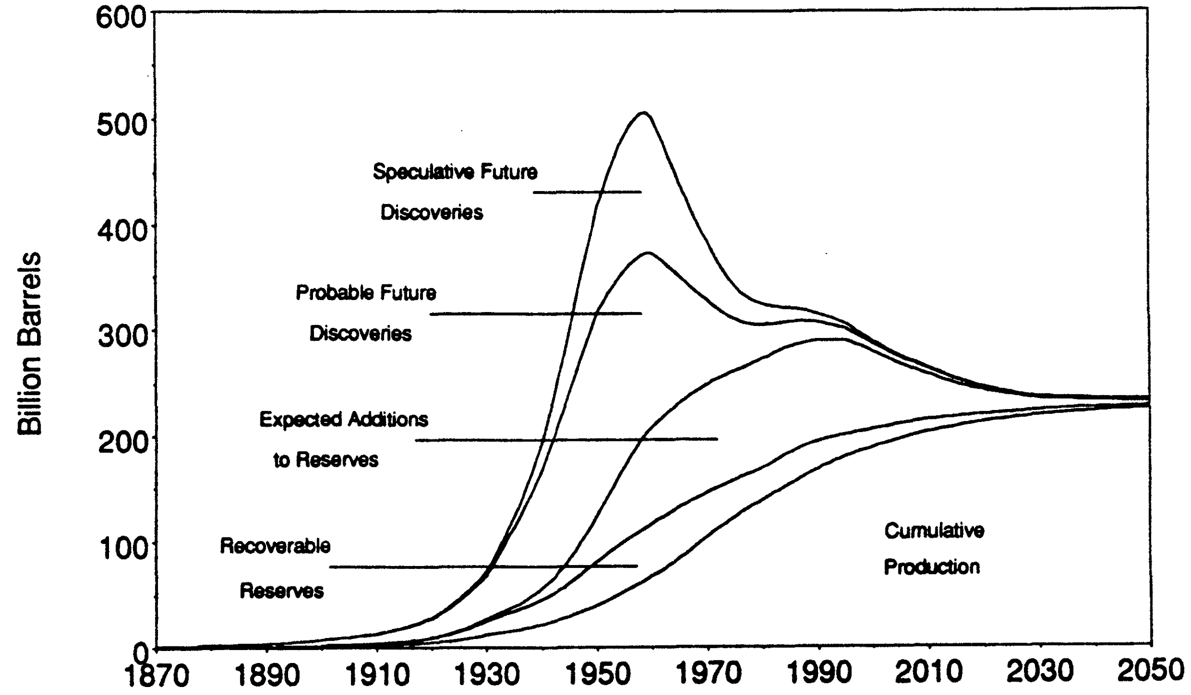

Exhibit 16: Actual and expected yield to exploration and fraction recoverable

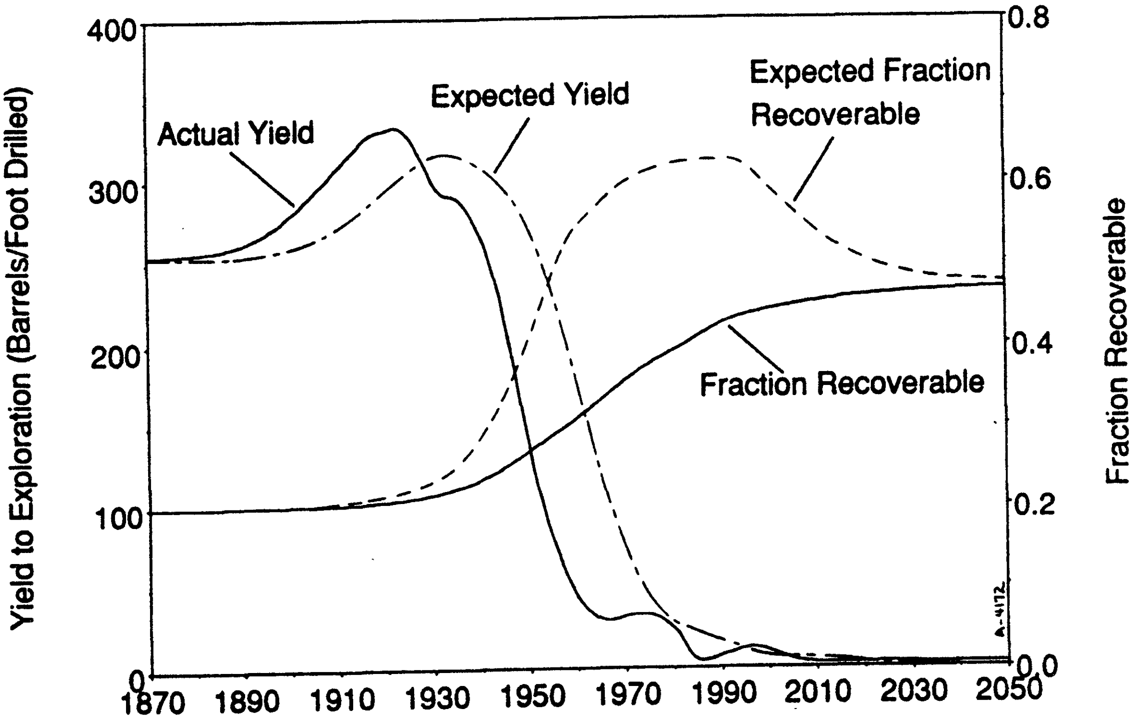

