D-3665

# A Skeptic's Guide to Computer Models

John D. Sterman
Sloan School of Management
Massachusetts Institute of Technology

February 1985

ABSTRACT

Increasingly, citizens and policymakers are faced with the results of computer models and must make judgments about the model's relevance and validity. How can such decisions be made in an intelligent and informed manner? Can modeling be made accessible to the ordinary person or will it remain the special magic of a technical priesthood? This paper offers tentative answers to these questions. It first highlights the characteristics and capabilities of computer models such as are used in foresight and policy analysis. The advantages and disadvantages, uses and misuses of formal models are presented. What are the fundamental assumptions of the major modeling techniques? How appropriate are these techniques for foresight activities? What are the crucial questions a model user or model consumer should ask when evaluating the appropriateness and validity of a model? The paper is designed to help model consumers peek inside these computerized black boxes.

But Mousie, thou art no thy lane,
In proving foresight may be vain;
The best-laid schemes o' mice an' men
   Gang aft a-gley,
An' lea'e us nought but grief an' pain,
   For promis'd joy!

Robert Burns, "To a Mouse"

## 1. The Inevitability of Using Models

Computer modeling of social and economic systems is only about three decades old. Yet in that time, computer models have been used to analyze everything from inventory management in corporations, to the optimal distribution of fire stations in New York City, to the performance of national economies, to the interplay of global population, resources, food, and pollution. Computer models have been front page news (as in the case of Limits to Growth [Meadows et al. 1972]), have been the subject of numerous congressional hearings, and have influenced the fate of legislation. Computer modeling has become an important industry, generating hundreds of millions of dollars of revenues annually.

As computers have become faster, cheaper, and more widely available, models have become commonplace in forecasting and public policy analysis, especially in economics, energy and resources, demographics, and other crucial areas. As computers continue to proliferate, more and more policy debates will involve the results of models, both in government and the private sector. Though we are not all going to be model builders, we are all becoming the consumers of models, whether we know it or like it. The ability to understand and evaluate computer models is fast becoming a prerequisite for the policymaker, legislator, lobbyist, and citizen alike.

Each of us will be faced with the results of models and will have to make judgments about their relevance and validity. How can such decisions be made in an intelligent and informed manner? Can modeling be made accessible to the ordinary person or will it remain the special magic of a technical priesthood?

This paper offers tentative answers to these questions. It first highlights the characteristics and capabilities of computer models such as are used in foresight and policy analysis. (Models of physical systems such as the models NASA uses to test the space shuttle are not discussed.) The advantages and disadvantages, uses and misuses of formal models will be presented. What are the fundamental assumptions of the major modeling techniques? How appropriate are these techniques for foresight and policy analysis? What are the crucial questions a model user or model consumer should ask when evaluating the appropriateness and validity of a model? The paper is designed to help model consumers peek inside these computerized "black boxes."

## 2. Mental and Computer Models

Fortunately, everyone is already familiar with models. People use models every day--mental models. Our decisions and actions are based not on the true state of affairs, but on mental images of the state of the world, how the parts of the system are related, and how our actions will influence the system.

Mental models have some powerful advantages. The mental model is flexible. It can take a wide range of information into account, not just numerical data. It can be adapted to new situations and modified as new information becomes available. The great systems of philosophy, politics, and literature are, in a sense, mental models. But mental models are not easily examined by others. Assumptions are hard to pin down in debate or discussion. Interpretations differ. Ambiguities and contradictions can go unresolved.

Of more concern is the fact that people are not very good at interpreting the assumptions of their own mental models. Psychologists have shown that people can take only a few factors into account in making decisions (Hogarth 1980, Kahneman et al. 1982). People often make errors in deducing the consequences of their assumptions. Research on the behavior of people in organizations (e.g. families, businesses, the government) shows that decisions are not made by rational consideration of objectives, options, and consequences. Rather, decisions are often made by rote, using standard operating procedures that evolve out of tradition and which adjust only slowly to changing conditions (Simon 1947, 1979). These decisionmaking rules often make sense given the role of the decisionmakers in the organization, the information available to them, and the limited time available to make decisions. The problem is that individual perspectives may be parochial, information incomplete, dated, or biased, and the time available to weigh alternatives insufficient. Decisions are strongly influenced by organizational context, authority relations, peer pressure, cultural perspective, and selfish motives. As a result many decisions turn out to be incorrect because the complicated puzzle of choosing the best course of action is too difficult. Psychologists and organizational observers have identified dozens of different biases that creep into human decisionmaking as a result of cognitive limitations and organizational pressures (Hogarth 1980, Kahneman et al. 1982). Hamlet exclaims (perhaps ironically) "What a piece of work is a man, how noble in reason, how infinite in faculties...!" But it seems that people, like Hamlet himself, are simply not capable of making rational decisions without error and without being swayed by societal and emotional pressures.

Enter the computer model. Computer models offer an improvement over mental models because:

- --They are explicit, and their assumptions are open to all for review.
- --They infallibly compute the logical consequences of the modeler's assumptions.
- --They are comprehensive, and able to interrelate many factors simultaneously.

These are powerful advantages. However, in practice, many models are

- --So complex and poorly documented that no one can examine their assumptions. They become black boxes.
- --So complex the user has no confidence the assumptions are consistent or correct.
- --Unable to deal with relationships and factors which are difficult to quantify, or for which numerical data do not exist, or which are outside the expertise of the specialists who built the model.

In part because of these problems, computer models have often been misused. Models have often been used to lend authority to an argument, to justify decisions already taken, or to provide a scapegoat when a forecast turns out wrong.

How can a policymaker know what kind of model is appropriate for the problem at hand? How can a prospective model user decide whether a model is appropriate for the purpose at hand, whether its results are valid or useful? How can one guard against the misuses of models? No single or comprehensive answer can be given, but some useful guidelines can be given.

## 3. The Importance of Purpose

A model must have a clear purpose. The purpose should be to solve a particular problem. A clear purpose is the single most important ingredient for a successful modeling study. Beware the analyst who proposes to model an entire social or economic system rather than a problem. What is the difference? For example, a model designed to understand how to stabilize the business cycle is a model of a problem. A model designed to understand how the economy can make a smooth transition from oil to alternative energy sources is a model of a problem. A model that claims to be a comprehensive representation of the economy is a model of a system. Why does it matter? All models are simplifications of the real system. A truly comprehensive model would be as complex as the real system and just as inscrutable.

The art of modelbuilding is knowing what to leave out. In this context, the purpose of a model is a logical knife. It provides a criterion for deciding what to cut out, leaving only the essential features necessary to fulfill the purpose. In the example above, the comprehensive model of the economy will likely be enormous. In order to answer all questions, it will include many factors irrelevant to the business cycle such as long-term population growth or resource depletion. And it will include factors irrelevant to understanding the energy transition such as short-term changes in unemployment, inventories, and interest rates. Because of its size, it will be next to impossible to examine the assumptions. The model builders, not to mention the intended consumers of its output, are unlikely to understand its behavior, thus its validity will be largely a matter of faith.

A model designed just to examine the business cycle or energy transition, on the other hand, can be much smaller. It can be limited to those factors thought to be important in understanding business cycles or energy. Its validity for its purpose can be assessed by asking how its assumptions relate to the most important theories of the business cycle or resource economics. Of course, a model with a clear purpose can still be incorrect, large, or difficult to understand. But a clear purpose allows model users to ask the questions that can reveal the utility of a model for solving the problem at hand.

