Analytical Techniques for Studying Substitution Among Materials (1982)

Pages
199

Size
8.5 x 10

ISBN
0309328454

Committee on Materials Substitution Methodology; National Materials Advisory Board; Commission on Engineering and Technical Systems; National Research Council

ANALYTICAL TECHNIQUES FOR STUDYING SUBSTITUTION AMONG MATERIALS

REPORT OF THE
COMMITTEE ON MATERIALS SUBSTITUTION METHODOLOGY

NATIONAL MATERIALS ADVISORY BOARD
COMMISSION ON ENGINEERING AND TECHNICAL SYSTEMS
NATIONAL RESEARCH COUNCIL

Publication NMAB-385
National Academy Press
Washington, D. C.
1982

NOTICE: The project that is the subject of this report was approved by the Governing Board of the National Research Council, whose members are drawn from the Councils of the National Academy of Sciences, the National Academy of Engineering, and the Institute of Medicine. The members of the committee responsible for the report were chosen for their special competences and with regard for appropriate balance.

The report has been reviewed by a group other than the authors according to procedures approved by a Report Review Committee consisting of members of the National Academy of Sciences, the National Academy of Engineering, and the Institute of Medicine.

The National Research Council was established by the National Academy of Sciences in 1916 to associate the broad community of science and technology with the Academy's purposes of furthering knowledge and of advising the federal government. The Council operates in accordance with general policies determined by the Academy under the authority of its congressional charter of 1863, which established the Academy as a private, nonprofit, self-governing membership corporation. The Council has become the principal operating agency of both the National Academy of Sciences and the National Academy of Engineering in the conduct of their services to the government, the public, and the scientific and engineering communities. It is administered jointly by both Academies and the Institute of Medicine. The National Academy of Engineering and the Institute of Medicine were established in 1964 and 1970, respectively, under the charter of the National Academy of Sciences.

This study by the National Materials Advisory Board was conducted under Contract No. J0199138 with the Bureau of Mines of the U.S. Department of Interior.

This report is for sale by the National Technical Information Service, Springfield, Virginia 22151.

Printed in the United States of America.

ABSTRACT

At the request of the U.S. Bureau of Mines, the Committee on Materials Substitution Methodology assessed the available analytical modeling techniques for their utility in treating and interpreting the mechanisms of materials substitution. The techniques were classified under eight categories, viz., extrapolation, case studies, optimization, econometrics, input/output, simulation, judgmental studies, and engineering studies. Invited papers on the techniques were presented and discussed at a workshop. Through extensive deliberations, the workshop and committee examined the strengths and weaknesses of the techniques and ranked each from three points of view: requirements of the technique, potential uses, and criteria for selection. The rankings assigned were high, medium, and low; they are displayed in tabular form in three tables.

While offering no explicit recommendations, the committee reached a number of conclusions. It concluded that no single technique can provide a comprehensive means of forecasting and analyzing substitution among materials. However, a good grasp of the strengths and weaknesses of the available techniques is necessary to understand the effects of various factors on materials substitution. It also concluded that the analyst is more important than the technique. Some modeling techniques are better suited to retrospective analysis (e.g., econometrics). Others are more suitable for accounting for future uncertainties and prospects (e.g., simulation or judgmental models). Materials substitution is most appropriately analyzed in terms of the long run.

In summary, the report outlines the potential of the available techniques for expanding our ability to treat and manage the substitution of one material for another. In light of this potential, the committee believes that the techniques examined merit vigorous development, refinement, and augmentation.

82-0142

0993199 PB83-215947

Analytical Techniques for Studying Substitution among Materials

(Final rept)

National Materials Advisory Board (NRC), Washington, DC.

Corp. Source Codes: 045528000
Sponsor: Bureau of Mines, Washington, DC.
Report No.: NMAB-385
Sep 82 190p
Languages: English Document Type: Conference proceeding
NTIS Prices: PC A09/MF A01 Journal Announcement: GRAI8319
Country of Publication: United States
Contract No.: J0199138

The available analytical modeling techniques are assessed for their utility in treating and interpreting the mechanisms of materials substitution. The techniques are classified under eight categories, viz., extrapolation, case studies, optimization, econometrics, input/output, simulation, judgmental studies, and engineering studies. Invited papers on the techniques were presented and discussed at a workshop. The report outlines the potential of the available techniques for expanding our ability to treat and manage the substitution of one material for another.

Descriptors: *Meetings; *Materials; *Substitution; Extrapolation; Simulation; Methodology; Econometrics; Forecasting; Optimization
Identifiers: NTISNASNRC
Section Headings: 8I (Earth Sciences and Oceanography--Mining Engineering); 48A (Natural Resources and Earth Sciences--Mineral Industries); 71GE (Materials Sciences--General)

PREFACE

The current interest in materials substitution has several sources. An important one is that most industrial nations, including the United States, are net importers of many materials. As such, they are vulnerable to disturbances of prices and supplies, including disturbances induced by producer nations, acting singly or as cartels. Materials-importing nations, therefore, are strongly interested in potential substitutes for imported commodities. In particular, they are interested in the time required to make specific substitutions and the disruptions likely to result from such shifts.

In the longer run, the possible depletion of natural resources creates a need to be able to assess the substitutability of reproducible or relatively abundant materials for scarce materials. It is also essential to be able to manage limited stocks of natural resources in the optimum manner.

Against the backdrop of such concerns, the National Materials Advisory Board*, at the request of the Bureau of Mines, formed the Committee on Materials Substitution Methodology. The committee's task was to assess the available analytical tools for their utility in treating and interpreting the mechanisms of materials substitution. Typical issues which these analytical techniques might be used to address include:

> At current prices, and assuming no radical change in technology or consumer preference, what will be the market share of competing materials in specific applications at various points in the future?
>
> If a disruption occurred in the market for a material, what would be the substitution dynamics in the end use applications? (i.e., How much could be substituted for at various prices? How long would it take to make the substitutions?)
>
> What will be the effects of technological change on the competition between materials in specific applications (e.g., How will the requirements for higher operating temperatures in the future affect the competition between cobalt and nickel in superalloys?)?

In addition to assessing the strengths and weaknesses of the techniques, the committee was to evaluate the potential for new research directions and the desirability of developing new analytical techniques.

To achieve these goals, the committee invited 30 to 40 persons to a workshop on June 11-12, 1980. The workshop was held at Woods Hole, Massachusetts, to minimize disruptive influences. Invited papers were presented during the first day of the workshop. The presentations were followed by deliberations by separate working groups that continued into the following day. During the last half of the second day a plenary session was held to hear and to discuss the reports of the working groups.

The purpose of this meeting was to examine the strengths and weaknesses of the available substitution methodologies when applied to national materials requirements, and to the forecasting and analysis of materials markets. The attendees were called upon also to define the methodologies needed for these problems and, where such methodologies did not exist, to discuss whether or not they could be developed. In addition, comments pertaining to the development of new data bases or the refinement of existing ones were solicited.

This report is a compilation of the workshop's findings (Part 1), the invited papers (Part 2), and the summary and conclusions of the committee.

The committee wishes to extend its thanks and appreciation to its technical advisors and other participants in the Woods Hole workshop. Particular thanks are due those who presented formal papers. The contributions of all who attended were invaluable to the committee's work.

Joel P. Clark, Chairman

* The National Materials Advisory Board is a unit of the Commission on Engineering and Technical Systems of the National Research Council. Its general purpose is the advancement of materials science and engineering in the national interest.

COMMITTEE ON MATERIALS SUBSTITUTION METHODOLOGY

Chairman

JOEL P. CLARK, Department of Materials Science and Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts

Members

MARTIN L. BAUGHMAN, Department of Electrical Engineering, University of Texas at Austin

JAMES C. BURROWS, Charles River Associates, Inc., Boston, Massachusetts

GILBERT Y. CHIN, Physical Metallurgy and Ceramics, Bell Laboratories, Inc., Murray Hill, New Jersey

HAROLD E. GOELLER, Oak Ridge National Laboratory, Oak Ridge, Tennessee

MORRIS A. STEINBERG, Technology Applications, Lockheed Aircraft Corporation, Burbank, California

JOHN E. TILTON, Department of Mineral Economics, The Pennsylvania State University, University Park, Pennsylvania

CHARLES S. TUESDAY, Research Laboratories, General Motors Corporation, Warren, Michigan

Technical Advisors

KENNETH C. HOFFMAN, Mathtech, Inc., Arlington, Virginia

HAROLD A. LINSTONE, Futures Research Institute, Portland State University, Portland, Oregon

Liaison Representatives

JOHN D. CORRICK, Bureau of Mines, U.S. Department of Interior, Washington, D.C.

KENNETH W. MLYNARSKI, Bureau of Mines, U.S. Department of Interior, Washington, D.C.

JOHN D. MORGAN, Bureau of Mines, U.S. Department of Interior, Washington, D.C.

ROBERT MEHRABIAN, Metallurgy Division, National Bureau of Standards, Washington, D.C.

MARGARET SLADE, Bureau of Economics, Federal Trade Commission, Washington, D.C.

NMAB Staff

RICHARD M. SPRIGGS, Staff Scientist

CONTENTS

FIGURES AND TABLES

SUMMARY AND CONCLUSIONS

SUMMARY

There is concern in the industrialized countries of the world about future problems in the markets for non-fuel minerals. These potential problems may be classified into two categories: resource depletion, and sudden interruptions in prices or supplies.

Substitution among materials is an important option that is often available to any industrialized society. In theory it can provide solutions to both of the potential problems mentioned above. However, in spite of its importance, there is much to learn about materials substitution. In particular, the driving forces for substitution, the time delays necessary for substitution, and interdependence of materials requirements are not well understood in most mineral markets.

There are a number of techniques that have been or can be applied to the study and analysis of materials substitution. The Committee on Materials Substitution Methodology has classified them under eight headings:

- Extrapolation
- Case Studies
- Optimization
- Econometrics
- Input/output
- Simulation
- Judgmental Studies
- Engineering Studies

The committee has examined these methods of analysis and ranked them in terms of their relative strengths and weaknesses. The methods are essentially modeling techniques, and the committee has ranked them from three points of view: requirements of the technique; potential uses; and criteria for selection for particular uses. The rankings assigned are high, medium, and low; they are displayed in tabular form in Tables 1, 2, and 3 (pp. 4-6).

CONCLUSIONS

The committee concludes from its deliberations that no single technique can provide a comprehensive means of forecasting and analyzing of substitution among materials. That is, there is no unique method which can be used to forecast substitution among materials or to analyze (1) the implication of a substitution on such factors as the demands for competing or complementary level, greater understanding of, and ability to develop, substitution strategies could promote effective use of natural resources. Furthermore, such techniques could enable industries to identify the points of their greatest sensitivity to materials supply and to act accordingly.

In sum, this report outlines the strengths and weaknesses of the techniques which are currently available as tools for forecasting and analysis of materials substitution.

Table 1. Comparison of Requirements of the Technique

	Extrapolation	Case Studies	Optimization	Econometrics	Input/ Output	Simulation	Judgmental	Engineering
Data Requirements	Low	High	High	High	High	Medium	Medium	High
Costs - Building	Low	Medium	Medium	Medium	Medium	High	Medium	Medium
Costs - Maintenance	Low	Low	Medium	Medium	High	Medium	Medium	Medium
Costs - Using and Documentation	Low	Low	High	Medium	High	High	Medium	High
Costs of Calculation	Low	NA	Medium	Medium	High	Medium	Low	Medium
Personnel	Low	Medium	Medium	Medium	Medium	Medium	High	Medium
Development Time	Low	High	High	Medium	High	High	Low	High

Requirements: Low means that the requirements are simple.
High indicates that the requirements are complex or detailed.
NA: Not Applicable

Table 2. Comparison of Potential Uses of the Techniques

	Extrapo-lation	Case Studies	Optimi-zation	Econo-metrics	Input/Output	Simu-lation	Judg-mental	Engi-neering
Forecasting Short Range	Medium	NA	Low	Medium	NA	Low	High	Medium
Forecasting Medium Range	Low	NA	Medium	Low	NA	Medium	Low	Low
Forecasting Long Range	Low	NA	Medium	Low	NA	Medium	Low	Low
Impact Analysis Technology	NA	Medium	High	Low	NA	High	Low	High
Impact Analysis Economic	NA	Medium	High	Medium	NA	High	Low	Medium
Impact Analysis Policy/Regu-latory	NA	Medium	Medium	Medium	NA	High	Low	Low
Materials Selection	NA	Low	Medium	Low	NA	Low	Medium	High
Historical Analysis	NA	High	Medium	High	Low	Medium	Medium	Medium
Implications of Substi-tution	NA	Medium	Medium	Low	Medium	Medium	Medium	NA

Potential Uses: Low means more difficult to use or less effective
High means easier to use or more effective
NA means not applicable

Table 3. Comparison of Criteria for Selection

	Extrapo-lation	Case Studies	Optimi-zation	Econo-metrics	Input/Output	Simu-lation	Judg-mental	Engi-neering
Application Desired:								
Forecasting	Medium	NA	Low	Medium	NA	Low	High	Low
Analysis	NA	Medium	High	Medium	NA	High	Low	High
Budget and Time Constraints	Low	High	Medium	Medium	Medium	High	Medium	High
Sensitivity to Changes in:								
Technology	Low	Low	High	Medium	Medium	Medium	Medium	Medium
Economic Factors	Low	Low	Low	High	Medium	Medium	Medium	Medium
Transparency:								
To Naive User	High	High	Medium	Low	High	Low	High	High
To Expert User	High	High	High	Medium	Medium	Medium	Low	Medium
Ability to Validate:								
Internal to Method	High	Medium	Low	High	Low	Low	Medium	Low
External to Method	Medium	Low	Medium	High	Medium	Medium	Low	Low
Level of Aggregation	High	Low	Medium to High	High	Medium to High	Medium to High	Low to High	Low

Selection Criteria: For application desired, a high rating is favorable. For those characteristics that are desirable, i.e., transparency, ability to validate, and level aggregation, a high rating is good. For those characteristics that are undesirable, i.e., budget and time constraints, sensitivity to changes in technology or economic factors, a high rating is bad.

NA means not applicable

INTRODUCTION

The essential role of mineral resources in the industrial and defense-related industries of the United States is well recognized. The adverse effects of supply disruptions and discontinuous changes in prices became painfully clear in the last decade, especially in the energy markets. Now there is considerable concern among the industrialized nations of the world about the potential for the occurrence of similar events in the non-fuel mineral markets in the current decade.

Potential problems related to non-fuel mineral markets may be classified into two categories: resource depletion, and sudden interruptions in mineral prices or supplies. There is some concern that resource depletion will severely constrain economic growth in the not-so-distant future. For example, it has been suggested (Skinner 1976 & 1979) that the few abundant materials in the earth's crust (e.g., iron, aluminum magnesium, titanium) must inevitably replace the scarcer ones in most engineering applications, and that the cost of extraction of geochemically scarce elements will undergo steep increases as lower grade ores are processed.

On the other side of the debate are those who argue that substitution among materials and technological change are the mechanisms that will provide our society with an inexhaustible source of raw materials in the future (Goeller 1976, Smith 1979). Indeed a number of economic studies have shown that substitution and new technologies have reduced the real prices of raw materials over the last century (Smith 1979).

Interruptions in the prices or supplies of minerals have occurred in the past due to a number of factors, including the formation of cartels, military actions, civil disturbances, embargoes and natural disasters. The potential for supply interruptions and/or large price increases has increased in a number of mineral markets in recent years. Of particular concern is the increasingly complex and unstable political situation in the central and southern regions of Africa. For instance, the invasion of the Shaba province of Zaire in 1978 by Katangan rebels was a major factor in the approximately seven-fold increase in the dealer price of cobalt during a one year period. Moreover, there are a number of other important materials, on which the United States is almost entirely dependent, that are produced predominantly in southern Africa. These include chromite, manganese, platinum, and palladium.

The ability to substitute for materials that are in short supply or that are subject to rapid price escalations is an important option that should be available to any industrialized society. That is not to say that substitution is technically feasible in every instance or that it is economically feasible in the range of historic materials prices. However, if material substitution were to occur in an ideal manner, we could expect to see a solution to the problem of long-run resource depletion (Goeller 1976) as well as a means of ameliorating the adverse consequences of sudden market disruptions. Despite its importance, we have much to learn about materials substitution. For instance, the primary driving forces for material substitution--relative prices, technological change, government regulation, consumer preferences--vary from material to material and from application to application. These driving forces have not been analyzed in depth for most material markets. Other factors which are usually poorly understood include the time response necessary for substitution to occur at various relative price ratios, the reversibility of the substitution process, and the interdependent nature of material requirements in particular, applications.

The ability of material users to substitute back and forth among materials as relative prices change, i.e., the reversibility, varies from application to application. There is a change in relative price that is required to induce substitution in the opposite direction once the original substitution has occurred. In many instances, a greater relative price is required to effect a substitution in the opposite direction because of the necessity to make expenditures for capital equipment in the original substitution process. For instance, if cobalt-base superalloys are substituted for by nickel-base superalloys, there will be an attendant change in requirements for "minor" alloying elements such as tantalum and columbium.

The purpose of this report is to critically assess the various methodological approaches which are currently available as tools for forecasting and analyzing substitution among materials with the following questions in mind:

What is now available?

What are the strengths and weaknesses of each approach?

What is needed to improve our capabilities?

The objectives of this effort were:

(1) To describe approaches which are suitable for forecasting and analysis of substitution among materials over the short, medium, and long terms.

(2) To provide an opportunity for the presentation and review of papers dealing with the use of the major methodological approaches in a forum where representatives of alternative points of view were present (i.e., advocates of different methodologies).

(3) To identify limitations which are inherent in the methodological approaches themselves or which result from applying a model in a particular context.

(4) To describe approaches which are suitable for forecasting and analysis of substitution among materials over the short, medium, and long terms.

The Committee on Materials Substitution Methodology classified the available methodological approaches under eight headings:

- Extrapolation
- Case Studies
- Optimization
- Econometrics
- Input/Output
- Simulation
- Judgmental Studies
- Engineering Studies

The committee pursued its efforts primarily by means of a two-day workshop. Invited papers on the various techniques were presented and then were analyzed by small working groups made up of invited participants as well as committee members. The committee used the reports of these analyses and further extensive deliberations to arrive at its consensus findings.

This report of the committee's findings is in two parts. Part 1 includes discussions of each of the eight analytical techniques examined and an evaluation of the relative strengths and weaknesses of each. The committee consensus as to the strengths and weaknesses is also presented in tabular form to permit ready assessment of the utility of a given technique for a given purpose. Part 2 of this report comprises the invited papers at the workshop.