Exhibit 17: Geologic analogy estimate of the recoverable resource remaining, compared to correct value.

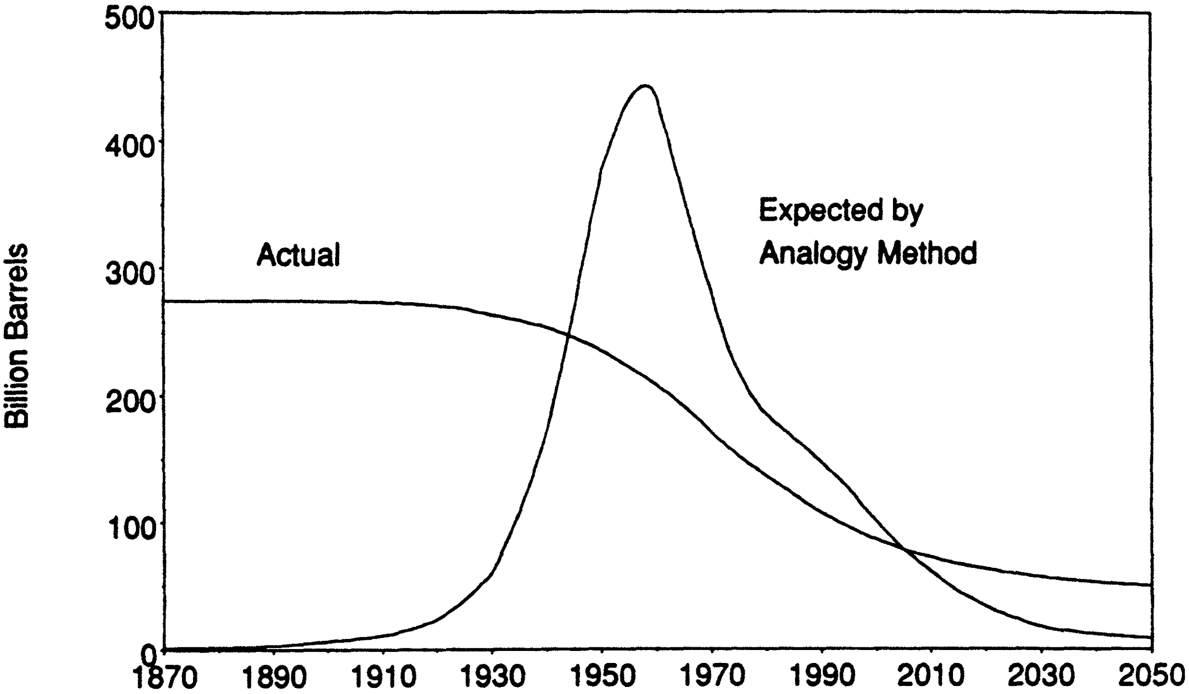

Exhibit 18: Comparison of the Hubbert and geologic analogy estimates