## 4. Two Kinds of Models

There are many types of models and they can be classified in many ways. Models can be static or dynamic, mathematical or physical, stochastic or deterministic. One of the most useful classifications, however, is to divide models into those that optimize versus those that simulate. The distinction between optimization and simulation models is particularly important since these types of models are suited for fundamentally different purposes.

4.1 Optimization The Oxford English Dictionary defines 'optimize' as "to make the best or most of; to develop to the utmost." The output of an optimization model is a statement of the best way in which to accomplish some goal. For example, a nutritionist would like to know how to design meals that fulfill certain dietary requirements but cost as little as possible. A salesperson must visit a certain number of cities and would like to know how to make the trip as short as possible, taking into account the available transportation between cities. Instead of trial and error, an optimization model may be used to determine the best way.

An optimization model typically consists of three parts. The objective function specifies the goal or objective. For the nutritionist, the objective is to minimize the cost of the meals. For the salesperson, it is to minimize the travel time or total mileage of the trip. The decision variables represent the choices to be made, for example the amount of potatoes in the diet or the order of cities to be visited. The optimal choices of the decision variables are the output of the optimization model. The constraints restrict the choices of the decision variables to those that are possible or acceptable. In the diet problem, the constraints would specify that consumption of each nutrient must exceed the minimum daily requirement. The constraints might also specify that you don't want potatoes more than three times a week. The constraints in the salesperson problem would specify that each city must be visited at least once, and would restrict the selection of routes to the available connections (e.g. if there were no direct flights from Boston to Cincinnati, the constraints would require you to pass through Cleveland or Pittsburgh or wherever on the way).

Thus an optimization model takes as input the goals to be met, the choices to be made, and the constraints to be satisfied. It yields as output the best decision that can be made given the assumptions of the model. Because optimization models tell you what to do in order to make the best of the situation, they are normative or prescriptive models. The purpose of an optimization model is not to tell you what will happen in a certain situation, but what ought to be done to optimize the objective.

Limitations of Optimization There are a variety of limitations and problems with optimization models which a potential user must bear in mind.

Whose Objectives? One obvious difficulty is the problem of specifying the objective function. It is clear that the dietician wants to minimize the cost of food, but what is the objective function of the mayor of New York City? How is the optimal population of the world to be defined? How can intangibles like quality of life be measured and incorporated in an objective function? How should conflicting goals and the differing agendas of special interest groups be balanced? The objective function embodies the values and preferences held to be desirable. Whose values and preferences should be used?

Because optimization is prescriptive, it always involves subjective value judgments. Users of optimization models should always scrutinize the objective function and constraints to examine the values they embody, both explicitly and by omission. For example, a water quality model may find the cheapest way to place sewage treatment plants along a river so as to meet water quality standards. The model user should ask how the model takes into account the impacts on fishing, recreation, wild species, and the development potential in the affected areas. Unless explicitly incorporated in the model, these considerations are implicitly held to be of no value.

Though difficult, the problem of choosing an objective function is not insurmountable. Intangibles like quality of life can often be quantified, at least roughly, by breaking them into measurable components. Quality of life in a city might be represented as depending on unemployment, housing adequacy, the crime rate, air quality, etc. A variety of techniques have been developed to help extract preferences from interviews and other impressionistic data. The attempt to make values explicit may itself have enormous value for the clients of a modeling project, and is a worthwhile exercise in any study.

Linearity: A more important problem relates to the verisimilitude of optimization models. Because a typical optimization problem is very complex, involving hundreds or thousands of variables and constraints, the mathematical problem of finding the optimum is extremely difficult. To render the optimization problem tractable, a number of simplifications are commonly introduced. One common simplification is to assume all the relationships in the system are linear. In fact the most popular optimization technique, linear programming, requires the objective function and all the constraints to be linear.

Linearity is convenient mathematically but almost always unrealistic. For example, a model of a firm's inventory distribution policies may contain a relationship between inventory and shipments. If the inventory of goods in a warehouse is 10% below normal, shipments may be reduced by, say, 2% because certain items will be out of stock. If the model required the relationship to be linear, then a 20% shortfall would reduce shipments by 4%, a 30% shortfall by 6%, and so on. But obviously, when the warehouse is empty (a 100% shortfall of inventory), no shipments are possible, while the linear relationship indicates shipments would be 80% of normal, an absurdity.

This may seem like a trivial example, but consider the sorry fate of the passenger pigeon, ectopistes migratorius. Before the colonization of North America, passenger pigeons were extremely abundant. Huge flocks of the migrating birds would darken the skies for days. They often caused extensive damage to crops and were hunted both as a pest and for food. For years, hunting had little impact on the population. The prolific birds reproduced fast enough to offset most losses to hunters. But the fertility of the pigeons depended nonlinearly on their population density. In large flocks they could reproduce at high rates. But in small flocks fertility dropped precipitously. As hunting gradually reduced the population, fertility fell, accelerating the decline in population. Lower population levels further lowered the birth rate, in a vicious cycle. By 1914, the passenger pigeon was extinct.

There are some techniques available to solve certain nonlinear optimization problems, and research is continuing. But in general, the nonlinearities that can be handled are limited, and the vast majority of optimization models assume the world is linear.

Lack of Feedback: Complex systems are highly interconnected. There is a high degree of feedback between sectors. For example, a water quality model may assume the sewage load to be treated is fixed, and compute the optimum size of treatment plants to be built. But if water quality improves as a result of treatment, the attractiveness of the river for development will increase, ultimately raising the sewage load. The results of the plant siting decisions feed back through the physical, economic, and social environment to alter the conditions that the policy was suited for.

A model that ignores feedback effects is said to have a narrow boundary. Such models tend to rely on exogenous variables. There are two basic kinds of variables in a model: endogenous and exogenous variables. Endogenous variables are those that are calculated by the model. They are the variables explained by the structure of the model, the variables for which the modeler has an explicit theory. Exogenous variables influence other variables in the model but are not calculated by the model. They are given simply by a set of numerical values over time. The values of exogenous variables may come from other models but are most likely the product of an unexaminable mental model.

Ignoring feedback can result in policies that are diluted, delayed, or defeated by the system, or which generate unanticipated side effects (Meadows 1982). An illustration is provided by the construction, in the 1950s and '60s, of interstate highway networks and freeways to alleviate congestion around major cities. In Boston, for example, it used to take a half an hour to drive from the neighborhood of Dorchester to the downtown area, a journey of only a few miles. With the construction of a limited access highway network, travel time dropped substantially. But by reducing congestion, outlying communities were opened up. The population in the suburbs soared. Today the rush hour journey from Dorchester to downtown often takes half an hour or more. The center city has become more congested and polluted. Its population has declined. Many businesses moved to the suburbs or were squeezed out by shopping malls. In the suburbs, farmland was paved over or turned into housing developments. The point is not to condemn these changes but to illustrate how a policy aimed at reducing highway congestion generated a wide range of side effects and was eventually undone by feedback effects which were largely unanticipated.

In theory, feedback can be incorporated in optimization models. But in practice, the resulting complexity and nonlinearity usually renders the optimization problem insoluble. As a result, many optimization models ignore most of the feedback effects. Model users should identify the degree to which important feedbacks are incorporated in the model and how excluded effects might alter the assumptions of the model and thus invalidate the results.