PART 1

WORKSHOP RESULTS

WORKSHOP RESULTS

EXTRAPOLATION

Extrapolative or pure time-series forecasting techniques use the historic behavior of a variable to predict its future. The distinguishing feature of these techniques is the forecasting function, which depends only on time. Usually, the variable of interest is observed at discrete, equally spaced intervals, and some systematic pattern is identified. This pattern is then forecast to continue. The simplest example of an extrapolative technique is linear trend analysis--fitting a straight line to historical data and projecting the continuation of the fitted trend. However, much more sophisticated techniques can be used to estimate a variety of nonlinear trends as well as cyclical deviations from these trends.

For extrapolative techniques to be useful, historical relationships must be relatively stable; we must assume that the future will be like the past, a possibility that becomes increasingly unrealistic as the forecasting period lengthens. Therefore, extrapolative techniques are often more appropriate for short-term than for long-term forecasting.

The Use of Extrapolative Techniques

In the area of materials substitution, extrapolative techniques have been applied most often to market-share data. If X_1 and X_2 represent consumption of competing materials for a particular purpose, then S_1, the market share of the first material, is $S_1 = X_1/(X_1+X_2)$ and $S_2 = 1 - S_1$. Extrapolative techniques can be used to predict the evolution of these shares, over time.

Several consequences of formulating the analysis in this way are:

(1) The market must be defined; i.e., the area of potential substitution must be known in advance.

(2) Appropriate units must be chosen for measuring X_1 and X_2.

(3) The forecasting function can only assume values between 0 and 1.

Market definition can be complicated by changes in technology and in consumer tastes. Choosing appropriate units presents a problem because materials possess many properties (strength, ductility, electrical and thermal conductivity, for example), but one summary measure of the materials' utility must be selected. The requirement that the forecasting function assume only values between 0 and 1 implies that linear trend fitting is rarely justified (market shares eventually become negative or greater than 1). However, a linear relationship can be used to forecast the consumption ratio, X_1/X_2.

Trends can be assumed to conform to a variety of nonlinear functional forms--exponential, logarithmic, or sinusoidal, for example. However, by far the most common functional form used to study trends in market shares is the logistic form,

$$S_1 = [1 + \exp - K\ (t-u)]^{-1},$$

where, at $S_1 = .5$, $t = u$, and the rate of change of the fraction substituted is .25K. The assumption underlying the use of a logistic-curve trend is that the percentage change in the new material's market share will be proportional to the old material's market share. S_1 initially increases slowly, but, as the new material takes hold, the increase is more rapid. However, as the market is saturated, the increase in S_1 slows again. The percentage change in S_1's market share decreases monotonically. However, the absolute change increases and then decreases.) When S_1 is plotted as a function of time, it takes on the familiar "S" shape. Fisher and Pry (1971) fitted logistic curves to 17 historic examples of substitutions (synthetic for natural fibers, for example) and found that, in each case, the "S" shape gave a reasonably good approximation to the data.

Fisher and Pry's method provides a description of the substitution process that has a certain intuitive appeal. However, it does not explain why some substitutions progress more rapidly than others. The logistic function is a family of curves, each completely defined by two parameters, K and u. Several authors have attempted to model the determination of the parameters K and u (or, equivalently, the speed of the new material's or product's penetration of the market). Mansfield (1961) looked at the percentage of firms adopting an innovation and found that differences in adoption rates were explained by the profitability of installing the innovation and by the size of the investment required to install it. He also looked at other determinants--the durability of capital, the rate of growth of the firm,and the phase of the business cycle--but found that these factors were statistically insignificant. Ayres, Noble, and Overly (1971) used the logistic curve to forecast interindustry patterns of trade (changes in input-output coefficients over time) and concluded that changes in relative prices explain differences in rates of substitution. They assume that (1) the price ratio, P_1/P_2, decays exponentially over time, (2) the percentage change in consumption ratio is linearly related to the percent range in price ratio, and showed that these assumptions imply logistic-curve time patterns for market shares.

More complicated time-series forecasting techniques attempt to explain a variable's systematic deviations from its fitted trend. For example, both Box-Jenkins and spectral techniques examine the behavior of a detrended time series (Mansfield 1961). (A Box and Jenkins forecasting equation consists of the ratio of two polynomials in the backshift operator. Detailed discussions are found in Box and Jenkins (1976). Spectral techniques are discussed by Bloomfield (1976). Box-Jenkins techniques look at behavior in the time domain and relate present values of a variable to its history. These techniques are particularly useful for predicting responses to random shocks. Spectral techniques look at behavior in the frequency domain. They decompose a time series into a set of orthogonal components, each associated with a frequency, and show the contribution of each frequency band to the total variance of the variable. Box-Jenkins and spectral techniques have been used extensively to forecast a variety of time series but, to the committee's knowledge, they have not been used to model substitution.

Strengths and Weaknesses

The principal advantage of extrapolative forecasting models is the extreme ease with which they can be constructed. With a logistic-curve trend for S_1, the market share of product 1, $S_1/(1-S_1)$ is an exponential function of time, and, therefore, its logarithm is a linear function of time. Ordinary least squares can thus be used to estimate the two parameters in the equation. If the trend assumes some other commonly used functional form, estimation is apt to be equally simple. In addition, data requirements are very modest. All that is needed is time-series observations on X_1 and X_2--consumption of the competing materials in the area of interest.

Because the models are so simple, however, their usefulness for forecasting and analysis is fairly limited. Extrapolative techniques have been used to forecast market shares over long periods (several decades). However, any major change in relative prices, technology, or government policy will upset these forecasts. These methods are most useful, therefore, for short-run forecasting when changes in the underlying structure are expected to be minor.

Extrapolative techniques provide a very convenient and easy-to-interpret summary of the history of market penetration. However, the models do not lend themselves to assessing the impacts of future events such as new government regulations, supply disruptions, or cost changes. (Box-Jenkins techniques can be used to assess the impact of exogeneous shocks. However, it must be assumed that the reaction will be the same regardless of the source of the shock.) The simple form of the models does not capture the complexity of factors or casual relationships that actually determine substitution.

Validation of model output is very simple, at least for the historical period--i.e., the observed market shares can be plotted against their fitted trends. However, because these models are more descriptive than causal, there is little structure to validate.

For a given application, therefore, the ease and low cost of constructing and using these models must be balanced against their limited usefulness for forecasting and analysis.

CASE STUDIES

Case studies of material substitution entail an in-depth analysis of changes in material consumption in one or a number of specific end uses. For example, the investigation by Demler and Tilton found in Part 2 of this volume identifies the amount of tin used annually to produce beer and soft drink containers in the United States during 1950-1977. In addition, case studies examine the factors responsible for material substitution. They attempt to assess the relative importance of changes in material prices, technology, government regulations, consumer tastes and preferences, and the other factors responsible for substitution.

In some studies the analysis of the underlying factors responsible for changes in material consumption is preceded by an intermediate step that breaks down the changes in material use into various causal components. For instance, Demler and Tilton in their analysis of beer and soft drink containers (Part 2 of this report), assess the effects on tin consumption of changes over time in (1) beer and soft drink consumption; (2) the portion of consumption shipped in packaged containers, such as bottles and cans, as opposed to kegs and other bulk containers; (3) the portion of the packaged beverage shipped in tinplate cans; (4) the number of tinplate cans of average size needed per barrel of beverage; and (5) the weight of the tin in an average tinplate beverage can. This intermediate step helps determine how much of the change over time of material usage is due to material substitution. It is also useful in assessing the importance of material prices and the other factors that ultimately are responsible for changes in material usage.

Use of Case Studies

Case studies are a particularly useful tool for analyzing material substitution. Their depth and detail can provide insights into the nature and the determinants of material substitution that other methodologies miss. Case studies can also be used to test more general hypotheses or models of material substitution. They may produce unanticipated findings regarding the factors affecting substitution, the nature of the functional relationship between these factors and material usage, the time lags and responses in these relationships, and other important considerations. Such findings may then stimulate the formulation of new and better hypotheses and models.

In addition to expanding our general understanding of material substitution, case studies may also provide useful information for contingency planning. The National Materials Advisory Board, government agencies, and other organizations on a number of occasions have estimated the extent to which material substitution could alleviate critical material shortages in the event that imports of particular commodities were cut off during a national emergency. Careful case studies of the major end uses of such commodities can assist in such evaluations. They can provide information on the extent to which alternative materials might be employed, the cost in terms of poorer performance or higher prices, the time required to make such changes, and the implications for the available supply of the alternative materials. In similar fashion, case studies can facilitate efforts to appraise the effects of changes in government regulations, material prices, or other variables on material usage.

Limitations of Case Studies

Aside from such contingency forecasting, case studies are not an appropriate tool for forecasting. By their nature, they focus on the past. It is true that in extending our knowledge of material substitution, case studies enhance our ability to specify causal models for forecasting and analysis of material requirements one, five, or 20 years in the future. In the process of conducting case studies, we may also enhance our ability to develop better judgmental models for assessing future material needs. Still, the case study approach by itself involves explaining, rather than predicting, material substitution.

Several other limitations of the case study approach should be noted. The scope of most such studies is quite narrow, since they focus typically on material consumption in one or several end uses. As a result, their findings may be specific to the particular end uses examined, and these may or may not be of much importance. It is possible, of course, to determine the extent to which conclusions can be generalized by conducting a large number of case studies, but this is costly both in terms of the funds and man-hours required.

Even when only a few end uses are considered, an extensive effort is often required to carry out a case study. Considerable data, much of which may not be readily available, must be gathered and analyzed. Where information is lacking, techniques for obtaining reliable estimates must be devised. Interpreting the data and drawing valid conclusions usually involves some judgment, which should be checked by those with knowledge and experience in the end use sectors examined. This typically involves interviewing and travel. As a result, case studies are expensive in terms of the quality and the quantity of research time they require.

Complementary Value

In considering the relative advantages and disadvantages of the case study methodology, it is important to point out that this approach should be considered more of a complement than a substitute for other material substitution methodologies. As noted earlier, case studies provide a useful means of checking or testing other methodologies. By their narrow focus and depth, they also provide insights into the nature of material substitution that enhance the analytical capability and forecasting accuracy of other approaches.

OPTIMIZATION TECHNIQUES

The optimization of a system, such as a materials system or subsector, implies a selection of individual processes or material forms from several alternatives on the basis of the specific optimization criterion, usually minimum cost. In optimization analyses, the critical factors are the characterization of the competing options and the specification of the objective function to be optimized. Optimization is a normative process, indicating the preferred decision or direction--what should be done--given the assumptions made in the analysis (behavior patterns, the form of the constraints, and the objective function, etc.). In some cases, particularly where uncertainties are not dominant and objectives are relatively simple, the optimization process also has some simulation capabilities.

The major optimization techniques that are employed in resource allocation, process selection, and substitution analysis are mathematical programming (linear and nonlinear), dynamic programming, and optimal-control theory. Of these, linear programming has been used extensively in view of its ability to include extensive technical detail on the system being analyzed. A major strength of these optimization techniques is the facility with which an analyst can include both technical and economic factors in his analysis. The descriptive equations of an optimization model are usually developed in physical terms--e.g., energy and material balances, equipment utilization, and manpower requirements--while the objective functions generally consider such economic considerations as cost minimization or profit maximization. Solutions to resource-allocation or process-selection problems formulated along these lines reveal the marginal value, or cost, of limited resources.

The role of optimization in the analysis of substitution possibilities among materials depends primarily on the point in the materials system at which the substitution may occur. The possibilities for insights through formal optimization are usually greatest when the competing resources or technologies produce almost identical products or services. Under these circumstances, the market shares among competitors can readily be analyzed as a relatively simple function of the cost of the use of the alternative material resources, the cost of conversion technologies, the efficiency of material resource use, the cost of ancillary inputs such as energy and labor, and the environmental or regulatory contraints. The variables can be handled in most process optimization models using a single profit- or welfare-maximizing function.

Strengths and Weaknesses

In the analysis of substitution within the materials system, optimization techniques have proven to be strongest in the analysis of resource inputs and process selection. For example, optimization techniques have been applied successfully to the steel, aluminum, paper, and other materials sectors to analyze the effect of changing costs of resources, energy, and labor on the plant or process involved in the production of these materials. The success of these applications reflects the flexibility of this technique as well as the practicality of its application. New information or expert opinion can be incorporated in these models without costly revisions, and the results can be compared readily against reality.

On the other hand, if the products from alternative resources or processes differ in quality, durability, esthetics, etc., the resulting multiattribute problem in optimization is more difficult to model. Techniques exist for multiobjective optimization, but they are difficult to apply in practice and the results show considerable sensitivity to uncertain technical parameters and judgmental weightings. In situations such as these, optimal solutions may be tested for "robustness" against the major uncertainties--e.g., how sensitive is the normative solution to changes in the state of the world or other relative factors. Under these circumstances, especial consideration must be given to those optimization results that display "flip-flop" behavior--i.e., a tendency to switch between extreme solutions following small changes in model parameters.

In practical applications, particularly where large uncertainties exist or objectives are complex, there has been considerable criticism of the oversimplification necessary to apply optimization techniques. Some of this criticism is deserved; in particular, the problem of data availability plagues this form of analysis, whose major advantage is its ability to include complex technical and economic interrelationships. However, alternative techniques do not handle such problems well either.

Use of Optimization

Solutions determined by optimization techniques can provide useful insights into the feasible limits of, say, cost minimization or profit maximization under idealized circumstances. While such solutions may not be practical, they do not provide a well-characterized benchmark against which practical comprises may be gauged. In this sense, normative optimization is similar to the idealized Carnot-cycle limits in thermodynamics. While the presence of irreversibility induced by friction and energy exchange complicate the design and analysis of real processes, it has still proven quite useful to consider the ideal Carnot-cycle limits as a benchmark for comparison.

This is not to imply that the current optimization techniques are the last word in the analysis of substitution. As the above paragraphs indicate, a great deal of research must be done before a satisfactory substitution modeling technique will be available. Among the problems are the incorporation of the dynamic nature of substitution into an optimization model and the coupling of dynamic and linear programming models. In addition, there are a variety of specific applications and materials sets for which substitution optimization needs to be done.

ECONOMETRICS

Empirical econometrics analysis is a way of describing the behavior of a particular market (or the economy as a whole) with a relatively small number of simple equations. Once these equations have been established, they can be easily used to predict the future values of relevant variables or to study the effects of changes in the values of variables such as taxes, import quotas, etc.

The strengths and weaknesses of econometric analysis are inherent in the methods used to establish the equations which constitute the econometric model. Hence it is useful to outline these methods; the substitution of aluminum for copper in the electrical conductor market will be used below to furnish an intuitive feel for the strengths and weaknesses of the methods.

Econometric analysis can be viewed as comprising three conceptually different phases: application and adaptation of economic theory; data collection; and application of econometric theory.

In the first phase, the analyst applies economic theory to the particular market under consideration. For example, copper and aluminum cables are intermediate goods used as inputs in the production of other goods. Economic theory links the demand for intermediate goods to the production function (or to the cost function) of the end-users. However, economic theory also has to be extended to take into account industry-specific features. In the copper-aluminum case, one such feature is the U.S. producer price system for copper in the 1960's and 1970's.

The outcome of the first phase is the specification of the equations which will be used to describe the market. In the copper-aluminum example, the analyst would develop one equation describing the aggregate demand for electrical conductors and a set of equations for the share of each conductor--copper, aluminum, or a possible third substitute--in the aggregate demand. Each equation would have some coefficients or parameters which would be estimated from historical data. The equations would usually have simple functional forms, the most commonly used forms being linear and log-linear.

The ability to specify the descriptive equations on the basis of rigorous economic theory is one of the major strengths of econometric analysis. When this is done with care, it is usually possible to interpret the equations in a simple, intuitive manner. This makes it possible for even lay users to appraise critically the spirit (if not the details) of the analysis.

Conversely, because of this reliance on economic theory, it is difficult to incorporate qualitative effects into simple quantitative equations. However, when the qualitative effects are strong and pervasive, the analyst can usually take them into account through some well-established--though ad hoc--procedures, such as "dummy variables."

The second phase of econometric analysis is obtaining historical data for the variables included in the equations specified in the first phase. While this is conceptually easy, there are many practical difficulties. For example, no information is available for the actual transactions price of aluminum prior to the early 1970's. Consequently, econometric models of this market have previously relied on some proxy variable such as the price of aluminum in Frankfurt. Further, the agencies that report the primary data sometimes change their methods of gathering and aggregating field data, so that the reported data do not constitute a consistent measure. Finally, the report data may simply be inaccurate; this problem is particularly serious when studying international markets, since the published data for some countries often reflect guesses and estimates rather than actual statistics.

The outcome of the second stage is a set of numerical, historical data on the relevant variables. This is another major advantage of econometric analysis, since the basis of the analyst's derivations is made clear and explicit. Further, other analysts can use the same data to test their notions of market behavior.

The reliance on historical data, however, also introduces a major limitation in econometric analysis. This limitation is that the equations based on historical data cannot incorporate any knowledge the analyst may have about likely changes in the future. For example, equations that explain the shares of aluminum and copper in the electrical conductor market (on the basis of the relative prices of aluminum and copper) would be useless if a third conductor--such as optical fibers for communications transmissions--were to become commercially feasible. Similarly, major political and social changes may also invalidate the equations. In short, if the structure of the market changes significantly, equations based on historical data will seriously misrepresent future market behavior.

The third phase is the application of econometric theory. The task here is to choose the appropriate mathematical and statistical methods to calculate the parameters (coefficients) of the equations, using the data collected in the second phase. In this phase, the analyst also takes into account any problems with the data. While many problems cannot be eliminated, it is sometimes possible to minimize their impact with the choice of an appropriate method of calculation.

The outcome of the third phase is a set of estimates of the parameters, together with a number of summary statistics which indicate--to other analysts, at least, if not to lay users--the level of reliability or the precision of these estimates. For example, the analyst may report an estimated elasticity of substitution between copper and aluminum of, say, 0.8, with a standard error of 0.05. This would mean that a 10 percent change in the copper-aluminum price ratio would be expected to lead to an 8 percent change in the copper-aluminum quantity ratio. Further, this expected change would almost surely be within the range of 7 to 9 percent.

The use of rigorously established, statistically sound methods to estimate the parameters and their precision levels is one of the attractive features of econometric analysis. When this feature is linked to equations based on economic theory, the result is that econometric analysis becomes a powerful tool with wide and general applicability. However, the technique can also be misused by analysts who, for one reason or another, are not careful enough to maintain the level of rigor required in each of the three phases described above. Further, even the most careful analyst cannot overcome the limitations inherent in relying on historical data.

In particular, great care must be taken if an econometric model must be used to make forecasts outside the historical range of the data collected in phase two. Furthermore, dramatic variations between observed and predicted market behavior could be an indication of a major structural change in the market, which could invalidate the model developed. This problem is of special concern when modeling materials substitution, since the object of the analysis is to predict structural shifts. In such cases, incorporating engineering and other information directly into the econometric analysis can greatly improve the usefulness of econometrics.

The major conclusion is that econometrics, when properly applied, can be a very useful tool in studying past materials substitutions, in forecasting the future if it will sufficiently resemble the past, and in incorporating certain kinds of engineering information into a consistent econometric analysis.

INPUT-OUTPUT

In any systematic approach to the question of materials substitution, two considerations occur. One, can a substitution be made? That is, within the context of technical feasibility and economic efficiency, which materials are admissable candidates as substitutes for each other, and under what conditions should one or more of them be chosen to the exclusion of others? Two, assuming that a substitution of one material for another has been implemented, what are the implications of that substitution for the system at large, within the context of institutional, economic, social, environmental, international, and physical constraints. Because of the complex interrelationships that characterize modern developed economies, it is impossible to know the direct and indirect consequences of a single change in the economy--i.e., the substitution of one material for another--without a firm understanding of the underlying structural relationships that constitute a modern economy.

The input-output technique is not designed to address considerations of the first kind: questions that center on what techniques of production should be used, when the switch to a new technique should be executed, or what materials are likely candidates for substitution. On the other hand, the input-output technique is in an advantageous position to consider questions of the second kind: what are the direct and indirect implications of a materials substitution for the system as a whole? In addition, since input-output tables can be very finely disaggregated, material by material, by observing both flows and coefficients, the method sheds some light on where to center research for possible future materials substitution by identifying the sectors and materials in which substitution from a technical point of view may be feasible.

SIMULATION

A simulation model is not always entirely distinguishable from the other model methodologies considered. To distinguish the differences from other modeling approaches--and the differences in practice are in degree, not in kind--it is useful to contrast the simulation approach to the other causal modeling methodologies considered, i.e., econometrics and optimization. An econometric model can be defined as one in which first, a relation between economic variables is hypothesized, and, secondly, statistical procedures are applied to determine whether or not the hypothesized interdependence can be accepted. The relations specified are sometimes dynamic and sometimes not, and econometric models may comprise a large number of equations. A process optimization model, on the other hand, is one in which the optimum mix of alternative substitutable inputs is determined to produce a well characterized output. Optimization models generally ascribe and attribute rational economic decision-making to the decision-maker being modeled, and can be either statically or dynamically specified. In its simplest form, a simulation model is nothing more than a hypothesis of interrelationships among variables, typically dynamically specified, in which the time behavior of the interrelations are revealed via simulation--the process of integrating the differential equations. (The term simulation is also applied to the process of solving both econometric and optimization models. Thus, the solution procedure is not the characteristic that distinquishes the alternative approaches.) A simulation model can be distinguished most clearly from the other causal modeling approaches as one not requiring hypothesis testing in the statistical sense nor attribution of rational economic decision-making to the actors being modeled.

Typically a simulation model will interconnect submodels that capture generally accepted behavior in the small in an attempt to simulate overall performance in the large. As practiced in some forms, the model will often articulate global constraints, generally accepted to exist, but quantified only approximately, usually because of limitations of knowledge, that have important implications for the time behavior of the model variables. For example, constraints that trigger turning points, or changes from growth to decline or vice versa, or changes from growth to stable steady states are favorite targets of analysis.

Probably the greatest strength of the simulation methodology is its flexibility. One is not constrained by data or economic rationality in constructing relationships of the model. Such flexibility is sometimes needed.

Often, important questions are questions simply because data are unavailable. The simulation approach allows the application of scientific thought to the problem at hand, nonetheless. Such an approach might take the form of analyzing the sensitivity of the system's behavior to an important missing element of data to determine the importance of the missing data. For this reason, the simulation approach is popular in a research environment. However, the model's results are always contingent upon accepting the "BIG" hypothesis that the model specification is correct--usually a matter of faith, since, because of limitations of data, the more scientific econometric approach cannot be applied.

As far as the ability of the simulation approach to deal with other than economic rationality is concerned, it is useful to observe that political processes often evolve solutions to economic problems that are politically acceptable but not economically rational. To the extent that the simulation approach to modeling can reflect the political realities, it might be preferred to the econometric or optimization approach as a descriptor of reality.

Especially when actions of government are likely to influence model variables, and government itself is internal to the structure of the model, simulation may be the only fruitful modeling approach. The limitation is that political science itself is a very inexact science and does not lend itself well to the precise formulation of interactions required in computer models.

It must also be said that the simulation modeling methodology does not preclude part of its structure's being derived through statistical techniques nor parts of the model's displaying economic rationality. A simulation model most artfully constructed may be comprised of components econometrically derived, especially for aggregates such as large numbers of consumers, and other parts based upon optimization or process models of particular industries, firms within an industry, or alternative production processes within an industry, all embedded in a simulating device where still other components of the model are based on neither econometric or optimization approaches. In such a model it is obviously inadvisable to distinguish the three alternative approaches to causal modeling as separate and mutually exclusive. They are not.

JUDGMENTAL

It has been known for a very long time that individual problem-solving ability varies and that individual planning horizons vary. The expert problem solver appears to be mentally guided by large numbers of patterns serving as an index to relevent parts of the knowledge store. Nobel laureate Herbert Simon and his associates suggest that the patterns are "rich schemata that can guide a problem's interpretation and solution and add crucial pieces of information. This capacity to use pattern-indexed schemata is probably a large part of what we call physical intuition." Larkin (1980) Linstone (1975) and Loye (1978) discuss the differential discounting which affects thinking about the future as well as the past.

Such analyses confirm that individuals may indeed be a source of valuable insights about the future. Much is made in the popular literature of predictions made by experts that have proven wrong; such cases appear to be more entertaining than correct predictions (a case of Schadenfreude). Nevertheless, forecasting by judgmental means remains in wide use in industry. A recent survey of more than 100 firms on the usefulness of 10 technological forecasting techniques (Balachandra 1980) placed expert opinion at the top. Fourteen of the 18 industries covered claim expert opinion as the most useful technique. The industries include construction, paper, chemicals, refineries, primary metals, fabricated metals, and machinery. Brainstorming also ranked high, much higher than, say, simulation (Balachandra 1980).

There are other considerations. The methods favored by scientists and engineers rely on data and models (or combinations thereof). This dependence is based on the great success in the use of the science-based paradigms in dealing with the well-structured problems that comprise purely technological research and development. However, sociotechnical problems are ill-structured, and the same paradigms and modes of inquiry prove quite inadequate. Assessment of the impacts of technological substitution in particular is illustrative of such ill-structured problems. Multiple perspectives are needed that sweep in other paradigms and inquiry systems (Linstone, Part 2 this report). Judgmental methods provide the opportunity to employ such perspectives.

Finally, multiple perspectives can fill another function--communication of output. Not surprisingly, they can significantly facilitate effective communication with parties whose involvement is vital but who are not scientists or engineers.

Intuitive Approaches

Expert opinion can be solicited by personal dialogue or interview, by group sessions or committees, and by questionnaires or remote conferencing. Lendaris (1979) lists three classes of techniques that can assist in drawing out ideas:

(a) Procedures that provide an atmosphere for freewheeling thinking, brainstorming, nominal group technique, Delphi, computerized conferencing.

(b) Structured semantic guidance, scenarios, functional analysis, attributed listing.

(c) Structured geometric guidance trees, digraphs or cognitive maps, networks, pattern recognition methods.

A representative example of each of these tools is discussed briefly below.

Type (a) Delphi

Delphi involves a group (in this application a panel of experts) in a remote, structured, conference procedure using iterative questionaires while maintaining individual anonymity. The use of Delphi in technology substitution is illustrated in Linstone (paper presented in Part 2); more details may be found in Linstone and Tureft (1975).

Advantages. Ease of use--no long training or technical background required to do Delphi.

Low cost--participants can remain at their own working places or homes.

Ability to bring in many views--experts can be drawn from distant places and in sizable numbers.

Multiple perspectives--facilitates introduction of several perspectives.

Flexibility--can be easily modified for specific needs (e.g., first a panel of research experts, then a panel of development or production experts).

No forcing of consensus--iteration of rounds proceeds until stability is attained whether consensus or not.

Disadvantages. Difficulty in experts' handling of environmental changes.

Individuals have problems in simultaneously projecting many changes. They can handle only a few and assume everything else remains unchanged (ceteris paribus). Thus multiple substitutions and future changes in values or criteria of utility tend to be ignored.

Possibility of manipulation--as with most tools, manipulation is a residual danger. The selection of panelists, the feedback data, and the wording of statements can all be altered.

Ease of misuse--ease of use also implies ease of misuse (e.g., sloppy execution).

Type (b) Scenarios

Scenarios are more useful in forecasting needs than capabilities. They represent an effort to provide a more holistic basis for forecasting using both trends and possible events. Scenarios are, in effect, future histories.

Advantages. Holistic approach facilitates inclusion of interactions--there is less likelihood of the ceteris paribus fallacy found with many techniques.

Ease of communication--scenarios are easily understood, in contrast with sophisticated modeling tools.

Disavantages. Low predictive power--the many variables included lead to low probability of occurrence of any one scenario. (Consider 20 events occurring or not occurring with a probability of 0.9 each. The likelihood of a scenario involving all is only 0.12.)

Writing problems--technologists are not usually adept at developing useful scenarios; their background militates against the use of multiple perspectives which add vital dimension to any scenario.

Low value for capability forecasts--scenarios are not well suited to forecasts of technological capability.

Type (c) Digraphs or Cognitive Maps

Digraphs are typical of structural modeling tools and are particularly appropriate for technology assessment (Linstone, et al 1979). For example, the tracing of consequences of a materials substitution can be facilitated by a directed graph or cognitive map. Such maps can augment intuition by clarifying the effects of a large number of interactions. Structural modeling inherently emphasizes structure rather than data, the geometric rather than the arithmetic

Digraphs can illuminate deviation-amplifying and deviation-counteracting (or dampening) loops of impacts as well as indicate in a semiquantitative way the dynamic behavior of key variables (by means of simple computer programs).

Advantages. Ease of use--simplicity of digraphs permits use with minimal training as well as ability to alter model readily.

Ability to consider many interactions--intuition can be effectively augmented by these tools.

Low-cost.

Both cumulative and proportional (or long-term and short-term) connections.

Disadvantages. Subtle assumptions--usually limited to pairwise interactions, transitivity.

Structural changes must be introduced externally.

Easy misuse--misinterpretation of semiquantitative as fully quantitative output in computer calculations.

We conclude that judgmental models have an important role in technological forecasting and assessment in general, and in technological substitution in particular. Their simplicity can be misleading, however, if skill and care are not exercised in their application.

ENGINEERING

The last of the causal methods considered by the committee was engineering analyses. Examples are described by Gordon and Noton in their presentations in Part 2 of this report. The Gordon method involves use of a relatively simple and straight-forward analytical technique already reduced to computer format; the Noton procedure involves manual use of a great deal of engineering and cost data on a variety of materials, forms, and fabrication techniques to evaluate substitution options in the aerospace industry. Noton's procedures can be greatly expanded for use in many other industries and could also be reduced to computer format for automatic selection of least-cost substitutes. Although somewhat less complex but with greater breadth than Noton's model, the Reference Materials System (RMS) described by Hoffman in his workshop presentation in Part 2 shows, in principal, how this might be done as related to substitution.

Engineering analysis methods differ from the other methods discussed in that they permit evaluation of physical criteria, production processes and methods, and engineering parameters as well as economic factors. Thus they can be used as subroutines in econometric and input/output models, to justify and validate costs and other economic factors commonly used in such purely economic models. In addition they can be used in conjunction with optimization and system simulation models to specify constraints, parameters and other relationships in the models. Most scientific and engineering relationships have well-founded, well-developed mathematical concepts and are often more readily treated analytically than are the softer issues in computer models. Like all modeling techniques, however, engineering analysis has difficulty in addressing many judgmental issues, such as the resolution of social and political problems.

In the simplest, least-cost substitution model, only the costs of the original and substitute material are considered. Gordon's model, however, goes on to evalute relative material quantities, additional (or decreased) capital and operating costs required for producing the substitute material or product, and additional (or decreased) life-cycle costs associated with use of the substitute. Gordon then illustrates use of the model in substitution of coke-produced for charcoal-produced pig iron in the production of bar iron (steel) in 17th century England and substitution of aluminum for copper in automobile radiators and in-house wiring. Gordon's paper shows that consideration of factors other than just materials costs can lead to very different conclusions about the value of the substitute.

Gordon's method suggests that factors other than materials cost are evaluated manually in somewhat the same way as in Noton's handbook procedure. However, this approach is not inherent in the method, since each factor can be reduced to a computer subroutine to find the appropriate cost.

Strengths and Weaknesses

The strengths of a Gordon-type model are: first, the model is quantified and gives numerical overall cost comparisons between materials; second, it is quite comprehensive and gives life cycle rather than just initial production costs for a substitute product; and third, it can readily be expanded to include multiple comparisons and coupled effects.

The Gordon-type model also has inherent weaknesses. It assumes technical feasibility of substitutions; in reality, these must often be proven through experience, particularly in total life-cycle evaluations. Secondly, the model fails to evaluate political, social, and environmental as well as purely economic factors, but this is a shortcoming in all other types of models as well. Further, it does not address substitute processes or functional substitutions but only material-material alternatives; inclusion of substitute processes in the model would be feasible, but inclusion of functional substitutes would be more difficult. Finally, the model does not permit easy inclusion of quantum changes, such as substitution of solar for conventional energy.

With regard to input considerations for engineering analysis models in general, two important factors are model development and operating costs, and the availability of data. The costs of developing a simple model such as Gordon's should be moderate, but costs increase rapidly as the number of subroutines for evaluating the various cost factors is enlarged and as the number of process steps is increased. The costs of developing models such as Hoffman's RMS model and of computerizing Noton's handbook procedure would be high because of their complexities. However, once developed, all of these models would be relatively inexpensive to use. With regard to data availability, basic chemical, physical, and engineering data are widely available. On the other hand, much of the production data are unavailable. For example, about 20 percent of U.S. Bureau of Mines national production data are withheld for proprietary reasons, and in many cases U.S. Department of Commerce Standard Industrial Classification information is insufficiently disaggregated to be totally useful. In addition, evaluation of newer, alternative processes will generally be difficult because the details of most new processes are tightly guarded by industry.

Current engineering models have forecasting value mainly for the short term (1 to 3 years) and to a limited extent for the intermediate term (next 3 to 5 years). This characteristic stems largely from lack of ability to predict quantum jumps in technology as well as limited availability of data on new processes under development. Engineering-analysis models are, or can be made, very flexible and have an added value in that they can evaluate and judge many technological factors that purely economic models cannot. For similar reasons they are probably the most pragmatic of models. Finally, the use of these models can be valuable in identifying technical and other areas where additional input must be developed to increase assurance of the model's validity. For example, it was pointed out that life-cycle substitute costs may be highly suspect in many cases. This shortcoming can be remedied by conducting accelerated tests and field trials on substitutes.

SUMMARY OF WORKSHOP RESULTS

The findings of the committee and the invited participants in the workshop at Woods Hole are summarized in Tables 1, 2, and 3 (pp. 4-6) and repeated here (pp. 28-30).

Table 1 compares the substitution methodologies examined on the basis of the requirements their implementation entails. These include financial requirements, data requirements, and manpower requirements.

Table 2 compares the applicability of these techniques to the spectrum of materials substitution problems. These include forecasting, impact analysis, and retrospective analysis.

Table 3 outlines the selection criteria to be applied during the selection of an appropriate methodology. Included are validation, flexibility, and transparency.

In each category, the methodologies are ranked in Tables 1, 2, and 3 in terms of low, medium, and high. The bases of the rankings are discussed below.

Requirements of the Methodologies

Table 1 summarizes the committee's assessment of the requirements of the eight methodologies discussed. These requirements may be broken down into four major divisions: data requirements, costs, support staff (personnel) requirements, and development time requirements.

Data Requirements

Extrapolative techniques were found to have the lowest data requirements. This result is not surprising in light of the modest objectives of such techniques.

Simulation and judgmental techniques were given medium rankings in this area. The emphasis of simulation techniques on measurable flows and transactions rather than the structural underpinnings of the system to be modeled explains their medium ranking. The assignment of the same ranking to judgmental models

Table 1. Comparison of Requirements of the Techniques

	Extrapolation	Case Studies	Optimization	Econometrics	Input/ Output	Simulation	Judgmental	Engineering
Data Requirements	Low	High	High	High	High	Medium	Medium	High
Costs - Building	Low	Medium	Medium	Medium	Medium	High	Medium	Medium
Costs - Maintenance	Low	Low	Medium	Medium	High	Medium	Medium	Medium
Costs - Using and Documentation	Low	Low	High	Medium	High	High	Medium	High
Costs of Calculation	Low	NA	Medium	Medium	High	Medium	Low	Medium
Personnel	Low	Medium	Medium	Medium	Medium	Medium	High	Medium
Development Time	Low	High	High	Medium	High	High	Low	High

Requirements: Low means that the requirements are simple.
High indicates that the requirements are complex or detailed.
NA: Not Applicable

Table 2. Comparison of Potential Uses of the Technique

	Extrapolation	Case Studies	Optimization	Econometrics	Input/Output	Simulation	Judgmental	Engineering
Forecasting Short Range	Medium	NA	Low	Medium	NA	Low	High	Medium
Forecasting Medium Range	Low	NA	Medium	Low	NA	Medium	Low	Low
Forecasting Long Range	Low	NA	Medium	Low	NA	Medium	Low	Low
Impact Analysis Technology	NA	Medium	High	Low	NA	High	Low	High
Impact Analysis Economic	NA	Medium	High	Medium	NA	High	Low	Medium
Impact Analysis Policy/Regulatory	NA	Medium	Medium	Medium	NA	High	Low	Low
Materials Selection	NA	Low	Medium	Low	NA	Low	Medium	High
Historical Analysis	NA	High	Medium	High	Low	Medium	Medium	Medium
Implications of Substitution	NA	Medium	Medium	Low	Medium	Medium	Medium	NA

Potential Uses: Low means more difficult to use or less effective
High means easier to use or more effective
NA means not applicable

Table [illegible]. Comparison of Criteria for Selection

	Extrapolation	Case Studies	Optimization	Econometrics	Input/ Output	Simulation	Judgmental	Engineering
Application Desired:								
Forecasting	Medium	NA	Low	Medium	NA	Low	High	Low
Analysis	NA	Medium	High	Medium	NA	High	Low	High
Budget and Time Constraints	Low	High	Medium	Medium	Medium	High	Medium	High
Sensitivity to Changes in:								
Technology	Low	Low	High	Medium	Medium	Medium	Medium	Medium
Economic Factors	Low	Low	Low	High	Medium	Medium	Medium	Medium
Transparency:								
To Naive User	High	High	Medium	Low	High	Low	High	High
To Expert User	High	High	High	Medium	Medium	Medium	Low	Medium
Ability to Validate:								
Internal to Method	High	Medium	Low	High	Low	Low	Medium	Low
External to Method	Medium	Low	Medium	High	Medium	Medium	Low	Low
Level of Aggregation	High	Low	Medium to High	High	Medium to High	Medium to High	Low to High	Low

Selection Criteria: For application desired, a high rating is favorable. For those characteristics that are desirable, i.e., transparency, ability to validate, and level aggregation, a high rating is good. For those characteristics that are undesirable, i.e., budget and time constraints, sensitivity to changes in technology or economic factors, a high rating is bad.