Lack of Dynamics: Many optimization models are static. They determine the optimal solution for a particular moment in time without regard for how the optimal state is reached or for the future evolution of the system. For example, in the late 1970s, the U.S. Forest Service constructed a linear programming model to optimize the use of government lands. The model was enormous, with thousands of decision variables and tens of thousands of constraints. It required the full use of a large computer for hours or even days at a time to find the solution. Typographical errors in the model's huge database required months of debugging. Despite the effort required, the model produced the "optimal" use of forest resources for a single moment in time. It did not take into account how harvesting a particular area would affect its future ecological development. It did not consider how land use needs or lumber prices might change in the future. It did not consider how long it would take for new trees to grow to maturity in the harvested areas, or the economic and recreational effects during this time. The model provided the optimal decisions for a single year even though those decisions would influence the development of forest resources for decades.

Not all optimization models are static. The MARKAL model, for example, is a large linear programming model designed to determine the optimal choice of energy technologies. Developed at the Brookhaven National Laboratory, the model produces as output the least-cost mix of coal, oil, gas, etc. in five-year intervals well into the next century. It requires exogenous inputs of future fuel prices, construction and operating costs for unconventional energy technologies, and energy demands. (Note that the model ignores feedbacks from energy supply to prices and demand.) The model is dynamic in the sense that it provides a "snapshot" of the optimal state of the system at five-year intervals. But it does not explain how the system moves from one optimal state to another. For example, it does not incorporate construction delays for energy production facilities, delays which are often much longer than five years. The model implicitly assumes that people, seeing what the optimal mix is for, say, the year 2010, would begin construction far enough in advance to have the required plants ready on time.

Delays are pervasive. It takes time to acquire capital plant and equipment, to clean up a waste dump, to acquire information. Delays are a major source of instability in complex systems. Delays in carrying out or perceiving the effects of decisions may cause overreaction or prevent timely intervention. Acid rain provides a typical example. Many scientists feel it will take years to determine whether and how incipient damage to the forests of New England, the Appalachians, and Bavaria is caused by acid rain or by natural forces. Until scientific and then political consensus emerges, legislative action is not likely to be strong. Implementation of pollution control programs, once passed, will take years. The lifetimes of existing power plants and other pollution sources is measured in decades. Settlement patterns and lifestyles dependent on the automobile change over even longer periods. By the time sulfur and nitrogen oxide emissions are reduced sufficiently, it may be too late.

Delays are crucial in determining the dynamic behavior of systems. But as with nonlinearity, it is difficult to incorporate delays in optimization models. When possible, delays are usually assumed to be of fixed length. The results of such models are of questionable value. Users of these optimization models may find, like the city tourist on the back roads of Maine, that "you can't get there from here."

<u>When to use optimization</u> Despite the limitations discussed above, optimization techniques can be extremely useful. But they must be used for the proper problems. Optimization models can substantially improve the quality of decisions in many areas, including computer design, airline scheduling, the location of factories, and the operation of oil refineries. Whenever the problem to be solved is one of choosing the best from among a well-defined set of alternatives, optimization should be considered. If

the meaning of "best" is also well-defined, and if the system to be optimized is relatively static and free of feedback, optimization may well be the best technique to use. Unfortunately, these latter conditions are rarely true for the social, economic, and ecological systems that are frequently of concern to decisionmakers.

Beware, however, the optimization model which purports to forecast actual behavior. The output of an optimization model is a statement of the best way to accomplish a goal. To interpret the results as a forecast of likely actual behavior is to assume that people in the real system will in fact make the optimal choices. It is one thing to say "to maximize profits the following decisions should be made" and quite another to say "people will succeed in maximizing profits, and therefore the following decisions will be made." The former is a prescriptive statement of what to do; the latter a descriptive statement of what will happen. The optimization model will only be valid for the latter purpose <u>if</u> people in fact optimize. It may seem reasonable to expect that people behave optimally--after all, it would be irrational to take second best when you could have the best. But the evidence on this score is conclusive: real people do not behave like optimization models. As discussed above, real people make decisions with simple and incomplete mental models, models that are often systematically incorrect, or that reflect goals and motives that are not captured in an optimization framework. Real people do not have the perfect information, foresight, and computational powers required to solve for the optimum solution. As Herbert Simon puts it,

> The capacity of the human mind for formulating and solving complex problems is very small compared with the size of the problem whose solution is required for objectively rational behavior in the real world or even for a reasonable approximation to such objective rationality (Simon 1957, 198).

Optimization models augment the capacity of the human mind to solve the problem of finding the objectively rational course of action. Nevertheless, even optimization models must make simplifying assumptions so as to be tractable--even with a computer the best we can hope for is a reasonable approximation to objectively rational behavior. But to model how people actually behave rather than how they ought to behave requires a different set of modeling techniques.

## 4.2 Simulation

The latin verb <u>simulare</u> means to imitate or mimic. The purpose of a simulation model is to mimic the real system so that its behavior can be anticipated and studied. A simulation model is a laboratory replica of the real system. By creating a representation of the system in the laboratory, experiments can be performed which are either impossible, unethical, or prohibitively expensive in the real world.

Simulations of physical systems are commonplace, ranging from simulations of weather patterns and the depletion of oil reservoirs to wind tunnel tests of aircraft designs. Similarly, economists and social scientists have used simulation to understand how cities evolve and respond to urban renewal policies, how energy prices affect the economy, how corporations grow, how population growth interacts with food supply, resources, and the environment. There are many different simulation techniques, including stochastic modeling, system dynamics, discrete simulation, and role-playing games. Despite the differences, all simulation techniques rely on a common approach to modeling.

A simulation model has two main components. It must include a representation of the physical world relevant to the problem under study. For example, to understand why America's large cities have continued to decay despite massive amounts of aid and numerous renewal programs, a model would need to include a representation of the physical components of the city: the size and quality of the housing stock, commercial structures, and other infrastructure; the size, skill mix, income, and other attributes of the population; the flows of people, money, etc., into and out of the city; and other factors which characterize the physical and institutional setting. The degree of detail needed depends on the specific problem to be addressed with the model. A model designed to understand why urban renewal programs have generally not worked in a variety of cities requires only an aggregate representation of the features common to such cities (Forrester 1969). But a model designed to improve the location and deployment of firefighting resources in New York City had to include a detailed representation of the streets and traffic patterns (Greenberger et al. 1976).

In addition to the physical structure of the system, a simulation model must portray the behavior of the actors in the system. Behavior in this context means the way in which people respond to different situations. The behavioral assumptions of a simulation model describe the way in which people make decisions. The decision rules are the input. The pattern of decisions is the output of the model. For example, in a pioneering simulation study of corporate behavior, Cyert and March (1963) found that department stores used a very simple decision rule to determine the floor price of goods. The rule was basically to mark up the wholesale cost of the items by a fixed percentage. When excess inventory piled up on the shelves, a sale was held and the markup was gradually reduced until the goods were sold. If sales goals were exceeded, prices were boosted. Prices were also adjusted towards those of competitors. The normal markup was determined by tradition--it adjusted very slowly towards the actual markup on the goods sold. Cyert and March found that when these rules for pricing were tested with actual store data, the model reproduced the pricing decisions of the floor managers quite well.