<u>NA</u> means not applicable

requires a bit more explanation. In establishing a judgmental model, the model makers are actually establishing a framework within which the experts selected can interact productively. The amount of data required to establish such a framework is, of course, quite low, although the model makers must collect sufficient data to assure that the framework they construct will be effective. However, the search for the requisite experts and the maintenance of their expertise entail significant data requirements.

The remaining five techniques were given high ranking in the data-requirements category. Engineering, input/output, and optimization techniques require detailed technical data, while econometric techniques require detailed statistical data. Case studies require all of the above and more.

Costs

The costs of implementing these various techniques were broken down into two main categories: the costs of creating a model and the subsequent cost of calculation. The creation costs were further broken down into three categories: the costs of building, maintaining, and using the model.

Costs of Creating a Model

Building the Model. Extrapolative techniques were given a low ranking in the cost of building a model. Relatively inexpensive software is readily available to perform any but the most complex extrapolative calculations, and most of these programs are written for the inexpert user.

At the other end of the spectrum, simulation models were given a high ranking. This reflects the normative nature of this technique, as well as the specialized nature of the model formulation. There is no such thing as a standard simulation of any materials sector, and every practitioner has his own preferred programming language and methodology.

Between these two extremes lie the other six techniques. The structure of each of these techniques is relatively consistent across all applications and the mathematical techniques are well established with the practitioners.

Maintaining the model. The costs of maintaining a model are related directly to the structure of the model developed. The more difficult the incorporation of new data, the more costly the model maintenance. Based on this criterion, input/output techniques were given the sole high ranking. The incorporation of every new piece of data requires a reformulation of the transformation matrix and can be very tedious and time-consuming.

Extrapolative and case-oriented studies were given a low ranking because of the ease with which new data can be incorporated. New and diverse data, of course, are the essence of case studies, so the low ranking is not surprising. Extrapolative techniques can incorporate new data by simply adding it to the data base the computer operates on.

The remaining five techniques were given a medium ranking. In each case, the incorporation of new data requires an addition to or a minor revision of the model. With engineering and optimization models, one would expect new data to be incorporated in the form of new or modified constraints, technical or economic, which can be readily included. Similarly, with econometric studies, new data imply either the inclusion of a new regression varible or the revision of a block of regression data. In some cases, simulation models may require only minor changes when it is desired to incorporate new data. The incorporation of new data may however, require extensive revisions of the model structure.

Using the Model. This category examined the costs a naive user would face in learning to use a model's results. The assessment included the difficulties associated with interpreting the model's results.

In the low category were placed extrapolative and case-oriented models. The low ranking of these models is readily explainable since their results and reasoning generally are clearly understandable and interpretable by the layman.

In the high category were placed optimization, engineering, input/output, and simulation techniques, which require a degree of technical or economic specialization on the part of the user. This requirement does not actually reflect the capabilities needed by the user, but users who are unaware of the technical or economic complexity of the techniques run a serious risk of misunderstanding or misapplying the results.

The remaining techniques, econometrics and judgmental, received a medium rank in this category. In the case of econometric models, this ranking reflects the ease with which the results may be presented, as well as the fact that there are subtleties in the use of such models which require some care. This is also the case with judgmental models, although the derivation of the results may be somewhat less apparent to the layman. It is important to note that this facile appearance may also be used to conceal biases and agenda.

Costs of Calculation

This category reflects the expected cost of obtaining results with a model and the cost of an analyst's time for interpreting these results. Extrapolative models were given a low ranking in this category, once again reflecting the moderate objectives and results of the technique. Judgmental models also received a low ranking. Although the costs of using experts may be high, the technique does not require full-time participation by them. Input/output models were also given a high ranking. This ranking reflects the extensive computing time required and the interpretive skills necessary. In particular, the inversion of the large transformation matrices entailed in these models is a considerable effort requiring a significant amount of computing time and ability. Given a medium ranking were econometric, simulation, optimization, and engineering models. Each of these models requires a degree of computation on the order of that required by the techniques in the high category, but far less work is needed to interpret the results.

Personnel Requirements

This category outlines the relative degree of technical competence required of the personnel involved in the construction, maintenance, and use of the various models.

In the high category are judgmental techniques. These techniques rely on avowed experts in the field(s) under study, who make judgments and determinations under a variety of conditions. It is self-evident that such techniques would make major demands on the competence of the participants.

In the low category are extrapolative techniques. Once again, this ranking is consistent with the view that an extrapolative technique is a good start, but hardly a rigorous treatment of the subject.

The other six techniques were placed in the medium category. This ranking perhaps understates the personnel requirements of these techniques, which are not by any means amenable to unsophisticated treatment. However, on a relative scale, the techniques are certainly less demanding than the judgmental method. In general, they require a user with significant insights into the technical and economic aspects of the system under study, as well as an appreciation of the vagaries and assumptions inherent in the techniques applied. Furthermore, an ability to validate the model's results is essential for any user.

Development-Time Requirements

This category compares the times required to develop the various models examined.

Extrapolative techniques, were placed in the low category for fairly obvious reasons. The data requirements are not substantial, and a number of mathematical techniques for treating the data are readily available. Also placed in the low category were judgmental models. While a certain degree of preparation is required, the main time requirement is the time needed to bring the experts together and give them an opportunity to interact, which is much less time than that required by the other techniques.

Case studies and engineering techniques were given high rankings. Case studies require significant blocks of time for collecting data, but much less for analyzing the data. The high ranking for engineering techniques is a result of similar considerations--the difficulty of collecting many kinds of engineering data, especially cost-related and proprietary technical data.

Optimization, input/output, and simulation techniques also received high rankings. Although these models employ fairly standardized techniques, the actual development of a model requires extensive examination of previous and current modeling efforts and other studies in order to formulate and refine the relationship on which the models will be based. Because of the reliance on previous work, these techniques also require significant development time.

Econometric techniques were given a medium ranking. This ranking reflects the relative standardization of the structures and the methods of these techniques. While the data requirements may be as great as those of the preceding techniques, far less time is required to develop the structure and application of these techniques.

Potential Uses of the Models

Table 2 depicts the committee's estimate of the appropriate uses for the modeling methods examined. As in Table 1, the techniques were ranked as low, medium, and high. The committee broke the potential applications of these modeling techniques into five categories:

1. Forecasting
2. Impact analysis
3. Material selection
4. Historical analysis
5. Implications for substitution

Forecasting

Forecasting was broken into three natural categories: short range (1 to 3 years), medium range (4 to 10 years), and long range (greater than 10 years). Case-oriented and input/output techniques were not included in this ranking because they are clearly inapplicable to forecasting.

Short Range. Judgmental models were given a high ranking in short-range forecasting, reflecting the committee's judgment that experts in a field are more likely to be able to balance the multitude of short-range factors that contribute to conditions in the near term. At the other end of the spectrum were optimization and simulation techniques. These methods were given low rankings based on their inability to reflect satisfactorily the dynamics of abrupt changes in the systems being modeled. Extrapolative and econometric models were given medium rankings because of the strong influence of past trends and incremental change in the results of these models. Engineering models were also given medium ranking in light of the fact that they rely on extant technology, which generally will not be subject to abrupt changes, regardless of other dynamics.

Medium Range. In medium-range forecasting, only medium or low rankings were given. Furthermore, most of the techniques lost favor as the time range expanded. Engineering, extrapolative, and econometric models dropped from medium to low rankings and judgmental models dropped from a high to a low ranking. In each case, the drop in ranking results from the very factors that made these techniques advisable in the short range. The reliance of the techniques upon past data and incremental change grows increasingly detrimental as the time horizon expands. Furthermore, techniques that are more likely to incorporate a strong present-oriented bias, such as judgmental techniques, tend to lose ground more rapidly. This is in direct contrast to techniques that are able to treat abrupt changes and dynamic situations more satisfactorily, such as optimization and simulation, which were given medium ranking.

Long Range. In long-range forecasting, much the same considerations hold as were outlined in the medium-range category, and the rankings reflect this similarity. It is important to remember that these rankings are only relative and that, in general, the longer the time horizon, the less certain one can be about the results of any predictive model. So, even though optimization models are given the same ranking in the long range as in the medium range, one would tend to give less credence to the results of a longrange study than to the results of a medium-range one.

Impact Analysis

Three types of impact analysis were specifically considered in this ranking of model applicability. These are:

a. Technology impacts
b. Economic impacts
c. Policy/regulatory impacts

Once again, certain techniques--extrapolative and input/output--were excluded from consideration. The exclusion of extrapolative techniques reflects the inability to treat such problems, while the exclusion of input/output techniques reflects the sheer difficulty of converting the transactions matrix to reflect one policy change, much less several.

Technology Impacts. Optimization, simulation, and engineering techniques were given high rankings in technology impact. Models constructed with each of these techniques can be easily altered to reflect alternative technologies. Econometric models were given a low ranking because of the high level of aggregation that marks econometric statistics. Because of this aggregation, it is very difficult to distinguish the impacts of specific technological changes. Judgmental techniques also garnered a low ranking, based on the perception that, while such models are able to point out areas in which technological changes are desirable, the full impacts of such changes are rarely considered beyond the immediate short range. Case studies were given medium ranking. This ranking was given because in these models the primary analytical means of treating technological change is personal, expert estimation which, while very subjective, can bring many otherwise unquantifiable factors to bear on the problem. However, this technique is very structure-oriented, and the impacts of the change in structure that would accompany a change in technology would be difficult to assess.

Economic Impacts. Optimization and simulation techniques were given high ranking in economic impact, reflecting the ability of both to create, so to speak, their own economic environment within which the relative impacts of many factors can be analyzed. Furthermore, this economic environment may be altered readily within the basic framework of the model. Case, econometric, and engineering studies were given medium rankings. The case-study technique, being a post facto type of study, is better suited to the analysis of observed changes in economic conditions than to the analysis of expected behavior in response to predicted economic changes. Similarly, econometric studies, while able to reflect the underlying structure of observed economic interactions, are less reliable when examining activity under economic conditions outside the norm.

Engineering models tend to submerge economic parameters inside of discounting or capitalization equations, of which the sensitivity to shifts in economic conditions is rarely dependent upon more than a discounting rate. Judgmental models were given a low ranking in this category as well. Such models were considered to be likely to reflect a broader perspective of the impacts of economic changes than would be useful for impact analysis. Furthermore, a complete picture of these impacts generally is beyond the scope of judmental models.

Policy/Regulatory Impacts. Simulation techniques received a high ranking in policy/regulatory impacts. This ranking results from the relative ease with which changes in policy can be incorporated into simulation models, especially where impacts on materials flows may be readily quantified. Case and econometric studies were given medium rankings for much the same reasons as in the economic-impact category. Optimization techniques were ranked medium in this category, primarily because of the difficulty associated with isolating the impact of specific policy on this technique's results. Engineering models were given a low ranking in this category, reflecting a wide perception that engineering models result in solutions that are policy independent. Judgmental models, as in the previous impact areas, also were ranked low.

Usefulness for Materials Selection

Engineering models were ranked high in usefulness for materials selection. The ability of these models to compare a variety of materials options on the basis of their engineering utility and other scientific criteria is a clear advantage over the other techniques considered.

Case, simulation, and econometric studies fared less well, receiving low rankings from the committee. Both case studies and econometric models are more indicative of past performance than of future tendencies. Simulation earned its low ranking for much the same reasons. These models tend to reflect the behavior of factors within the context of the very structures that would be eliminated by a technological or engineering innovation.

Optimization was given a medium ranking. While this technique is limited in the same fashion as simulation models, sufficient flexibility can be introduced into the model to reflect the responses of optimal decision-makers to changes in the availability of alternative materials. Judgmental models were also given a medium ranking. While these models can reflect many perceptions and criteria, the effects of personal biases on the model's results cannot be discounted.

Historical Analysis

Case studies and econometric analyses were given high rankings in historical analysis. Since these techniques are based on historical analysis, one would expect them to perform best in this area. Input/output models were low rank because of their "present-orientedness." These models are designed to describe current material flows rather than to outline the path that led to the present situation. Optimization and engineering analyses were ranked medium because of their relative inability to trace the path of transition from one situation to another. These models, once a clearly better approach is made available, select the newer approach, and frequently do not reflect the lag times involved in such changes. Simulation and judgmental techniques were also ranked medium. Once again, these techniques are "present-oriented" and would not be expected to yield results as useful as those given by more backward looking techniques.

Implications of Substitution

In this category, the committee attempted to rank the available techniques on the basis of their ability to offer insight into the impacts that materials substitution would have on the system under study. Most techniques received tentative rankings of medium except for econometrics, which was ranked low in view of its insensitivity to major future changes in consumption. Of the remaining techniques, input/output was given serious consideration in light of its ability to reflect minutely the impacts of changes in materials-consumption patterns. On the other hand, optimization and engineering techniques were found to be more able to quantify the benefits and costs of such changes, while simulation techniques were expected to offer major insights into the effects of policy changes.

SELECTION CRITERIA

The criteria for selecting a way to forecast and analyze materials substitution for both the expert and the naive user are detailed in Table 3. The rankings (low, medium, and high) are meant to indicate the relative strengths and weaknesses of each technique, and the interpretation as to whether low means the technique is or is not less desirable depends on the specific criteria described below. These criteria comprise the questions that a potential user should be asking, for there are fundamental differences among the techniques and no single one is superior in all respects for all uses. The most fundamental question is what will the analysis and results be used for and who will be using them and for how long. Use of the selection criteria in Table 3 should permit the user to address these considerations.

The eight methods at hand are applied almost without exception to forecast trends and events or to understand and analyze past occurrences. The purpose of forecasting is obvious; however, the purpose of analyzing the past goes beyond simply an intellectual understanding. Once the process of materials substitution has been modeled, questions relating to hypothetical changes in various factors--such as new technologies, changes in prices, and so forth--can be posed and their probable implications traced as they affect the model. Some techniques are more useful in conducting these "what if" questionings than others. Clearly the success of any particular application depends on the quality of the specific model and the data employed to estimate it.

Model building is a way to abstract systematically from the complexities of the real world and to stress the critical or core relationships that are thought to describe that reality. Model-building techniques such as those described in this report are the most widely used and have proven successful in the past. However, each technique imposes both explicit and implicit assumptions and restrictions on the model builder in addition to those abstractions he controls directly when specifying a model. The criticality of the technique-related assumptions surfaces in the Application Desired and the Transparency criteria of Table 3.

Application Desired

Forecasting. In forecasting one makes assumptions about the future state of the world and attempts to trace the effects of the events or conditions assumed. Hence the farther into the future one is forecasting, the less meaningful past experience becomes. In forecasting with extrapolative techniques the analyst assumes that past trends will continue, and consequently the method is ranked medium for short-term forecast (less than five years as denoted in Table 3) and low for forecasts of greater length. Case studies and input/output analysis are unranked for forecasting in Table 3. The methods involved in case studies are primarily ad hoc and depend entirely on past behavior; therefore, they cannot be used in, forecasting. In input/output analysis, the economy is assumed to remain in long-term equilibrium and technology is rigidly defined, implying the technique is inappropriate for forecasting. Econometric estimates rely on past data, but allow the investigator to extract information about a wide variety of relationships, some or all of which may continue into the future. This ability earns econometric analysis a medium rank. The limitations of optimization, simulation, and the engineering methods include restrictive implicit assumptions and failure to consider the effects of key variables which may cause change in the future; these techniques earn a low rank. Judgmental techniques produce forecasts based on expert opinion, and their accuracy depends on the quality of the chosen experts or, equally important, the current state of thinking within the industry or discipline. Judgmental methods earn a high rank, for if specialists in the real world have formed opinions about the future that are not likely to be realized, then forecasting is a vacuum.

The contribution of analysis, in addition to understanding, is alternative outcomes produced by different "what ifs" and sensitivity about the results. Techniques that accommodate a wide variety of assumptions about technology and economic conditions and that allow for discontinuities from the past and normative judgments about the future are most productive in this respect. Optimization, simulation, and engineering approaches rank high because they produce this output. Case studies are ranked medium because each is unique. Econometric models are ranked medium because the estimated relationships among variables rely on the past and thus are less able to accommodate structural changes or values for variables outside of the range of past observations. Input/output methods can trace through implications of substitution, but in themselves are not suited to analysis. Judgmental models are ranked low unless the questions posed to the participants specifically include the "what if" questions. When this is the case, judgmental models would advance to medium, but not to high, because subsequent "what if" questions entail another replication of the method.

Budget and Time Constraints

Constraints on budget and the time available to produce results vary directly, and the rankings in Table 3 summarize those found in Table 1.

Sensitivity to Changes

Technology. The sensitivity of a technique to changes in technology and economic factors differs from analysis because these changes may be posed hypothetically, may be assumed to follow recent behavior, and may be extended to ranges far outside historical experience. Clearly a method must explicitly consider technology in order to be sensitive to changes in it. Consequently, optimization, simulation, and engineering rank high in this respect in Table 3. Case studies, where frequently the technology is invariant, are ranked low. Econometrics, where technological change is sometimes implicitly built into the model and less often described explicitly, must be ranked medium. In judgmental methods, technological change must be posited and responses as to its effect elicited. This additional complexity gives judmental methods medium ranking. In extrapolation, technological trends are assumed to continue as they had in the past, earning it a low ranking with respect to technological or economic changes. Although most input/output models imply that technology is frozen, it is possible to account for technological change, which nudges the techniques to medium.

Economic Factors. Econometric methods invariably include--and, in fact, stress--economic relationships and thus rank high in sensitivity to economic changes. Simulation may include economic factors, giving the respective high and medium rankings. In all other techniques, the roles of prices and economic equilibrium are usually ignored or implicitly assumed to be constant, thus making them unresponsive to economic variables. However it is possible to incorporate assumed economic changes in both judgmental and engineering techniques, which would raise their classsification to medium. When one wishes to trace through the effects of changes in demand or economic factors other than price, input/output indicates the primary and secondary effects in a detailed way, giving it a medium rank.

Transparency

The level of transparency is the degree to which the structure of the relationships among variables can be clearly understood and the effects of changes in variables on forecasts and simulations are logical and consistent. For a method to be transparent, the assumptions and premises that underly the method, data, and relationships should be made explicit and their implications should be intuitively plausible. Although techniques may involve sophisticated statistical or mathematical methods, some may be more easily explained to naive nontechnical users than others. Furthermore, some methods may not be transparent even to experts because of their complexity, particularly in how clearly the model and use of the method are described in technical documentation.

Expert Users. Although econometrics and input/output techniques are well understood, it is possible to introduce unnoticed assumptions, either by accident or by design, that may deceive an expert, earning these techniques a medium ranking. Simulation and engineering approaches are ad hoc techniques, which means confusion is possible, placing them also in the medium category. In judgmental techniques, no one, not even the investigators, knows what is in the minds of the respondents, what their premises, prejudices, and the like may be. Therefore, these are ranked low in transparency. The methods of extrapolation, case studies, and optimization are explicit and have been replicated so many times that deception is unlikely, placing them in the high category.

Naive Users. Transparency to the lay person depends on how logical and straightforward the technique appears to be when explained in nontechnical terms. Extrapolation, case studies, input/output, judgmental, and engineering approaches can be relatively easily explained which earns them high rankings.

Optimization is complex but comprehensible with a moderate amount of explanation, so it gets a medium. Both econometrics and simulation can most likely be explained only by means of black box analogies and mystifying jargon, which earn them a low ranking. (Note that "transparency" in this context does not mean that the naive user has an understanding of the inner workings or even the important relationship in the models. It merely is a measure of how easily one can grasp the overview of the technique. Because of this inconsistency in the distinctions between transparency to expert and naive users, one arrives at the situation where it is specified that the transparency to naive users of certain techniques (i.e., judgmental) is high, but to expert users it is low.)

Ability to Validate

In rating the ability to validate the output of a model, it should be recognized that there is no way to guarantee the accuracy of forecasts or sensitivity analyses. However, some methods explicitly account for variability of past data, which provides a means of measuring how accurately they explain past events. Although other techniques fail to check for accuracy internally, the structure of the model and its output are quantified in a manner that allows using past data to check the accuracy of predictions. External validation is possible for all techniques by means of review by panels of experts.

Internal to Method. Internally derived validation occurs for the statistically based methods of extrapolation and econometrics, which gives them a high rank; however, a unique or limited data base diminishes the value of any statistical internal or external test. Because history and analysis occur coincidently in case studies, the method is given a medium. If iterative methods such as Delphi are used in judgmental models, the coincidence of initial responses and the speed of convergence to consensus or a stable diversity indicates the confidence experts have in their judgments; the method earns a medium. All others ranked low.

The ability to use external information and data to validate model output is potentially high for all methods. However, we chose to limit this criterion to formal, statistical methods of validation. In this case, only econometrics ranks high, because there are numerous formal tests of validation. Some of these tests may be applied to extrapolation, which earns a medium rank. The quantifiable outputs of optimization, and input/output, simulation make them susceptible to external validation, which earns them a medium. Judgmental, engineering, and case studies are ranked low, either because they fail to produce quantifiable outputs or because the outputs may be in a form that makes them difficult to test.

Level of Aggregation

The final criterion that might be employed for selection is the level of aggregation for which the technique is most suitable. Extrapolation and econometrics depend on plentiful, accurate data and a wide range for the variables being observed. Furthermore, the abstract behavior assumed in economics and specified in econometric models is more likely to be revealed at an industry or higher level of aggregation for cross-sectional data and in relatively long trends for time series data, thus requiring a high level of aggregation. Econometrics and extrapolation, therefore, rank high in level of aggregation. Conversely, case studies and engineering models most frequently apply to a single situation or technology and can take many unique and detailed factors into consideration, making them most suitable for individual firms and the lowest level of aggregation. Judgmental models are unique in that they are suitable for all levels of aggregation. Optimization, input/output, and simulation, while nearly as flexible, are classified as being most applicable for medium to high levels of aggregation.

Value of Judgment

A final point should be emphasized. The eight techniques were considered individually in Table 3. However, judgment and feedback from practioners and experts will enhance any of the other seven techniques. Forecasting and analysis are both art and science and require simplification and abstraction. Human judgmental inputs can be used effectively to identify defects in the models themselves and in the forecasts and simulations undertaken with them. The blind application of a single technique is ill advised and at best should be used to stimulate thought and discussion among those who are affected and those who are making decisions.

SUMMARY

The preceding pages have served to outline both the strengths of existing techniques for the study and analysis of materials substitution and those areas of substitution about which these techniques provide little or no insight. The task is to develop new techniques and to refine the existing ones in order to narrow or to eliminate these blind spots.

The incentives that resulted in the formation of the Committee on Materials Substitution Methodology are no less valid today than when the committee was formed. Resource depletion and political upheaval have always been factors to be considered, but the potential for supply disruptions and price validity resulting from effective cartels and/or military actions has enhanced our sensitivity to the need for more effective planning and understanding at the national level.

The benefits that may be derived from the development of these techniques are not limited to the national level. At the industry level, greater understanding of and ability to develop substitution strategies could promote more effective use of natural resources. Furthermore, such techniques could enable industries to identify the point of their greatest sensitivity to materials supply and to act accordingly.

This report outlines the potential of the techniques available for expanding our ability to treat and, potentially, to manage the substitution of one material for another. Further study of these techniques in light of this potential is merited; this report has described a number of specific areas, the exploration of which should broaden their applicability.

CONCLUSIONS

It is clear from the foregoing that no single technique can provide a comprehensive means of forecasting and analyzing substitution among materials. Nevertheless, if one is to understand the effects of such factors as technological change, relative prices, and changes in the regulatory environment on materials substitution, one must have a good understanding of the strengths and weaknesses of the available techniques, as summarized in Tables, 1, 2, and 3.

One of the most strongly held views of the committee was that the analyst is more important than the technique. When the analyst lacks a detailed understanding of the technique, the technology or the market forces, the results are often confusing, misleading, and inconsistent. When successful, the techniques explored here enable the user to anticipate shifts in materials demand and to evaluate technologic and economic opportunities.

Some modeling techniques are better suited to retrospective analysis (econometrics, for example). Others are more suitable for taking future uncertainties and prospects into account (simulation and/or judgmental models, for example). In either case, the most appropriate time frame for analyzing materials substitution is the long run. Since technology for the most part is embedded in capital (machinery, processing equipment, etc.), the capability for materials substitution is severely constrained over the short run when the capital stock is fixed. Therefore, materials substitution, with some exceptions, should be viewed with a telescope rather than a microscope. In other words, the emphasis should be on the long run.

Materials substitution depends on technology embodied in past investments and, therefore, occurs at different rates in each industry as changes in the economic environment make substitution profitable. This implies that disaggregation by material, industry, and product shape is sometimes necessary to understand why substitution has taken place in the past and what its prospects are in the future. Modeling based on data on product shape and end use enables the researcher to track past changes and to predict future ones. Aggregate indexes, in contrast, are inappropriate because of differing technologies and rates of adaptation of technological change among industries.

The analysis and evaluation of the techniques examined in this report reveal, as noted earlier, that no single one is ideally suited to modeling materials substitution. Were such an ideal technique to exist, it would take into account the physical laws of nature, the properties of alternative materials, and the realities of the marketplace. The capital and materials input decision is both an economic and an engineering one. To concentrate on one without considering the other is likely to result in inaccurate or misleading perceptions.

The benefits that may be derived from the development of existing techniques are not limited to the national level. At the industry level, greater understanding of, and ability to develop, substutition strategies could promote more effective use of natural resources. Furthermore, such techniques could enable industries to identify the points of their greatest sensitivity to materials supply and to act accordingly.

In summary, this report outlines the potential of the techniques available for expanding our ability to treat and, eventually, to manage the substitution of one material for another. In light of this potential, the committee believes that the methodological approaches which currently exist are useful for limited purposes. Their utility would be improved by further development and refinement. Moreover, there is considerable room for new conceptual approaches, especially ones that can successfully integrate aspects of economic theory, operations research techniques, and engineering principles.

REFERENCES

Ayres, R. U., S. Noble, and D. Overly. 1971. Effects of Technological Change on and Environmental Implications of, an I-O Analysis for the U.S., 1967-2020, International Research and Technology Report, 219-R/T.

Balachandra, R. 1980. Perceived Usefulness of Technological Forecasting Techniques, Technological Forecasting and Social Change, Vol. 16:155-166.

Bloomfield, P. 1976. Fourier Analysis of Time Series: An Introduction. John Wiley and Sons, New York.

Box, G. E. P. and G. M. Jenkins. 1976. Time Series Analysis, Forecasting, and Control. Holden-Day, San Francisco.

Fisher, J. C. and R. H. Pry. 1971. A Simple Substitution Model of Technological Change, Technological Forecasting and Social Change, Vol. 3.

Goeller, H. E. and A. M. Weinberg. 1976. Science, 191:683-689.

Larkin, J., J. McDermott, D. P. Simon and H. A. Simon. 1980. Expert and Novice Performance in Solving Physics Problems. Science, Vol. 208:1335-1342.

Lendaris, G. 1979. Human Aspects in Structural Modeling, Technological Forecasting and Social Change, Vol. 14:329-351.

Linstone H., and M. Turoff, Editors. 1975. The Delphi Method, Addison-Wesley Publ. Co., Reading, MA.

Linstone H., et. al. 1979. The Use of Structural Modeling for Technology Assessment, Technological Forecasting and Social Change, Vol. 14:291-329.

Loye D. 1978. The Knowable Future. Wiley-Interscience.

Mansfield, E. 1961. Technological Change and the Rate of Imitation, Econometrica, Vol. 29.

Skinner, B. J. 1976. American Scientist, 64:258-269.

Skinner, B. J. 1979. Earth Resources, Proc. National Academy Sciences, 76, No. 9:4212-4217.

Smith, Kerry V. 1979. Scarcity and Growth Revisited, Johns Hopkins Press.

MATERIAL SUBSTITUTION AND TIN CONSUMPTION IN THE BEVERAGE-CONTAINER INDUSTRY*

Frederick R. Demler and John E. Tilton
Department of Mineral Economics
Pennsylvania State University

Much has been written in recent years concerning the importance of material substitution. It is often argued, for example, that the use of relatively abundant materials in place of increasingly scarce materials helps offset, and may even postpone indefinitely, the tendency for real material costs to rise as a result of resource depletion.

Forest products provide a striking illustration of the benefit of material substitution. During the 19th century, the abundance of timber in the United States led to its widespread use as both a structural material and a fuel. By the end of that century, however, the country had consumed much of its available timber stands and faced a shortage of this vital resource that threatened to interrupt the nation's rapid economic development. The substitution of coal, petroleum, and natural gas for wood as a fuel, and iron and steel, aluminum, cement, and plastics for wood as a material, however, averted a severe shortage and allowed the country to continue its rapid rate of economic expansion during the 20th century (Rosenberg 1973).

In addition to alleviating long-run shortages caused by resource depletion, material substitution may in some instances help soften the effects of abrupt and unexpected shortages caused by wars, embargoes, and other types of short-run supply interruptions. Material substitution also plays a major role in the growth of material demand and so must be considered in forecasting future requirements as well as the need for new mining and processing capacity.

---

* This article draws heavily on Demler (forthcoming) and Tilton (1979). It is based on a research that the Pennsylvania State University is conducting on material substitution in tin-using industries (Tilton, forthcoming) under a grant from the National Science Foundation. It is one of four case studies. The others are examining material substitution in vegetable and fruit containers (Grubb, forthcoming), in various uses of solder (Canavan, forthcoming), and in the use of tin chemicals as stablizers in the production of polyvinyl chloride (PVC) plastic (Gill, forthcoming).

Despite its importance, we still have much to learn about the nature of material substitution. For example, is the replacement of one material by another usually motivated by a change in their relative prices? Or, are government regulations, consumer preferences, technological breakthroughs, product competition, and other factors more frequently of overriding importance? What are the principal factors affecting the time requirements for substitution? What kinds of substitutions can be made quickly, and so relieve abrupt and unexpected shortages? What types require years or decades to effect completely?

Tin Consumption in Beverage Containers

To provide insights into such questions, we are examining material substitution and its effects on tin consumption in beer and soft-drink containers in the United States since 1950. The first step in this analysis documents the amount of tin, chromium, steel, aluminum, glass, and plastic used each year. The findings (Figure 1) reveal that tin consumption in beverage containers dropped sharply from 1950 through 1957 and then partially recovered before declining once again during the 1970's. This figure also indicates that, while all of the tin used for beverage containers went into beer cans in 1950, the softdrink market by the end of the 1970's was consuming more than twice as much tin as the beer market.

The Apparent Determinants of Tin Consumption

The second step in the analysis identifies and assesses what we call the apparent determinants of tin consumption. For beer or soft drink containers the five apparent determinants are the barrels of beverage consumed in any given year; the proportion of that consumption that is shipped in packaged containers, such as bottles and cans, as opposed to bulk containers, such as kegs; the proportion of the packaged beverage that is shipped in tinplate cans; the number of average volume tinplate cans required ber barrel of beverage, and the weight of the tin in an average-sized tinplate can. Since the product of these determinants forms an identity with tin consumption, once the values of the former are known, the latter is determined. Thus, changes over time in the amount of tin used in beer or soft drink containers can, in the first instance, be explained by changes in one or more of these determinants.

Figures 2 and 3 show that the consumption of both beer and soft drinks has grown consistently and substantially since 1950, tending to stimulate the use of tin. The figures also indicate that the percentage of total consumption shipped in packaged containers has fallen for beer and risen for soft drinks. So the effect of the second determinant differs for the two beverages, tending to increase the use of tin in beer and to decrease its use in soft drinks.

Packaged sales account for the largest share of both the beer and soft-drink markets, and it is here that the most strenuous competition among materials is found. As Figures 4 and 5 indicate, the returnable bottle dominated the mix of packaged containers in 1950. Its markets share since, however, has deteriorated rapidly in beer and more gradually in soft drinks.

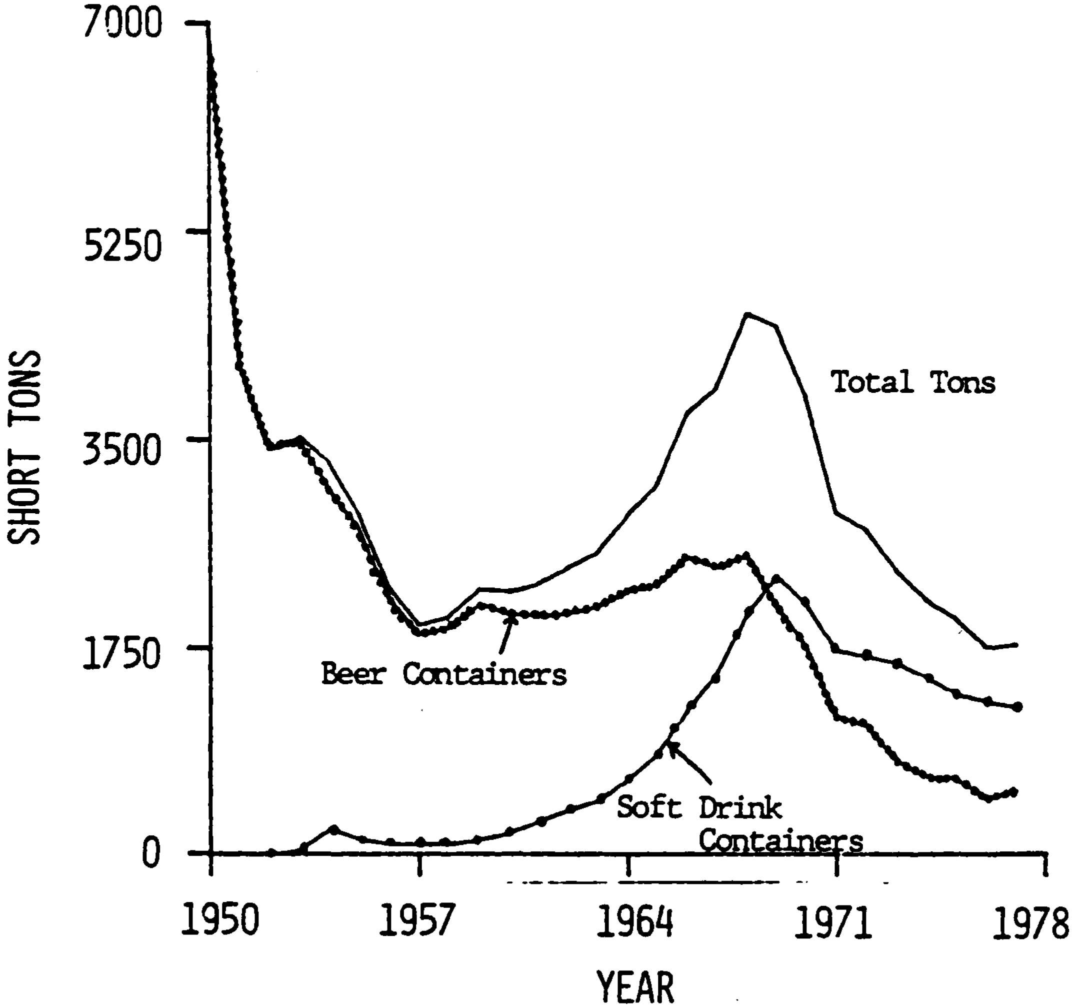


Figure 1. Tons of Tin Consumed in Beverage Containers (Source: Demler)

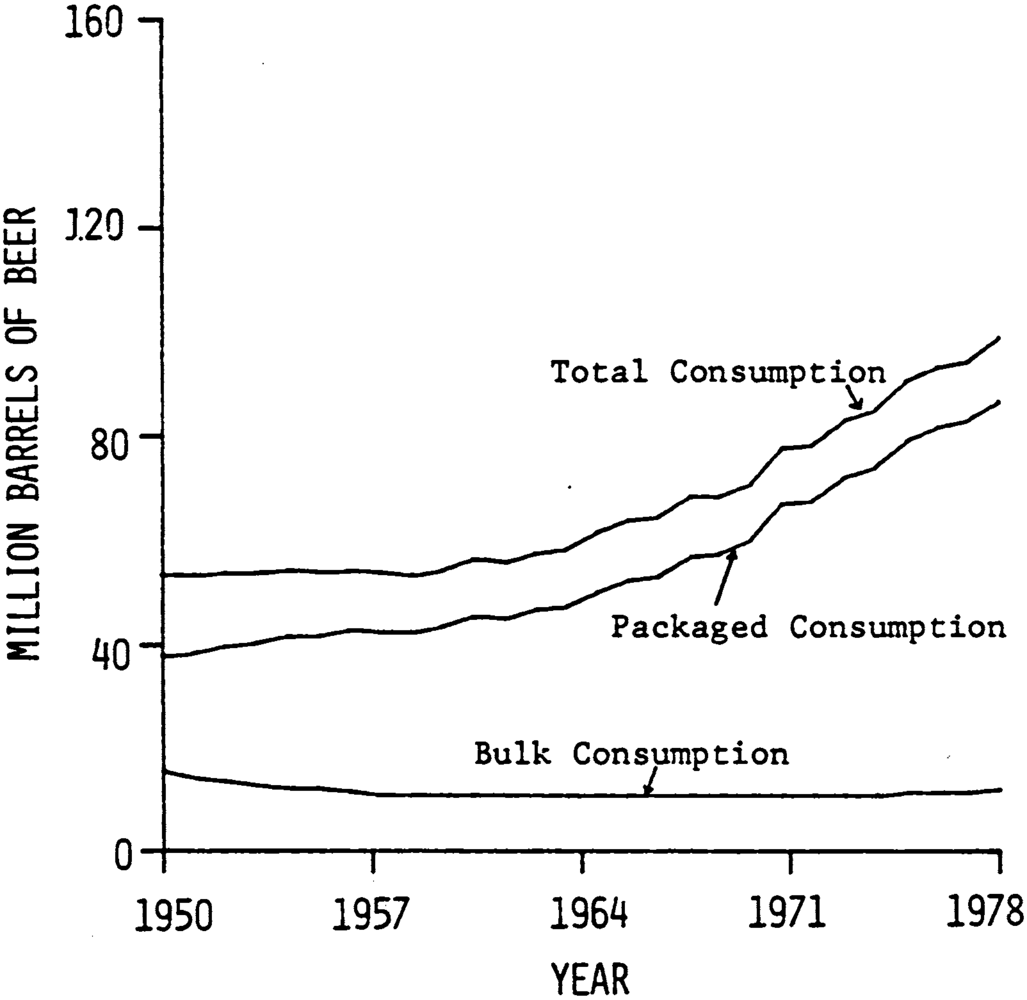


Figure 2. Estimated Annual Consumption of Beer (Source: Demler)