Thus the inputs to a simulation model are assumptions about the physical structure of the system and the procedures people use to make decisions. The state of the system determines the nature and quality of the information available to decisionmakers. The model plays the role of the decisionmakers, using the available information to mimic their decisions. These decisions then feed back and alter the state of the system, giving rise to new information and new decisions. Simulation models are "what if" tools. They are descriptive models. The purpose of a simulation model is not to tell a policymaker what should be done, but what would happen in a given situation. Often such "what if" information is more important than knowledge of the optimal decision. For example, during the 1978 debate over natural gas deregulation, President Carter's original proposal was modified dozens of times before a final compromise was passed. During the congressional debate, the Department of Energy used a system dynamics model to evaluate each version of the bill (DOE 1979). The model did not indicate what ought to be done to maximize the economic benefits of natural gas to the nation. Congress had its own ideas on that score. But by providing an assessment of how each proposal would affect gas prices, supplies, and demands, the model generated useful ammunition for the administration in lobbying for its proposals.

## Limitations of Simulation

Like any model, a simulation model is only as good as its assumptions. Naturally, a good simulation model should have an adequate representation of the physical system it represents. In general, simulation models are quite flexible and can portray the physical environment with detail and accuracy sufficient for their purposes. Unlike optimization, simulation models can easily incorporate feedback effects, nonlinearity, and dynamics. The structure of simulation models is not rigidly determined by mathematical limitations as those of optimization models often are. Indeed, one of the main uses of simulation is to identify how nonlinearities, physical delays, and the limited information available to decisionmakers interact to produce the troubling dynamics that have persistently resisted solution (for examples, see Sterman 1985, Morecroft 1983, Forrester 1969).

Accuracy of the decision rules A potential trouble spot is the accuracy of the decision rules portrayed in simulation models. Simulation models must represent human decisionmaking as it is, even if it is not optimal. The decisionmaking heuristics and strategies people use, including their limitations and errors, must be modeled. Only if a model mimics the response of decisionmakers to changing circumstances will it respond to policy interventions in the same way the real actors would. In principle there are few limitations on the accuracy of the decision rules portrayed in models. In practice, discovering those rules is often difficult. Decisionmaking rules cannot be determined from aggregate statistical data, but must be investigated first hand. Primary data on the behavior of the actors can be acquired through observation of decisionmaking in the field, that is, in the boardroom, on the factory floor, along the salesperson's route, in the household. The modeler must discover what information is available to each actor, examine its timeliness and accuracy, and infer how that information is processed to yield a decision. The skills of the anthropologist and ethnographer are often needed. Fortunately, psychologists, behavioral scientists, sociologists, and other social scientists have developed an extensive body of primary data which describes how decisionmaking is made. The best simulation modeling draws on a wide variety of disciplines as well as first hand observation of the system to elicit the decision rules of the people in the system.

Soft variables Because simulation models must portray decisionmaking as it is, they must often include variables which are difficult to quantify. It is frequently necessary to represent intangibles such as product quality, optimism, reputation, expectations, desires, and so on. Again, there is no limitation in principle to the inclusion of such soft variables, and many simulation models do. Unfortunately, some modelers limit the factors they include to those variables that are measurable, and often measurable by numerical data. This practice is sometimes defended as more scientific than "making up" the values of parameters and relationships for which no numerical data are available. Without numerical data, how can statistical tests be performed? How can parameter values be estimated?

The overwhelming majority of all data is descriptive and qualitative. And the majority of this data has never been written down. Yet they are crucial for understanding and modeling complex systems. Imagine trying to operate a school, factory, or economy solely on the basis of the available numerical information. Without the mental, descriptive knowledge of operating procedures, political subtleties, organizational structure, and so on, the result would be chaos. To leave out of a model a relationship known to be important but for which no numerical data are available is just as much an unscientific value judgment as using judgment to estimate the relationship. Ignoring the relationship implies it has a value of zero--probably the only value known to be wrong! (Forrester 1980)

Model Boundary A great strength of simulation models is the ability to capture the important feedback relationships that shape the behavior of the system and govern its response to policies. In practice, however, many models ignore factors outside the expertise of the modelbuilders or the mission of the sponsoring organization.

The consequences of omitting feedback are often serious. For example, many energy models assume the economy is unaffected by the price of energy. The PIES model (Project Independence Evaluation System), used in the 1970s by the Federal Energy Administration, and later by the Department of Energy, provides a typical example. The PIES model assumed that economic growth, the costs of unconventional fuels, interest rates, inflation, and world oil prices were all unaffected by domestic energy prices, production, or policies. A full embargo of imported oil or doubling of oil price would have no impact on the economy, according to the model. Yet the FEA described the model's purpose in the following way:

> [Energy] strategies are evaluated in terms of their impact on:
>
> * Development of alternative energy sources
> * Vulnerability to import disruptions
> * Economic growth, inflation, and unemployment
> * Environmental effects
> * Regional and social impacts (FEA 1974, 1)

By treating the economy exogenously, the PIES model was inherently contradictory. The model showed that the investment needs of the energy sector would rise substantially as depletion raised the development costs of new sources of oil and as synthetic fuels were developed. But at the same time, the model assumed that the higher investment needs of the energy sector could be satisfied without reducing investment or consumption in the rest of the economy and without raising interest rates or inflation. In effect, the model let the economy have its pie and eat it too. In part because it ignored the feedbacks between the energy sector and the rest of the economy, the PIES model consistently proved to be overoptimistic. In 1974 the PIES model projected that by 1985 the US would be well on the way to energy independence. Energy imports would be only 3.3 million barrels per day, production of shale oil would be 250,000 barrels per day, all at an oil price of about $22 per barrel (1984 dollars) and with vigorous economic growth. In reality oil imports are about 5.5 million barrels per day. A shale oil industry remains a dream. All this despite huge reductions in oil demand caused by oil prices that have exceeded $30 per barrel and the most serious recession since the Great Depression (see the appendix in Stobaugh and Yergin 1979 for a good discussion of the limitations of the PIES and other energy models).

Narrow model boundaries are not limited to energy analysis. The Global 2000 report (CEQ 1980) showed that most of the models used by government agencies rely significantly on exogenous variables. Population models assumed food production was exogenous. Agriculture models assumed that energy prices and other input prices were exogenous. Energy models assumed that economic growth and environmental conditions were exogenous. Economic models assumed that population and energy prices were exogenous. And so on. Because important intersectoral feedbacks were ignored, the models produced inconsistent results.

A broad model boundary that includes important feedback effects is more important in a model than a great amount of detail in the specification of individual components. It is worth noting that the PIES model provided breakdowns of energy supply, demand, and price for dozens of fuels, each for different regions of the country. Yet its aggregate projections for 1985 aren't even close. One can legitimately ask what purpose was served by the effort devoted to forecasting the demand for jet fuel or naphtha in the Pacific Northwest when the basic assumptions were so palpably inadequate and the main results so woefully erroneous. (In fairness, the PIES model is not unique in the magnitude of its errors. Nearly all energy models, of all types, have consistently been wrong about energy production, consumption, and prices. The evidence shows clearly that energy forecasts actually lag behind the available information, reflecting the past rather than anticipating the future [DOE 1983].)