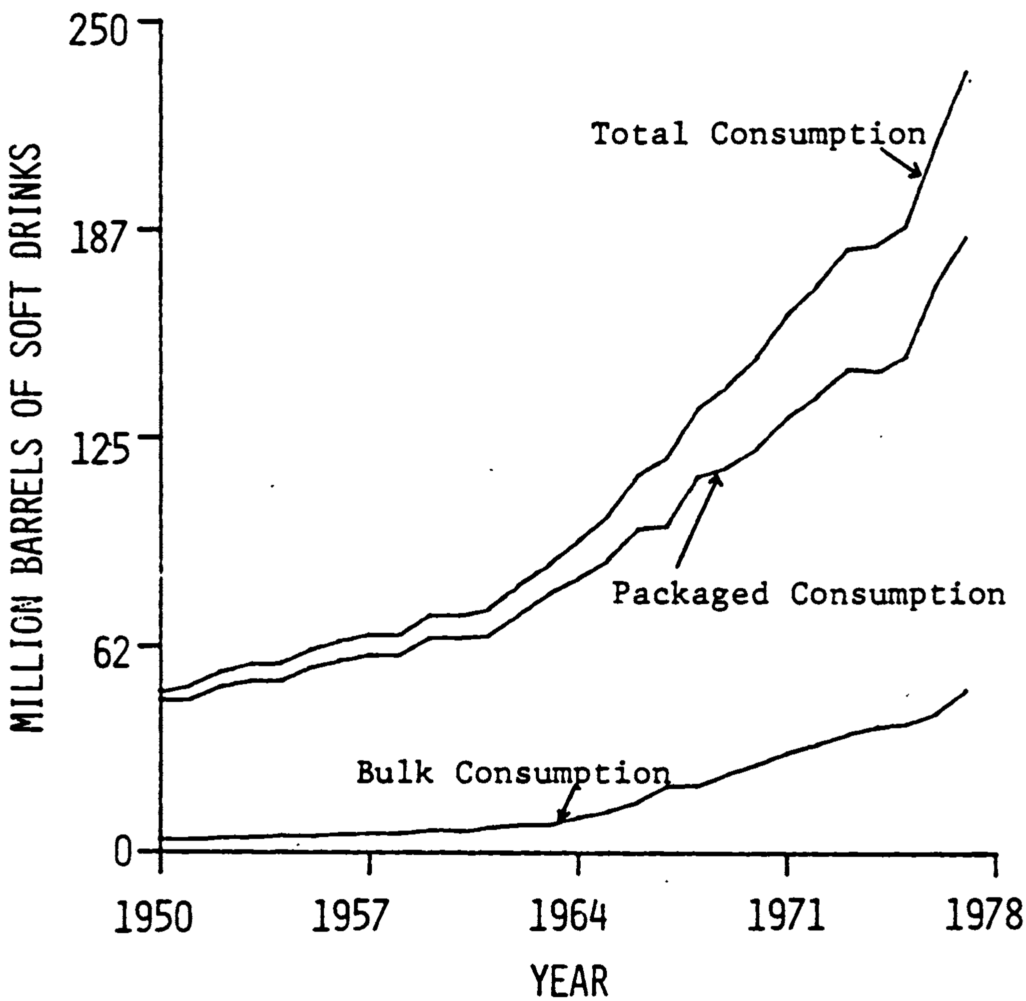


Figure 3. Estimated Annual Consumption of Soft Drinks (Source: Demler)

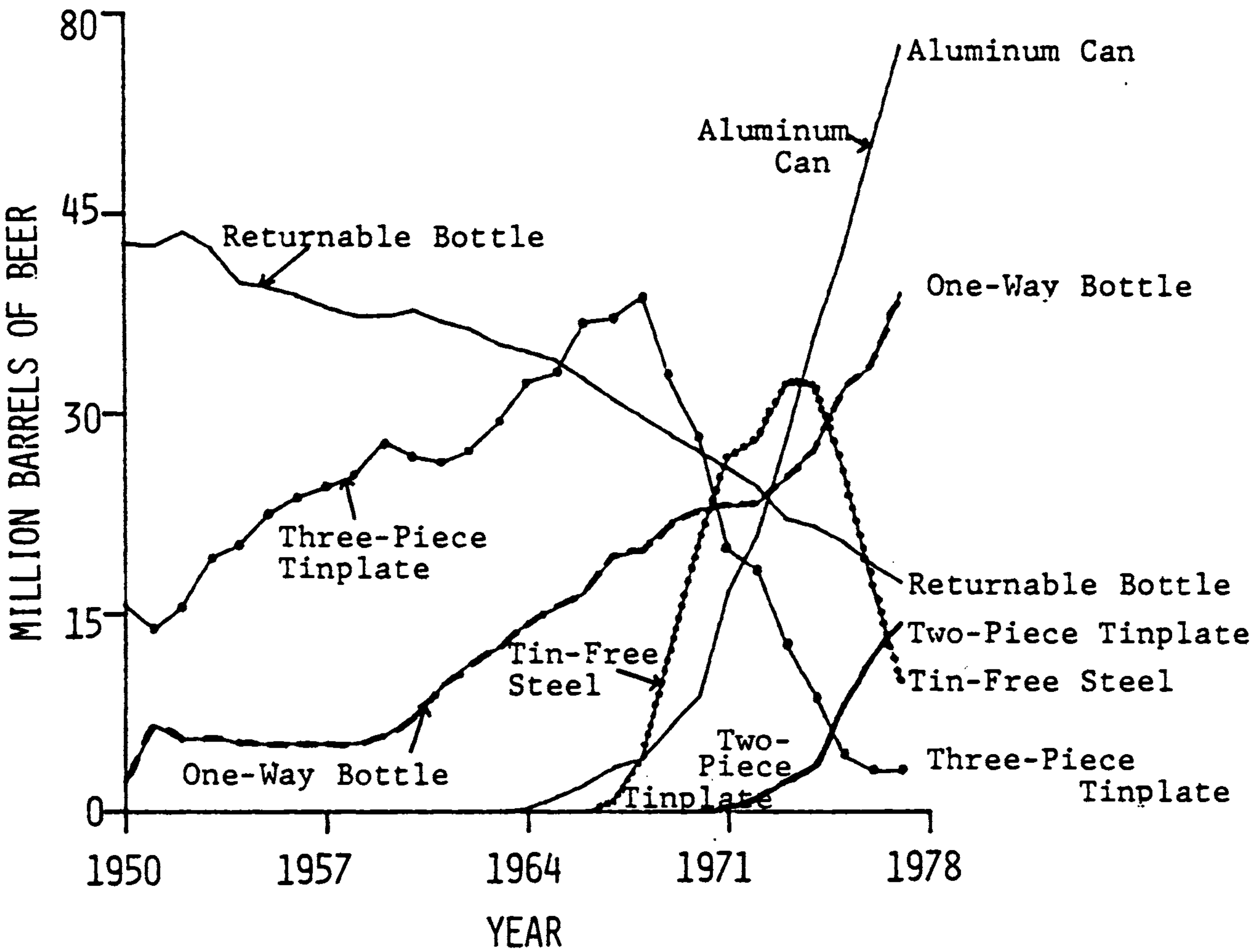


Figure 4. Barrels of Beer Per Container (Source: Demler)

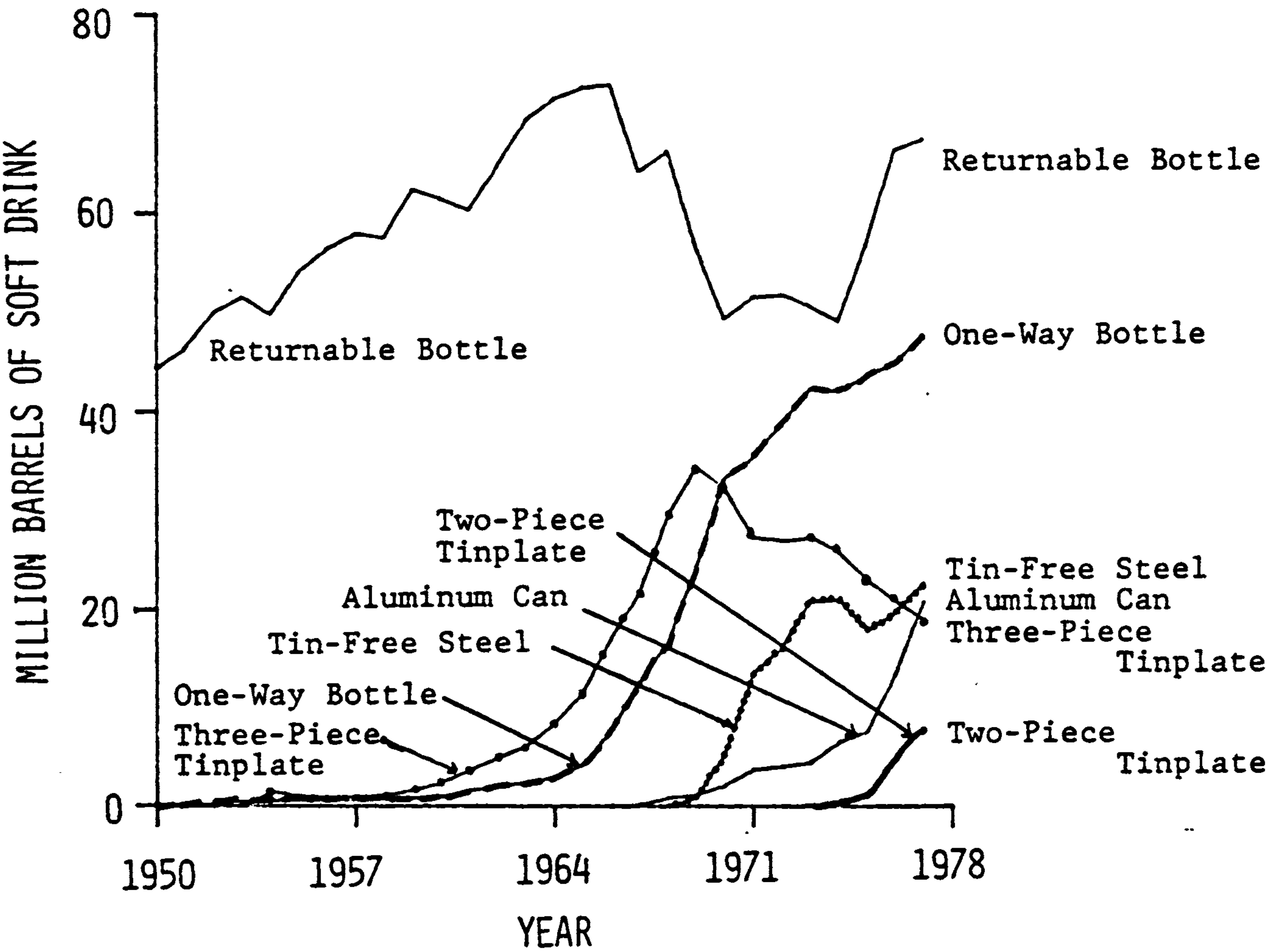


Figure 5. Barrels of Soft Drink Per Container (Source: Demler)

In the beer market, the returnable bottle was surpassed by the tinplate can in 1965, the tin-free steel can (which uses chrome rather than tin to cover the steel sheet) in 1971, and the aluminum can and one-way glass bottle in 1973. Today more beer is shipped in aluminum cans than in any other type of container. In the softdrink market, the new types of containers have been introduced more slowly. Here the returnable bottle is still the most widely used container, though it has faced increasing competition since the early 1960's from the one-way bottle, the tinplate can, the tin-free steel can, the aluminum can, and more recently the plastic bottle.

Between 1950 and 1977, the share of the packaged market for beer containers held by the tinplate can declined slightly, reducing the use of tin. In contrast, its share of the softdrink market increased substantially over this period and in the process stimulated the consumption of tin.

The fourth apparent determinant of tin usage, the number of average-sized tinplate can required per barrel of beverage, changed little over the period. This is because the average-sized tinplate can remained very close to 12 ounces for both beer and soft drinks.

The final determinant is the weight of tin in the average sized-tinplate can. As Figure 6 illustrates, this fell very rapidly between 1950 and 1957 and then continued to decline at a more leisurely pace over the next two decades. As a result, the tin required per can in 1977 was less than 10 percent of the amount in 1950.

Overall, the major determinant fostering the use of tin in beverage containers was the growth in beer and soft-drink consumption. For soft drinks, the penetration of the tinplate can into the market for packaged containers was also an important development stimulating tin usage. The major determinant reducing tin consumption was the substantial reduction in the tin content of the average-sized can.

Underlying Factors Affecting Tin Consumption

While the apparent determinants indicate how tin consumption has changed over time, they are affected in turn by underlying factors--the price of tin and alternative materials, technological change, consumer preferences, government regulations, and a multitude of other considerations. The third step in our analysis involves identifying and assessing the relative effects of the major underlying factors that are ultimately responsible for the changes in tin usage occurring over time in in beverage containers. While this step cannot be carried out with the same empirical precision as the second step, one can nevertheless obtain a fairly reliable picture of the important underlying factors from the trade literature and interviews with industry personnel.

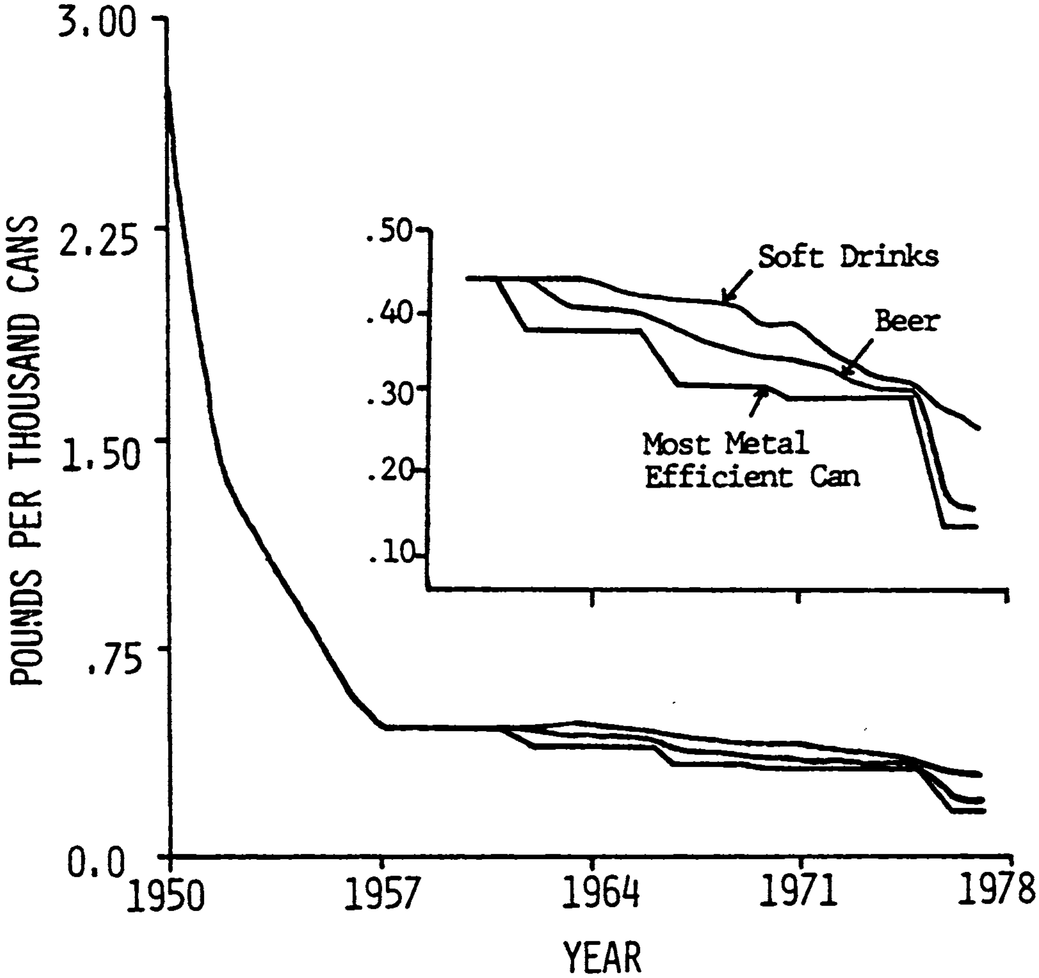


Figure 6. Tin Content of Tin Can (Source: Demler)

Almost all of the growth in beer and soft-drink consumption since 1950 can be explained by changes in three factors; the population of the United States, the age distribution of that population, and per-capita income. The overriding underlying factor affecting the tin content of tinplate cans and the share of the package-container market held by tinplate cans--the two other major determinants of tin consumption--is technological change. The sharp drop in the tin content of tinplate came between 1950 and 1957 (Figure 6) occurred as electrolytic tinning replaced the older, more tin-intensive, hot-dip method of producing tinplate. The more recent reductions in the tin content of the tinplate can can be directly related to the introduction of the aluminum flip top in 1961, the tin-free steel bottom in 1967, the two-piece tinplate can in 1971, and the lighter tin coatings supplemented by enamels during the first half of the 1970's.

The dominant role of new technology in altering the mix of package containers and the share of the tinplate can in that mix is even more readily apparent. In the early 1950's, the tinplate can fabricated by the three-piece technology with the soldered side seam monopolized the can market. Aluminum could not be economically soldered and was thus excluded from the market. In 1958, Coor's Brewing and Beatrice Foods introduced the impact-extruded two-piece aluminum can. While this technology enabled aluminum to be used as a canmaking material, because of lower cost. In 1963, Reynolds Aluminum first produced an aluminum can using the draw-and-iron process, which proved to be a competitive two-piece technology. Between 1963 and 1977, the market share of the aluminum can consistently increased (Figure 4 and 5).

The aluminum challenge forced other container manufacturers to respond. Steelmakers introduced the tin-free steel can, a chrome-plated steel can. Since this can could not be economically soldered, new joining techniques were required. Continental Can developed a welded seam and American Can a cemented seam. With these developments, the use of the tin-free steel can expanded rapidly, largely as a substitute for the more expensive three-piece tinplate can.

Steel producers and canmakers also were experimenting with two-piece steel cans. Pure steel (blackplate), tin-free steel, and tinplate were tested as potential materials for the two-piece technology. Blackplate, the least expensive material, proved to lack the necessary properties of corrosion resistance and lubricity. Lubricity is essential for proper forming of the can body in the two-piece process. Tin-free steel was also relatively inexpensive, but again lacked lubricity. Only tinplate, because of its tin content, had the needed lubricity as well as corrosion resistance, and so it became the material used to produce two-piece steel cans. These cans first appeared for beer in 1971 and for soft drinks in 1973. As Figure 4 and 5 indicate, their growth since has helped the tinplate can maintain or recover its share of the package market in beer and softdrink containers.

Material prices have also been important, as one could expect on the basis of economic theory. However, their influence has been largely indirect, exerted through the incentives they provide for developing new cost-saving technologies. Very few instances can be found where one material can be substituted for another using the same technology and production facilities. One significant exception is the dual canmaking facility, which can switch between tinplate and aluminum sheet in about four hours. Such can lines, however, still account for only a small proportion of total capacity, having been first introduced in 1976.

A number of other underlying factors have also influenced tin consumption in beer and softdrink containers. Legislation requiring deposits on containers to encourage their recycling has favored the returnable bottle, and among cans, the aluminum container, which is composed of only one material and so is easier to recycle. Changes in social customs that have led to the decline of the local tavern and the rise of the fast-food chains have altered the relative importance of packaged and bulk containers. The desire for convenience has favored the use of the aluminum flip top, and the taste preferences of certain consumers have helped maintain the use of the totally inert glass bottle. The larger size of beer plants compared to soft-drink bottlers has encouraged the earlier introduction and diffusion of new containers in that beverage. Although these factors have all influenced tin consumption, the most important underlying factors are technology and material prices, along with the trends in demography and per-capita income which are largely responsible for the substantial growth in beer and soft-drink consumption over the past 30 years.

Implications

While it is hazardous to generalize on the basis of one or even several case studies, our investigation of tin consumption in beer and softdrink containers, suggests that material substitution greatly affects material requirements over the longer run, a period of 20 years or more, and may even substantially alter requirements within a period as short as several years. Thus, forecasting techniques that do not explicitly take into account the effects of material substitution are likely to perform poorly.

In addition, the finding that technological change is the dominant underlying factor causing material substitution and that changes in material prices largely influence material demand indirectly by influencing the rate and direction of new technology has two important implications. First, it suggests that material substitution greatly complicates forecasting future requirements, for the ultimate effects of the new technology induced by price changes, along with the speed with which it will be generated, are highly uncertain and difficult to predict. Second, it implies that material substitution may not be great help in alleviating shortages due to wars, embargoes, strikes, cyclical surges in demand, inadequate investment in new mines and processing facilities, and other causes that tend to persist for only a few years. By the time shortages and the resulting higher prices have stimulated the technological developments needed to permit the substitution of more readily available and cheaper materials, such shortages are likely to have passed. In the interim, when the dislocation to the economy may be quite severe, material substitution may be limited in the contribution it can make.

REFERENCES

Canavan, P. D. Solder, Forthcoming.

Demler, F. R. Beverage Containers, Forthcoming.

Gill, D. G. Tin Stablizers and the Pipe Industry, Forthcoming.

Grubb, T. J. Fruit and Vegetable Containers, Forthcoming.

Rosenberg, N. 1973. Innovative Responses to Materials Shortages, American Economic Review, 63, 2:111-118.

Tilton, J. E. Editor, Materials Substitution: The Experience of Tin-Using Industries, Resources for the Future, Washington, D.C., Forthcoming.

Tilton, J. E. 1979. Understanding Material Substitution, Earth and Mineral Sciences, 48, 7:50, 52-54.

ANALYSIS OF THE MATERIALS SYSTEM--THE ROLE OF OPTIMIZATION

Kenneth C. Hoffman
Mathtech, Inc.

The materials system is composed of integrated set of resources and technological processes operating in a complex institutional framework. Substitution among resources is an important feature of the system and occurs in response to relative price changes, scarcity, technological change, and regulation.

Substitution may take place at several points in the materials system. Resources may be substituted to produce virtually identical products or close equivalents--e.g., natural rubber from guayule or hevea, and synthetic rubber from oil or coal. Substantially different materials may be sutstituted in end-use functions--e.g., aluminum for steel in automobiles. More subtle but highly significant substitutions occur due to more drastic changes in the end-use function--e.g., the displacement of energy intensive vacuum tubes and associated electric supply systems with transistors and later, intergrated circuits.

Clearly, the analysis of substitution possibilities among materials is an extremely complex affair. The role of optimization will vary, depending primarily on the point in the materials system at which the substitution may occur. In general, the possibilities for insights through formal optimization are greatest when the competing resources or technologies produce almost identical products or services. Under these circumstances, the market share among competitors can be analyzed as a relatively simple function of the cost of the alternative material resources, cost of conversion technologies, efficiency of use of material resources, cost of ancillary inputs such as energy and labor, and environmental or regulatory constraints. These variables can be handled in most process optimization models using a single profit or welfare maximizing objective function. On the other hand, when the products from alternative resources or processes differ with respect to quality, durability, esthetics, etc., it is difficult to handle this multi-attribute problem in optimization models. Techniques exist for multi-objective optimization; however, they are difficult to apply in practice and give rise to considerable sensitivity to uncertain technical parameters and judgmental weightings.

The discussion up to this point has indicated the complexity of the materials system and the differences in approach that apply at various points in the system. It will be helpful to clarify the nature of the energy system and, for this purpose, the Reference Materials System (RMS) will be employed. The RMS is a network description of the physical flows of materials from resource to end use. The RMS evolved from previous work on an analytical framework for energy analysis--the Reference Energy System. The RMS was developed using similar principles as an information framework in support of the Committee on Renewable Resources for Industrial Materials of the National Research Council.

In any analytical activity it is necessary to first define the structure and scope of the system being analyzed and to develop an information base containing all relevant and available data. No such comprehensive data base exists for the materials sector of the economy and this is a serious obstacle to the analysis or optimization of substitution possibilities. Several excellent individual studies have been done in specific sectors; however, these cannot be assembled or integrated into a comprehensive picture due to the absence of standardized units and definitions that cover the entire materials system. A substantial portion of this paper will be devoted to the outline of a materials information system based on the RMS. The intent is to stimulate comment and thinking on the structure and form of a practical information system using the RMS as a "straw man." Following that discussion of the structure of the materials system, the role of optimization in analyzing substitution possibilities at various steps in the system is reviewed.

Elements of the Materials System

It is traditional to review the economy of a nation from the perspective of financial institutions with production, trade and consumption expressed in monetary units. Many of the policy levers available to governments are of a monetary or fiscal nature, so it is quite understandable that most information systems dealing with major sectors of the economy stress this type of economic data. As resource problems arise in specific sectors of the economy, attention must be focused on the physical aspects of production, trade, and consumption. In addition, the recognition of the need for long term research and development to solve resource supply and substitution and conservation problems lead to an increased need for comprehensive information on energy and materials through economy.

Information on the physical aspects of resource supply, conversion, and utilization does not, of course, replace financial economic data, but is complementary to such data in providing a complete picture of the structure of the economy of a nation. This paper outlines a framework that may be employed to organize information on the physical flow of materials from their harvesting, or extraction, through the conversion steps required to produce useful materials, to their utilization, maintenance, and recycling in specific sectors. The incorporation of the utilization step is of special importance since it is this portion of the materials systems that governs the conservation of materials and the substitution of abundant materials for scarce ones. While the information system organized about the physical flow of materials through the economy, other factors of production in the economy such as energy, labor, and capital, may be incorporated along with environmental effects.

The materials information system outlined here is compatible with a large variety of data systems and analytical models. Coupling of the information system to simulation models and economic models has been demonstrated in a conceptual way.

The availability of materials for housing, durable goods, industrial construction, transportation systems, and energy is central to the life-style and prosperity of a nation. The materials system is quite complex in view of the existence of a large number of natural sources of renewable and nonrenewable character, and the multitude of technical activities operating within a complex institutional framework. The technical activities include the extrapolation for a wide range of materials resources, conversion of these resources into useful products, operation and maintenance of these products over their life spans, and, finally, recovery or recycling of these products back into the resource stream. Although the materials system itself is a vital element of the nation's economy, this system has close relationships with other sectors including its effect on employment, energy needs, capital requirements, and the environment. Technical and policy options designed to deal with specific issues may alter the trade-offs among these sectors.

While energy problems occupy much of the nation's attention and are dealt with by a cabinet-level agency, the Department of Energy, there is no focal point for the formulation and coordination of materials policies. Supply, demand, and allocations within the U.S. materials system are largely determined by independent forces working through the market in the private sector. However, the problems arising from growing environmental concern and changing patterns in the international supply and demand of resources generally induce changes in resource markets that are outside the scope of the decision-making capacity of the private sector. Government support for research and development in the materials system is increasing but is still quite fragmented. Government policies as well as private sector decisions must be based on improved up-to-date knowledge of the technical, economic, and environmental parameters of the materials system. This kind of information is also sought by scientists and engineers who need technical data on materials properties and processes, and by industrial managers who seek information on materials supply, demand, and potential markets.

A large number of formal and informal materials information systems have been devised, both in private and public sectors. Unfortunately these systems, in addition to being quite disparate and incompatible, are generally deficient in that they consider only isolated aspects of the materials system. The need to address the broad technical and policy questions in both the public and private sectors points toward the requirements for a framework within which economic, environmental, and technical factors involved in the supply and utilization of all alternative materials may be simultaneously considered for analysis of the materials system. The objective of this section is to outline a comprehensive framework, the RMS, that may be used to organize relevant information. In addition, the framework is compatible with a wide variety of analytical methods that may be employed to assess the broad impacts of materials policies. The RMS represents the supply and demand balance in the materials system and the technologies employed to produce and utilize materials.

An important feature of this framework is the incorporation of the utilization, maintenance, and recycling portion of the system at the same level of detail as the supply side. These portions of the materials system are often ignored in policy analysis.

Many studies have been performed on the energy and environmental aspects of materials production. Makino and Berry (1973) and Midwest Research Institute (Hunt and Welch 1974) have published information on the energy inputs to the production of glass, aluminum, and plastic container materials, and Ayres and Kneese (1969) has analyzed environmental impacts associated with materials production. Hannon (1972) has considered the direct and indirect energy inputs to materials using input/output modeling in the analysis of recycling policies. The RMS format provides a comprehensive and standard format in which the results of such process analysis of specific materials and production steps may be displayed. The methodology is similar to the Reference Energy System which has been coupled to inter-industry models of the economy (Hoffman 1975) and can be used in a similar manner to provide a generalized coupled process and economic model for use in technology and policy analysis.

The RMS concept has been employed as the central systems analysis approach by the Committee on Renewable Resources of the National Research Council. The thrust of the study was to identify the most promising areas for substituting nonrenewables by renewables which in turn would highlight the research and development programs needed to overcome the barriers to production and use of renewable resources. The RMS approach has also been adopted for a study (Kearney 1977) in Ireland concerned with the use of biomass as a source of energy. Although the specific emphasis on the various policy objectives will vary from country to country depending upon its stage of development, mineral base, etc., the RMS, because of its general nature, can be adapted as a policy and planning tool to any national situation. For example, trade-offs between the labor requirements and capital expenditures as influenced by a particular technology will be somewhat different in an industrialized country as compared with a developing country where the policy objectives may differ. Such policy objectives are exogenous to the RMS and may be formulated independently.

The nation's materials system can be thought of as consisting of an integrated set of technical activities such as exploration, refining, conversion, transportation, fabrication of material resources into useful products, and, finally, the maintenance and recycling of these products. The RMS is a network representation of the physical flow of materials through all of the production and utilization steps that a resource must go through to be used for a specific purpose in the economy. The scope of the RMS is outlined in Table 4. At the left-hand side is a listing of resources, both renewable and nonrenewable, while the products and end uses, defined at the functional level, are listed on the right side. The definition of the use of materials for specific functions and purposes is central to the RMS concept. Only at this level can conservation and substitution opportunities be analyzed with any technical reliability. Engineering properties such as strength-to-weight ratios, corrosion resistance, and durability must be considered.

Table 4. Scope of Reference Materials System and Associated Data Elements

Resource Base	Production (Growing)	Harvesting or Extraction	Processing	Transportation (Aggregate)	Fabrication and Recycling	Product Identification	Additional Fabrication (e.g., erection) &	End Use and Recycling
RENEWABLES		Water Use				Lumber		Commercial &
		-Consumptive						Industrial
Forest Resources		-Nonconsumptive				Plywood		Structures
Grazing & Rearing Land		Land use				Paper		Housing
Resources								
- Birds		Energy				Particle Board		Transportation
- Cattle						& Fiberboard		Furniture &
- Sheep		Fertilizer &						Upholstry
		Chemicals				Chemicals		
Crop Land Resources								
-Cotton		Labor				Fibers & Woven		
Energy								
-Cereal & Sugar Cane						Fabrics		-Fuel
		Environmental	Data elements to be identified					-Power
Forest Resources		-Air Pollutants	for each resource/activity			Nonwoven Fabrics		
-Coconuts		-Water Pollutants	combination.					Books
Publication								
-Citrus Peel		-Solid Waste				Elastomers		Producer Goods
-Gum		-Risks				Fuels		
Marine Resources,		Capitol Goods						Fabrics
Including Agricultural						Plastics		-Clothing
Types		Operating Main-						-soft goods
-Algae		tenance Costs				Aluminum Mill		(footwear)
-Menhaden, etc.						Products		-Packaging
		Institutional &				Steel Mill		Communication
NONRENEWABLES		Organization				Products		Disposable
Products								
						Concrete		-Packaging
Aluminum								-Other
Iron & Steel								Chemicals
								Recreation
(Competes								
Cement & Concrete								for use of
Land)								
Oil & Gas								

The completed RMS, involving a network representation of the flow of materials from the resource side through all of the "activities" listed along the top to a specific end use, such as building and construction, for the year 1977 is shown in Figure 7. This figure is quantified in terms of the mass of material flowing annually through each activity. While the material flows on the supply side were obtained from the Statistical Abstract (1978) and the annual statistical reports, of The Aluminum Association (1977) and the Iron and Steel Institute (1978), put out by various trade associations, the data on the demand side were mostly estimated using the product mixes and conversion ratios, as they existed in the year 1974, from the materials source book (Bhagat 1976). The network can also be quantified in terms of energy use, cost, labor, and environmental effects associated with each activity. A path from a specific resource to a specific end use is called a "trajectory." Each "activity" in the trajectory represents a technical process or production step that is characterized by both a material flow element (and material losses) and the data elements listed--e.g., energy requirements, other material inputs, labor and capital needs, and environmental effects. The activity category involving "installation, erection, and maintenance," not relevant in the energy system, is of special importance in the case of a materials system for evaluating life-cycle usage characteristics of materials. Opportunities for recycling of materials are identified in terms of activities characterized by material flows and data elements. Imports and exports of resources and products can be indicated by flow vectors from and into the appropriate nodes.

The RMS illustrated in Figure 7 is simplified and aggregated for presentation purposes only. Additional detail is provided in versions of this system that have been developed for policy studies. An example of additional information that is needed is alloying materials such as chromium, molybdenum, and cobalt that provide desired strength and corrosion resistant properties for certain applications.

It is feared by many that resource scarcity will limit future economic and social development. Analysis of the role of materials in our society requires the extension of the RMS to a general economic framework. The conventional input/output framework provides a detailed picture of the structure of the economy of inter-industry flows. While normally quantified in monetary units, input/output tables have also been quantified in physical terms (mass flows, energy flows, etc.). The RMS provides the basis for estimating the technological coefficients and material substitutions represented in the input/output tables. Figure 8 shows the format of a modified input/output table. The flow of materials resources through the materials conversion processes into the other nonmaterial industry sectors and the final demand sectors is represented by coefficients representing the mass of specific materials required per dollar or physical unit of output in the industry sectors. The summation of total outputs in dollar terms represents the Gross National Product (GNP) of the nation. This framework then provides the analytical link between GNP (which when exhibited in terms of individual sector elements is representative of a life-style pattern) and the requirement for specific materials. When presented at this level of detail, the results of engineering analysis may be represented in a policy framework. This step of introducing the physical representation of a technical system in an economic framework has been accomplished for the energy system but not as yet for the materials system.