### 4.3 Econometrics

Strictly speaking, econometrics is a simulation technique. But it deserves separate discussion for two reasons. First, econometrics evolved out of economics and statistics, while most other simulation techniques emerged from engineering or operations research. The difference in pedigree leads to large differences in purpose and practice. Second, econometrics is one of the most widely used formal modeling techniques. Pioneered by Nobel Prize winning economists Jan Tinbergen and Lawrence Klein, econometrics is taught in nearly all business and economics programs, and ready-to-use statistical routines for econometric modeling are now available for many personal computers. Econometric forecasts are regularly reported in the nation's media.

Econometrics is literally the measurement of economic relations, and originally involved statistical analysis of economic data. As commonly practiced today, econometric modeling consists of three stages. These are specification, estimation, and forecasting. In the first step, the structure of the model is specified. Structure means the set of relations between variables, both those that characterize the physical setting and those that describe behavior. For example, an econometric model of the macroeconomy will typically contain accounting relations that specify how GNP is composed of consumption, investment, government activity, and international trade. It also will include behavioral equations that describe how these quantities are determined. The Phillips curve is an example of such a behavioral relation. If the model contains a Phillips curve, one of the equations will specify that the rate of inflation depends on the amount of unemployment. Presumably the modeler expects that high unemployment reduces inflation and vice-versa. An econometric model will typically consist of a set of such equations, with many interrelationships between the variables. For example, another equation may relate unemployment to the demand for goods, the wage level, worker productivity, etc. Still other equations may explain these in terms of other factors. A large econometric model may have hundreds or even thousands of equations.

Not surprisingly, econometrics draws on economic theory to guide the specification of models. The validity of the models thus often depends on the validity of the underlying economic theory. Though there are many flavors of economic theory, a small set of basic assumptions about human behavior are common to most (especially modern neoclassical theory and the "rational expectations" school). These include:

<u>Optimization</u>: People (economic agents, in the jargon) are assumed to be concerned with just one thing--maximizing their profits. Consumers are assumed to maximize the "utility" they derive from their resources. Decisions about how much to produce, what goods to purchase, whether to save or borrow, are assumed to be the result of optimization by individual decisionmakers. "Non-economic" considerations (defined as any behavior which diverges from profit or utility maximization) are ignored or treated as local aberrations and special cases.

<u>Perfect information</u>: To optimize, economic agents need accurate information about the world. The information needed goes beyond the current state of affairs. Also needed is complete knowledge of the available options and their consequences. For example, to determine the optimal mix of labor, machines, energy, and other inputs in the production process, a firm must know not only the wages of workers and the prices of machines and other inputs, but also how much could be produced with different combinations of people and machines, even if those combinations have never been tried. Such knowledge is assumed to be freely and accurately known in most economic models. Many go further, assuming people know not only the current situation, but future prices, technologies, and possibilities as well, including the ability to perfectly anticipate the consequences of their own actions or those of competitors.

<u>Equilibrium</u>: The pioneers of mathematical economics were primarily concerned with the net result of optimization by individuals and firms. The net result defines the equilibrium of the market or economy. The crucial questions of theory involved the nature of the equilibrium state for different situations. Given people's preferences and the technological possibilities for producing goods, at what prices will commodities be traded, and in what quantities? What will wages be? What will profits be? How will a tax or monopoly power influence the equilibrium? These questions proved difficult enough without tackling the more difficult

problem of dynamics. Indeed, dynamic theory, including the recurrent fluctuations of the business cycle, of the growth and decline of industries and nations, of inflation, remained primarily descriptive and qualitative long after equilibrium theory was completely mathematized. Consequently, dynamic behavior in economics tends to be seen as a transition from one equilibrium to another. The transition is usually assumed to be stable.

The rich heritage of static theory in economics left a legacy of equilibrium for econometrics. Many econometric models assume markets are in equilibrium at all times. When adjustment dynamics are modeled, variables are usually assumed to adjust in a smooth and stable manner toward the optimal, equilibrium value. The lags are nearly always fixed in length. For example, most macroeconometric models assume the capital stocks of firms in the economy adjust to the optimal, profit maximizing level with a fixed lag of several years. The lag is the same whether the industries that supply investment goods have the capacity to meet the demand or not (see, for example, Eckstein 1983 and Jorgenson 1963). Yet clearly, when the supplying industries have excess capacity, orders can be filled rapidly. When capacity is strained, customers must wait in line for delivery. Analysis shows that there are significant differences between a model that assumes a fixed investment lag regardless of the physical capability of the economy to fill the demand and one that explicitly models the determinants of the investment delay (Senge 1980). In general, models that explicitly portray delays and their determinants will yield different results from models that simply assume smooth adjustments from one optimal state to another.

Economists acknowledge the idealization and abstraction of their assumptions about human behavior, information, and equilibrium, but point to the powerful results that have been derived from them. However, a growing number of prominent economists argue that these assumptions are not only abstract but false. In his Presidential address to the Royal Economics Society, E. H. Phelps-Brown said:

> The trouble here is not that the behaviour of these economic chessmen has been simplified, for simplification seems to be part of all understanding. The trouble is that the behaviour posited is not known to be what obtains in the actual economy. (Phelps-Brown 1972, 4)

Nicholas Kaldor of Cambridge University is even more blunt:

> ...in my view, the prevailing theory of value--what I called, in a shorthand way, 'equilibrium economics'--is barren and irrelevant as an apparatus of thought.... (Kaldor 1972, 1237)

As mentioned earlier, a vast body of empirical research in psychology and organizational studies has shown that people do not optimize or act as if they optimize, that they don't have the mental capabilities to optimize their decisions, that even if they had the computational powers necessary they lack the information needed to optimize. Instead, they try to satisfy a variety of personal and organizational goals, use standard operating procedures to routinize decisionmaking, and ignore much of the available information to reduce the complexity of the problems they face. Herbert Simon, in his acceptance speech for the 1978 Nobel Prize in economics, concludes:

> There can no longer be any doubt that the micro assumptions of the theory--the assumptions of perfect rationality--are contrary to fact. It is not a question of approximation; they do not even remotely describe the processes that human beings use for making decisions in complex situations. (Simon 1979, 510)

The second stage in econometric modeling is statistical estimation of the parameters of the model. The parameters determine the precise strengths of the relationships specified in the model structure. In the case of the Phillips curve, for example, having assumed in advance that unemployment affects inflation, the modeler would then use the past data on inflation and unemployment to estimate precisely how strong that relationship has been. Sophisticated statistical routines are used to estimate the parameters of the model. In essence, these routines, known generally as regression, are simply fancy curve-fitting techniques. They use the historical data to find the parameter values that best match the data itself, for example, matching the inflation rate in terms of the unemployment rate.

The use of statistical procedures to derive the parameters of the model is the hallmark of econometrics, and distinguishes it from other forms of simulation. All modeling methods must specify the structure of the system and estimate parameters. But the focus in econometrics on statistical parameter estimation to the exclusion of other techniques imposes a strong discipline on the model builder. It gives econometricians an insatiable appetite for numerical data, for without numerical data the statistical procedures used to estimate the models are useless. It is no accident that the rise of econometrics went hand in hand with the quantification of economic life. For example, the development of the national income and product accounts by Simon Kuznets in the 1930s was a major advance in the codification of economic data, permitting consistent measures of economic activity at the national level for the first time. To this day all major macroeconometric models rely on the national accounts data, and indeed, macroeconomic theory itself has adapted to the national accounts framework.

It is obvious that policy evaluation and foresight depend on an accurate knowledge of the state of the world and of its history. Econometrics has been a valuable stimulus to the development of much-needed data gathering and measurement by government and private companies alike. But at the same time, the relentless focus on numerical data blinds econometric modelers to less tangible but no less important factors. Econometric models portray the behavior of people. But aggregate statistical data measure only the result of the decisions made, not how or why those decisions were made. Statistical data do not reveal the nature and quality of the information people used to make decisions, and therefore models based on such data cannot be used to indicate how changes in that information would alter future decisions.

Reliance on statistical procedures to estimate the parameters forces econometricians to exclude from their models variables for which no numerical data exist such as soft variables and unobservable concepts like desires, goals, perceptions, and so on. Potentially observable quantities that haven't been measured must also be ignored or handled with proxy variables for which data do exist. For example, the literacy of a population may be proxied by education expenditures per capita, though the connection between the two may be tenuous.

Another problem is the failure of econometric techniques to distinguish between causal relationships and correlations. Simulation models must portray the causal relationships in the system if they are to mimic its behavior, especially in new situations or in response to new policies. But the statistical techniques used to estimate parameters in econometric models only reveal the degree of past correlation between the variables. Statistical techniques can never tell the modeler whether a relationship is causal. The problem in using correlations is that the correlations may change or shift as the system evolves (Lucas 1976 makes the same point in a different context). Consider the Phillips curve as an example. The Phillips curve stopped working sometime in the early 1970s. Inflation rose and at the same time unemployment worsened. Many economists argued that structural change had occurred. By structural change they meant that the underlying causal structure of the system had changed. In fact, the Phillips curve was never a structural relationship at all--it never represented the causal forces that determine inflation or wage increases. Rather, the Phillips curve was nothing more than a way of restating the past behavior of the system. In the past, the curve said, low unemployment had tended to occur at the same time inflation was high, and vice-versa. Naturally, when the inflation of the 1970s swept prices to levels unprecedented in the industrial era and as people learned to expect continuing price increases, the historical correlation broke down. The behavior of the system had changed. But the underlying structure of causal relationships need not have changed. As inflation worsened, causal relationships that had been present all along but which were dormant in an era of low inflation gradually became active determinants of behavior. In particular, the ability of people to adapt to continuing inflation existed all along but was not tested until inflation became high enough and persistent enough. (These causal relationships involved learning to deal with high inflation through indexing, COLAs, inflation-adjusted accounting, etc.--they were the result of an adaptive feedback process of learning). Because econometric models rely on historical correlations, a modeler's appeal to "structural change" usually means the inadequate structure of the model had to be altered because it failed to anticipate the behavior of the system.

A related problem caused by the reliance on statistical estimation arises from the limited range of historical data usually available. Aggregate statistical data do not provide a guide to behavior outside the historical range of experience or under a different set of policies or incentives. Historical relationships are assumed to remain valid in the future. Consequently, many econometric models are not robust--changes in policies or conditions that carry the system outside the range of historical data often cause the models to break down. To illustrate, in 1979 the DRI model was used to test policies to eliminate oil imports. The model assumed that the response of oil demand to the price of oil was rather weak--a ten percent increase in oil price caused a reduction of oil demand of only two percent, even in the long run. To reduce oil consumption by 50 percent (enough to cut imports to zero at the time), the model indicated that oil had to rise to $800 per barrel. Yet at that price, the annual oil bill for the remaining 50% would have exceeded the total GNP for that year (see Sterman 1981). Today, with the benefit of hindsight, economists agree that oil demand is much more responsive to price than was earlier believed. But considering the behavior of the model in extreme conditions could have revealed the inconsistency of the original assumptions much earlier.

The validation of econometric models is also strongly influenced by the reliance on numerical data. Because the micro-level data that describe how decisions are made are commonly ignored in econometrics, the criterion for the goodness of an equation or model becomes the degree to which it fits the data. (The model's predictive accuracy is also a criterion, but this is never known in advance--at best one knows how well a model predicted in the past.) The statistical routines used to estimate parameters indicate the degree of fit between the estimated and actual variables, and tell the modeler if the relationship between the variables is statistically significant. When a relationship fails to be significant, the modeler may try another specification for the equation, hoping for a better statistical fit. Without recourse to the descriptive, micro-level data, the resulting equations may be ad hoc and bear only slight resemblance to either economic theory or actual behavior. Alternatively, the discrepancy between the model and data may be explained by faulty data, exogenous influences, or other factors. The Phillips curve again provides an example. When the Phillips curve broke down, numerous revisions of the equations used to predict inflation were made, with limited success. Some analysts pointed to the oil price shock, Russian wheat deal, or other one-of-a-kind events as the explanation for the change. Still others argued that there had been structural change which caused the Phillips curve to shift out to higher levels of unemployment for any given inflation rate. Others argued that the Phillips "curve" was really a vertical line--that in the long run, the rate of inflation was solely dependent on monetary policy and had no relationship to unemployment at all.

Econometrics texts teach that the statistical significance of an equation is an indicator of the correctness of the relationship (e.g. Pindyck and Rubinfeld 1976). But this is a mistaken view. Statistical significance does not mean a relationship is a correct or true characterization of the way the world works, but simply indicates how well the equation fits the observed data. A statistically significant relationship indicates the variables in the equation are highly correlated--and that the apparent correlation is not likely to have been the result of mere chance. But it does not indicate that the relationship is causally correct or even that it is causal at all. While the criterion of statistical significance as a yardstick for judging models seems plausible, failure to find a statistically significant relationship may simply indicate that there aren't enough data, or that the data don't contain enough information to allow the statistical procedures to discriminate between competing hypotheses. Or there may be statistical limitations. The regression procedures used to estimate parameters only yield unbiased estimates under certain conditions. These conditions are known as maintained hypotheses because they are assumptions which must be made in order to use the statistical technique. The maintained hypotheses can never be verified, even in principle, but must be taken as a matter of faith. In the most common regression technique, ordinary least squares, the maintained hypotheses include the unlikely assumptions that the variables are all measured perfectly, that the model being estimated corresponds perfectly to the real world, and that the random errors in the variables from one time period to another are completely independent. More sophisticated techinques do not impose such restrictive assumptions. But they always invoke other hypotheses, shifting the locus of the inevitable a priori, but never eliminating it.

The restrictive assumptions and mixed results of econometrics have generated serious criticism from within the economics profession. Phelps-Brown notes that because controlled experiments are generally impossible in economics, "running regressions between time series is only likely to deceive" (Phelps-Brown 1972, 6). Lester Thurow notes that econometrics has failed as a method for testing theories and is now used primarily as "a showcase for exhibiting theories." But as a device for advocacy, econometrics imposes few constraints on the prejudices of the modeler. Thurow concludes

> By simple random search, the analyst looks for the set of... variables and functional forms that give the 'best' equations. In this context the 'best' equation is going to depend heavily upon the prior beliefs of the analyst. If the analyst believes that interest rates do not affect the velocity of money, he finds a 'best' equation that validates his particular prior belief. If the analyst believes that interest rates do affect the velocity of money, he finds a 'best' equation that validates this prior belief. (Thurow 1983, 107-8)

But the harshest assessment of all comes from Nobel laureate Wassily Leontief:

> Year after year economic theorists continue to produce scores of mathematical models and to explore in great detail their formal properties; and the econometricians fit algebraic functions of all possible shapes to essentially the same sets of data without being able to advance, in any perceptible way, a systematic understanding of the structure and the operations of a real economic system. (Leontief 1982, 107; see also Leontief 1971)

But surely such theoretical problems matter little if the econometric models provide accurate predictions. Unfortunately, econometrics fails on this score as well. The predictive power of econometric models, even over the short-term (one to four years) is poor and virtually indistinguishable from that of other forecasting methods. There are several reasons for the failure to predict accurately.

To forecast, the modeler must provide estimates of the future values of the exogenous variables, that is, those variables which influence the other variables in the model but which are not in turn influenced by the model.

An econometric model may have dozens of exogenous variables. Each must be forecast before the model can be used to predict. The source of the forecasts for these variables may be other models, but is usually the intuition and judgment of the modeler. Ensuring consistency, much less correct forecasts for the exogenous variables, is difficult.

Often the forecasts produced by the models don't square with the modeler's intuition. Many modelers, including those at the "big three" econometric forecasting firms, Data Resources, Inc., Chase Econometrics, and Wharton Econometric Forecasting Associates, routinely adjust their forecasts whenever they feel the model output is wrong. This fudging, or "add-factoring," as they call it, is extensive: the late Otto Eckstein of Data Resources admitted that their forecasts were "60% model and 40% judgment" (Wall Street Journal, 17 February 1983). "'There is no way of telling where the Wharton model leaves off and [model developer] Larry Klein takes over'" according to another economist (Business Week, 30 March 1981). Worse, the adjustments are often colored by the personalities of the modelers:

> "Mr. Eckstein concedes that sometimes his forecasts reflect an optimistic view. Data resources,...'is the most influential forecasting firm in the country,' he declares. 'If it were in the hands of a doom-and-gloomer, it would be bad for the country.'" (Wall Street Journal, 17 February 1983)

Add-factoring has been attacked by other economists as unscientific. The mental models used to add-factor, though they are the mental models of seasoned experts, are subject to the same cognitive limitations other people face. And whether good or bad, the assumptions behind add-factoring are unexaminable.

In a shocking experiment, the Joint Economic Committee of Congress (through the politically neutral General Accounting Office) asked the three leading econometric forecasting firms (DRI, Chase, and Wharton) to make a series of simulations with their models. One set of forecasts was "managed" or add-factored by the forecasters at each firm. The other set consisted of pure forecasts, made by the GAO, to examine the untainted results of the models. The models were run under different assumptions about monetary policy. As an illustration of the inconsistencies revealed by the experiment, consider the following: When the money supply was assumed to be fixed, the DRI model forecast that after ten years, the interest rate would be 34 percent, a result totally contrary to both economic theory and historical experience. The forecast was then add-factored down to a more reasonable 7 percent. The other models fared little better, revealing both the inability of the pure models to yield meaningful results and the extensive ad-hoc adjustments made by the forecasters to render the results palatable (JEC 1982).

The failures of econometric models have not gone unnoticed. A representative sampling of recent articles in the business press on economics and forecasting includes the following headlines:

"1980: The year the forecasters really blew it"
(Business Week, 14 July 1980)

"Where the big econometric models go wrong"
(Business Week, 30 March 1981)

"More or less oil will go up or down or maybe it won't:
Energy experts are gun-shy after forecasts,
for years, haven't turned out well"
(Wall Street Journal, 5 May 1982)

"Where have all the answers gone?
Economists seem bankrupt just when their ideas are needed most"
(Time, 17 January 1983)

"Economists, too, find themselves in disarray"
(US News & World Report, 7 Feb. 1983)

"Forecasters overhaul models of economy in wake of 1982 errors"
(Wall Street Journal, 17 Feb. 1983)

"Business forecasters find demand is weak in their own business
Bad predictions are factor"
(Wall Street Journal, 7 Sept. 1984)

"Economists missing the mark: more tools, bigger errors"
(The New York Times, 12 Dec. 1984)

The result of these failures has been to erode the credibility of all computer models no matter what their purpose, not just econometric models designed for prediction. This is unfortunate, for foresight does not depend on the ability to predict the future. In fact, there is substantial agreement among modelers of global problems that exact, point prediction of the future is neither possible nor necessary (Meadows et al. 1982, 279):

> ...at present we are far from being able to predict social-system behavior, except perhaps for carefully selected systems in the very short term. Effort spent on attempts at precise prediction is almost surely wasted, and results that purport to be such predictions are certainly misleading. On the other hand, much can be learned from models in the form of broad, qualitative, conditional understanding--and this kind of understanding is useful (and typically the only basis) for policy formulation....If your doctor tells you that you will have a heart attack if you do not stop smoking, this advice is helpful, even if it does not tell you exactly when a heart attack will occur or how bad it will be.

When to use econometrics: Econometric models do not seem to be well-suited to the types of problems of concern in poicy analysis and foresight. The prime purpose of econometric models is short-term prediction of the exact future state of the economy. Most of the attributes of econometrics have evolved in response to this need, including the reliance on regression techniques to pick the "best" parameters from the available numerical data, the extensive reliance on exogenous variables, and add-factoring. Though in practice econometric models do not predict very well, they are about as good as anything else for that purpose.

Though econometric models purport to simulate human behavior, they in fact rely on unrealistic assumptions about the motivations of real people and the information available to make decisions. Though they must represent the physical world, they commonly ignore dynamic processes, disequilibrium, and the physical basis for delays between actions and results. Though they may incorporate hundreds of variables, they ignore soft variables and unmeasured quantities. Foresight is most often concerned with longer time horizons than are common in econometrics. The feedback relationships between environmental, social, and demographic factors are usually as important as economic influences. Often the numerical data needed to model these effects are not available. The need to consider the long term means the system is likely to leave the historical region of behavior, making historical correlations an unreliable basis for analysis.

## 5. Checklist for the Model Consumer

The preceeding discussion has focused on the limitations of various modeling approaches in order to provide potential model consumers with a sense of what to look out for when choosing a model. Despite the limitations of the various modeling techniques, there is no doubt that computer models can be and have been extremely useful foresight tools. Well-built models offer significant advantages over the often faulty mental models currently in use.

To further assist the model consumer, the following checklist presents key questions a model user should ask to help evaluate the appropriateness of a model for a particular purpose.

--What is the purpose of the study? What problem does the model address?

--What is the boundary of the model? What factors are endogenous? Exogenous? Excluded? Are soft variables included? Are feedback effects properly taken into account? Does the model capture possible side effects, both harmful and beneficial?

--What is the time horizon relevant to the problem? Does the model include as endogenous components factors that may change significantly over the time horizon?

--Are people assumed to act rationally and to optimize their performance? Does the model take non-economic behavior into account (organizational realities, non-economic motives, political factors, cognitive limitations)?

--Does the model assume people have perfect information about the future and about the way the system works, or does it take into account the limitations, delays, and errors in acquiring information that plague decisionmakers in the real world?

--Are appropriate time delays, constraints, and possible bottlenecks taken into account?

--Is the model robust in the face of extreme variations in input assumptions?

--Are the policy recommendations derived from the model sensitive to plausible variations in its assumptions?

--Are the results of the model reproducible? Or are they adjusted ("add-factored") by the model builder?

--Is the model documented, and is the documentation publicly available?

## 6. Conclusions

The arguments above have crucial implications for the design of governmental and private foresight and policy analysis capabilities. Foresight requires the intelligent use of different models designed for specific purposes, not a single, comprehensive model of the world. Foresight is not a well-intentioned way to back into an Orwellian world of centralized control. To repeat a dictum offered above, "Beware the analyst who proposes to model an entire social or economic system rather than a problem." It is simply not possible to build a single, integrated model of the world, into which mathematical inputs can be inserted and out of which will flow a coherent and useful understanding of world trends. To be used responsibly, models must be subjected to review and debate. To foster that process, a cross-disciplinary approach is needed. Models designed by experts in different fields and for different purposes must be compared, contrasted, and criticized. The foresight process should foster such review.

The history of global modeling provides a good example of such a process. The initial global modeling efforts, published in World Dynamics (Forrester 1971) and The Limits to Growth (Meadows et al. 1972) provoked a storm of controversy. A number of critiques appeared, and soon after, other global models were developed. Over ten years, the International Institute for Applied Systems Analysis (IIASA), near Vienna, conducted a program of analysis and critical review designed to bring the modelers together. Six major symposia were held. Eight major global models were examined and discussed. The models had different purposes, used a range of modeling techniques, and were built by persons with widely varying backgrounds. There remain large areas of methodological and substantive disagreement. Yet despite the enormous differences in perspective, consensus has emerged on a number of crucial issues. These include:

(1) The physical and technical resources exist to satisfy the basic needs of all the world's people into the foreseeable future.

(2) Population and material growth cannot continue forever on a finite planet.

(3) Continuing "business as usual" policies in the next decades will not result in a desirable future, or even the satisfaction of basic human needs.

(4) Technical solutions alone are not sufficient to satisfy basic needs or create a desirable future. (Paraphrased from Groping in the Dark, Meadows et al. 1982)

The IIASA program on global modeling represents the most comprehensive effort to date to use computer models as a way to bootstrap human understanding. It has created agreement on crucial issues where none existed. It has guided research and sped progress. It offers a model for the effective conduct of foresight in both the public and private sectors.

The primary function of modelbuilding should be educational rather than predictive. No one should make decisions on the basis of a computer model whose results are simply presented, take 'em or leave 'em.

Towards that end, the role of modeling should be redefined as a process rather than as a technology for producing an answer. The common mode of computer-based analysis, in which a study is commissioned by a client who then waits for the final report, largely ignorant of the methods, assumptions, and biases that go into the conclusions, virtually guarantees failure. Such a procedure places the policymaker in the role of a supplicant before the oracle, awaiting the prophecy. Like King Croesus before the Oracle at Delphi, there is a nearly overwhelming temptation for policymakers to interpret such pronouncements in accordance with their preconceptions, or easier still, to simply ignore unfavorable results.

Worse, the model-as-oracle attitude so prevalent today rightfully alarms many who see blind acceptance of models as an abdication to the computer of the responsibility for judgments that should be human (Weizenbaum 1976). Models should not be used as a substitute for critical thought, but as a tool for improving judgment and intuition.

Yet for all the pitfalls of formal modelbuilding, it must be remembered that the alternative is continued reliance on the mental models that have failed to resolve the pressing problems with which public policy is concerned. While far from perfect, the computer model is often superior to the alternative mental models currently in use.

Indeed much of the value of formal models derives from the difference between the results of the formal model and those of the mental model. By exploring the reasons for the differences between the results of the mental and formal models, both can be improved. Improving the mental models upon which decisions are ultimately based is the proper goal of computer modeling. The success of such a dialectic, however, depends on the ability to understand the assumptions of the computer model. Foresight must foster that dialectic and stimulate education, aided by the computer, but ultimately relying on informed human judgment, not computer printouts.

ACKNOWLEDGMENTS

Many of the ideas expressed here emerged from discussions with or were first formulated by, among others, Jay Forrester, George Richardson, Peter Senge, and especially Dana Meadows, whose book Groping in the Dark was particularly helpful.

REFERENCES

Council on Environmental Quality, Gerald O. Barney (ed.) (1981) The Global 2000 Report to the President, (3 vol) Washington, D.C., CEQ.

Cyert, R., and March, J. (1963) A Behavioral Theory of the Firm, Englewood Cliffs: Prentice Hall.

Department of Energy (1983) Energy Projections to the Year 2000. Washington: Department of Energy, Office of Policy, Planning, and Analysis.

Department of Energy (1979) National Energy Plan II. DOE/TIC-10203. Washington, DC: Department of Energy, May 1979.

Eckstein, O. (1983) The DRI Model of the US Economy. New York, McGraw Hill.

Federal Energy Administration (1974) Project Independence Report, Washington, D.C., FEA.

Forrester, J.W. (1969) Urban Dynamics, Cambridge, The MIT Press.

Forrester, J.W. (1971) World Dynamics, Cambridge, The MIT Press.

Forrester, J.W. (1980) "Information Sources for Modeling the National Economy," Journal of the American Statistical Association, 75 (371), September, 555-574.

Greenberger, et.al (1976) Models in the Policy Process, New York, Russell Sage Foundation.

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

Joint Economic Committee (1982) Three Large Scale Model Simulations of Four Money Growth Scenarios. Prepared for subcommittee on Monetary and Fiscal Policy, 97th Congress 2nd Session, Washington, D.C.

Jorgenson, D. W. (1963) "Capital Theory and Investment Behavior," American Economic Review, 53, 247-259.

Kahneman, D., Slovic, P., and Tversky, A. (1982) Judgement Under Uncertainty: Heuristics and Biases, Cambridge, Cambridge University Press.

Kaldor, N. (1972) "The Irrelevance of Equilibrium Economics," The Economic Journal, 82, 1237-55.

Leontief, W. (1971) "Theoretical Assumptions and Nonobserved Facts," The American Economic Review, 61, (1) March, 1-7.

Leontief, W. (1982) "Academic Economics," Science, 217, 9 July, 104-107.

Lucas, R. (1976) "Econometric Policy Evaluation: A Critique," The Phillips Curve and Labor Markets, Brunner, K. and A. Meltzer (eds), Amsterdam, North-Holland.

Meadows, et.al. (1972) The Limits to Growth, New York, Universe Books.

Meadows, D.H. (1982) "Whole Earth Models and Systems," CoEvolution Quarterly, 34 (Summer 1982), 98-108.

Meadows, D.H., J. Richardson, and G. Bruckmann (1982) Groping in the Dark. Somerset, NJ: Wiley.

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

Phelps-Brown, E. H. (1972) "The Underdevelopment of Economics," The Economic Journal, 82, March, 1-10.

Pindyck, R., and Rubinfeld, D. (1976) Econometric Models and Economic Forecasts, New York, McGraw Hill.

Simon, H. (1947) Administrative Behavior, New York, MacMillan.

Simon, H. (1957) Models of Man, New York, Wiley.

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

Sterman, J. (1985) "A Behavioral Model of the Economic Long Wave," Journal of Economic Behavior and Organization, 6, 1.

Sterman, J. (1981) "The Energy Transition and the Economy: A System Dynamics Approach," unpublished PhD dissertation, Cambridge, MIT.

Thurow, L. (1983) Dangerous Currents, New York, Random House.

Weizenbaum, J. (1976) Computer Power and Human Reason: from Judgment to Calculation. San Francisco: W.H. Freeman.