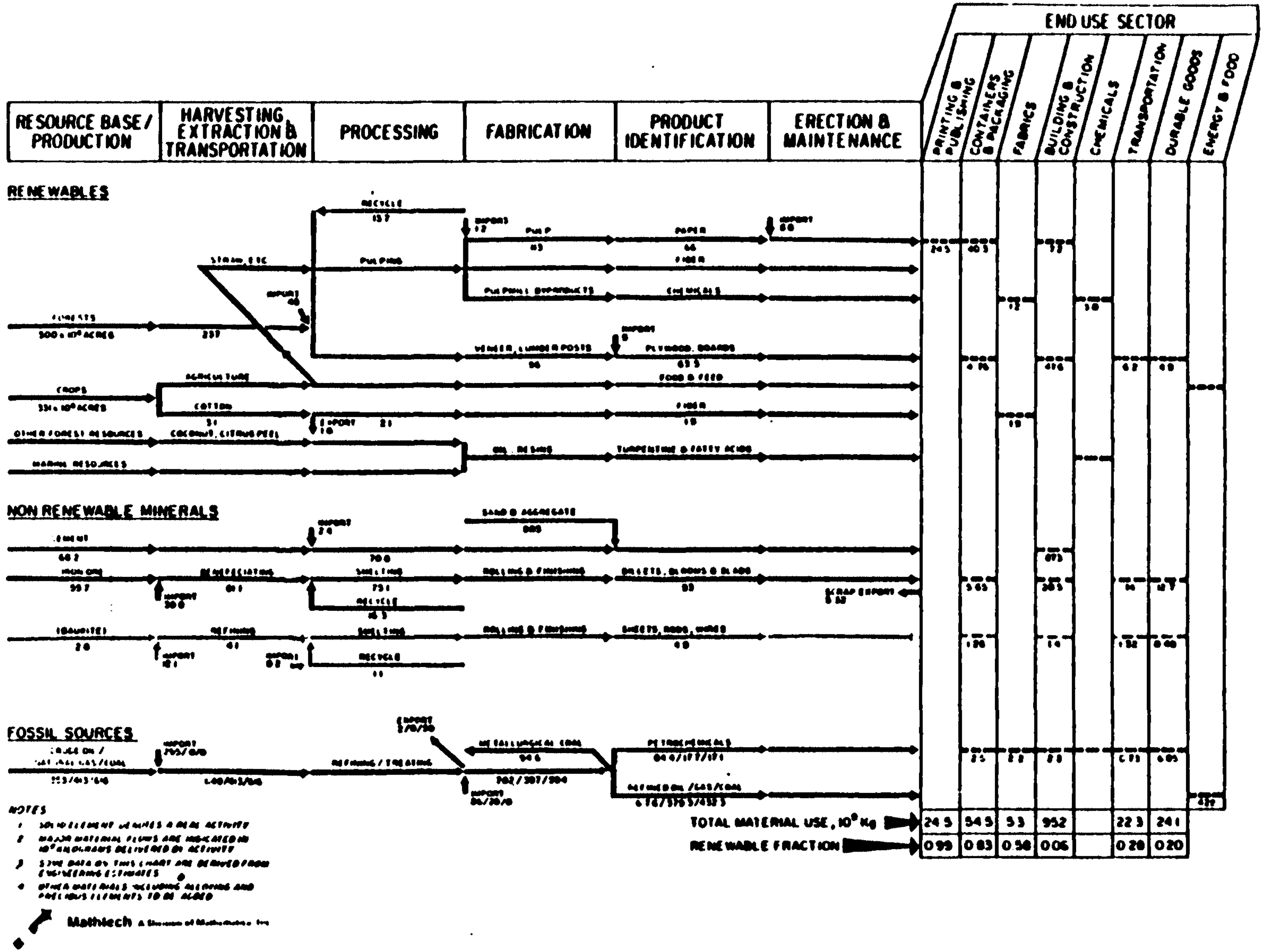


Figure 7. U. S. Materials System, 1977

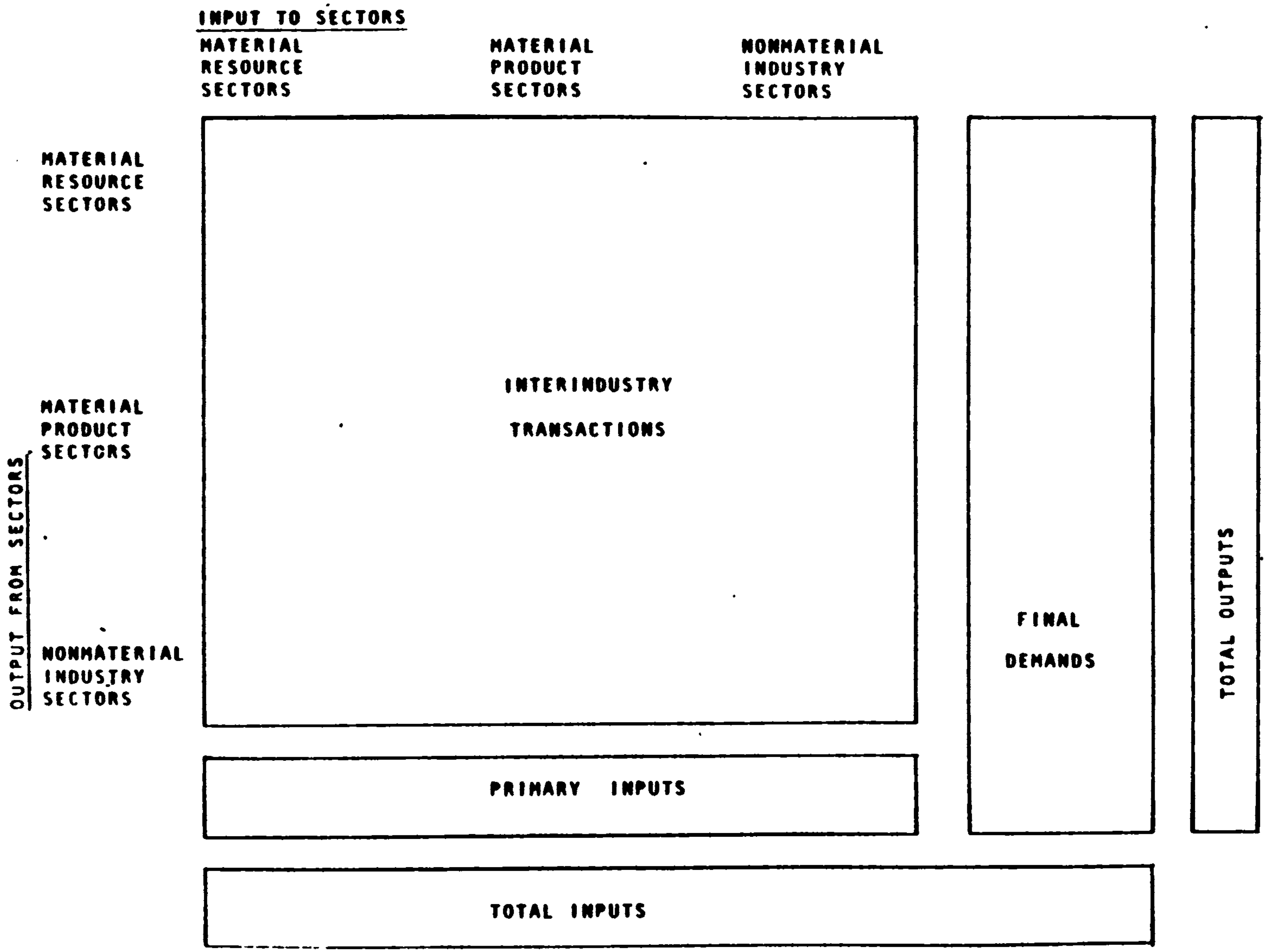


Figure 8. Material Utilization in the Economy, Data Format (Source: Demler)

The logic of incorporating a physical representation of a technical system in an economic framework along with consideration or resource, labor, capital, and environmental factors is illustrated in Figure 9. This figure illustrates the way in which resources and technology underlie the economy of a nation and affect its environment. Starting at the bottom, resources are employed in technological systems to produce goods and services in the economy. Environmental effects are also produced that must be balanced against benefits of production. Policy actions or decisions taken at any level can affect the need for and use of the materials and technology employed in the nation's economy.

RMS projections of material flows, compatible with the economic forecasts for future years, say 1985 and 2000, can be prepared, assuming a natural evolution of technologies and no new federal policy initiatives. This projected system can then be used as a base case for substitution analysis and technology assessment as discussed in the following sections. The RMS can be prepared to represent the flow of materials through an industry, regions of the country, or the entire country.

Analysis of Materials Utilization and Substitution

The RMS and the associated data can support models and analysis of materials utilization and substitution. The overall incremental effects of the substitution determined by optimization or other appropriate methods can be ascertained by adjusting the material flows and attendant energy, economic, and environmental implications indicated on the RMS diagram and backup data sheets.

The overall analysis involves these steps, following the definition of a base, or most likely, case, in the RMS format:

1. Determination of the specific supply or end-use sectors of the RMS to be considered in the analysis.

2. Definition of new processes and resources to be analyzed in the affected trajectory from the resource to the specific end use (definition of losses, energy, labor and capital requirements, and environmental effects).

3. Analysis, using optimization or other appropriate models of flows through the affected trajectories in the RMS to reflect the revised utilization or substitution of materials and/or new processes.

4. Accumulation and tabulation of resource, energy, labor, capital, and environmental consequences of the utilization or substitution in the RMS format.

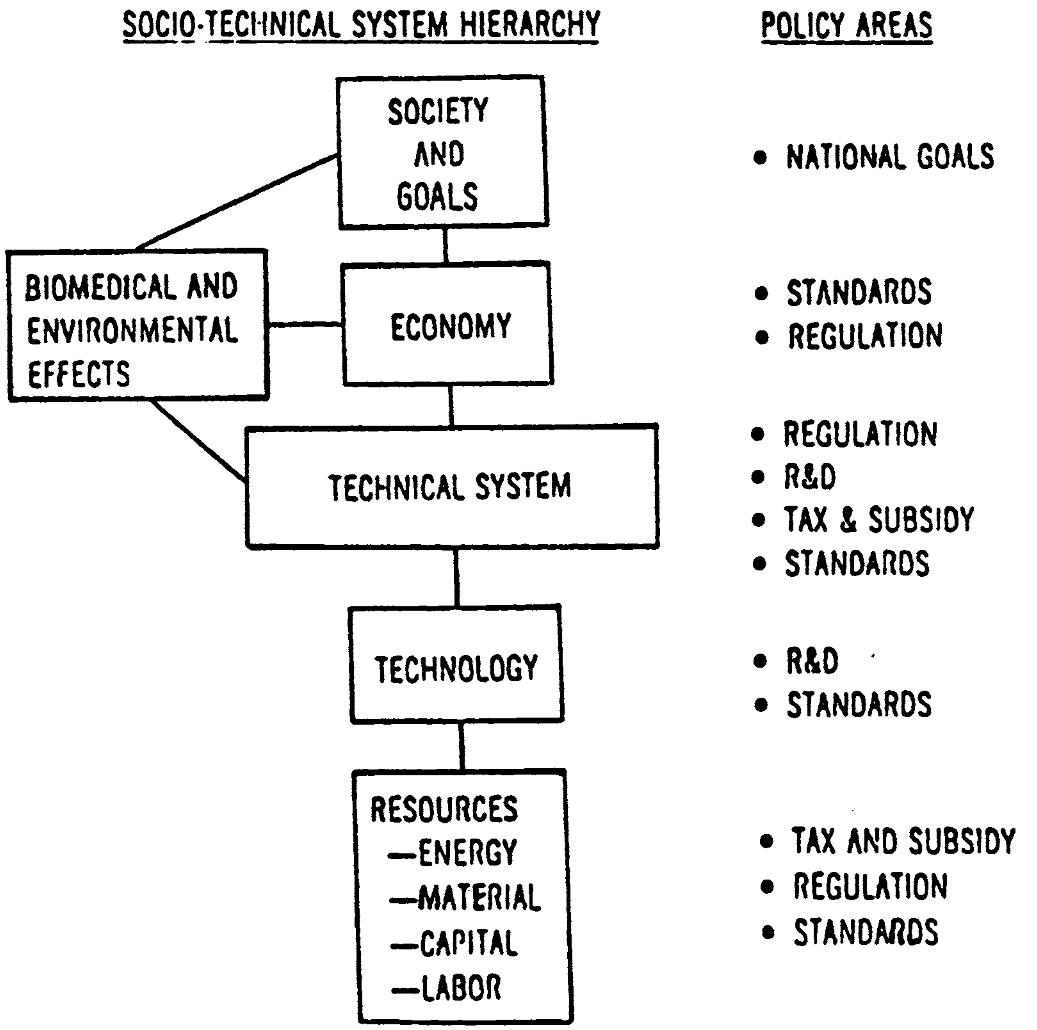


Figure 9. SocioTechnical System Hierarchy

In analyzing the specific nature of the substitution, it is necessary to address the specific application. The mass ratio of substitution (e.g., kg of paper that would replace a kg of plastic) depends on the specific application and the nature of the material. Thus, at the point of end use, one would have to focus on paper bags, for example, as a substitute for polyethylene bags. The determination of these substitution ratios must be done exogeneously, using optimization, simulation, or judgmental techniques, and the results reflected in the revised or perturbed RMS. In certain instances, material preferences and substitution may be constrained or influenced by such factors as esthetics and codes or standards.

The parameters of the technical characteristics of new processes must also be developed for inclusion in the data base by people with a process background. The intent of the RMS format is to capture those characteristics of the technology that are important to materials policy formulation and make them available in a consistent and comprehensive format.

Following these steps, the pertubation of the appropriate trajectories and the accumulation of information on detailed consequences is straightforward using the RMS. In the case of an analysis of the substitution of paper bags for polyethylene bags, for example, the flows through the wood-to-paper trajectory would increase by the appropriate amount while the flow of crude oil and natural gas through the petrochemical trajectory would be decreased. The full materials system implications may then be traced all the way back to the forest and the source of the oil, imported or domestic. The results of the analysis may then be used as a basis of support or revision of the original utilization of substitution measure.

When used in this fashion, the RMS can be a useful technique for the analysis of materials policy. It must be recognized that the technique focuses on the physical structure of the system and its requirements. Thus, although substitution analysis may be performed in a rather direct manner, in cases of more general policy analysis, the effects of a policy action on the supply or demand for materials use and on the physical structure of the system must be developed or estimated prior to use of the RMS.

The case study to evaluate the energy implications of substitution of plastics by paper products for certain kinds of packaging and containers follows.

Case Study of Material Substitution in Containers and Packaging Sector

Packaging is used for three major classes of goods: durable, nondurable, and foodstuffs. The overwhelming fraction of durable goods is packaged in corrugated cardboard. Corrugated cardboard is also most commonly used as a packing material in the case of durables. Nondurables consists of clothing, textiles, and chemicals and require a variety of packaging characteristics. Foodstuffs, the third major area for packaging, represents about 15 percent of the production activity of the U.S. economy and account for 60 percent of the total shipment value of the entire range of goods that are packaged. This sector involves the widest variety and largest amount of packaging materials, apart from corrugated cardboard, produced from renewable resources. In the following discussion, specific examples have been chosen for which both nonrenewables and renewables can be interchangeably used to meet certain packaging requirements. Such examples are: sanitary food containers used for milk, butter, margerine, frozen foods, ice cream, shortening, etc; trays for packaging meats, eggs, and produce; and flexible containers--e.g., bags and sacks.

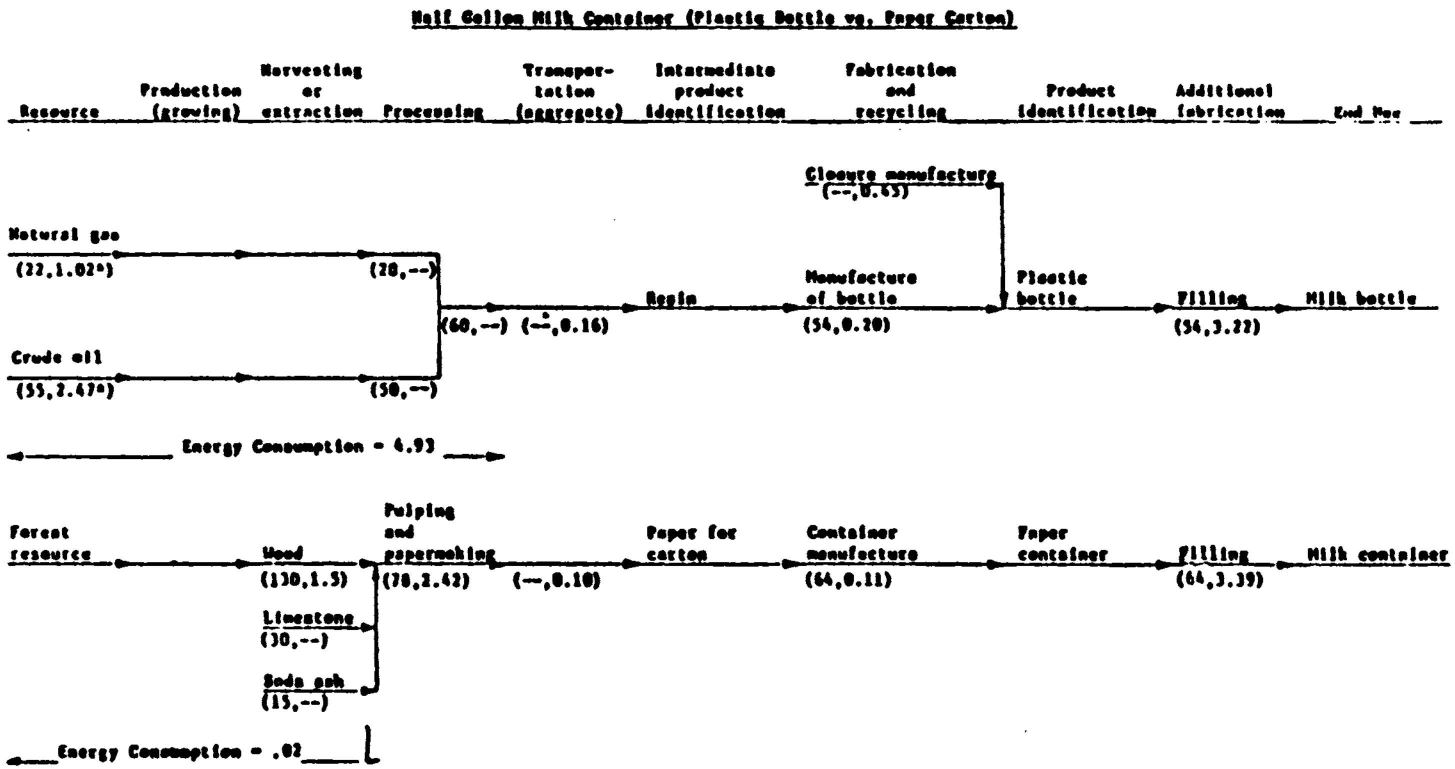
Although labor requirements and capital costs are also important considerations in the comparison of alternative materials, attention is focused exclusively on energy implications in this case study of materials for containers and packaging.

In connection with sanitary food containers, two RMS trajectories are shown in Figure 10. These correspond to the special case of half-gallon containers made of plastic and paper. Mass flows and energy values shown in the figure under each activity link refer to requirements for manufacture of one container of each type. Energy data are in terms of the "gross" value of energy requirement. Summing all the energy components along the two trajectories, one can see that plastic bottle weighing 54 grams needs about 8.4×10^6 joules, whereas an equivalent paper carton weighing 64 grams needs 6.4×10^6 joules. Also, the plastic bottle requires 22 grams and 55 grams of natural gas and crude oil, respectively, as chemical feedstock, while an equivalent paper carton needs 130 grams of groundwood. Adding the energy content of raw materials, the total energy inputs to a plastic bottle and an equivalent paper carton work out to 11.9×10^6 and 7.9×10^6 joules, respectively.

In Figure 11, two trajectories for the manufacture of size 6 meat trays from styrofoam and from molded wood pulp are shown. The energy requirement in the two cases add up to about the same value, 0.9×10^6 joules each. Here again, taking into account that 2.3 grams of natural gas and 7.2 grams of crude oil are needed as chemical feedstocks in the case of the polystyrene tray and 30 grams of groundwood is needed as raw material for one pulp tray, the total energy values increase to 1.3×10^6 joules, remaining the same in both cases.

In the case of flexible containers, polyethylene is used for plastic bags and Kraft paper for paper bags. The energy cost of Kraft paper is (Table 5) $\approx 48 \times 10^6$ joules/kg, and that of polyethylene $\approx 160 \times 10^6$ joules/kg or 3.3 times as much. But, because medium-weight polyethylene bags weigh only half as much as an equivalent paper bag, the ratio of energy consumption of plastic and paper bags is $\approx 1.65:1$.

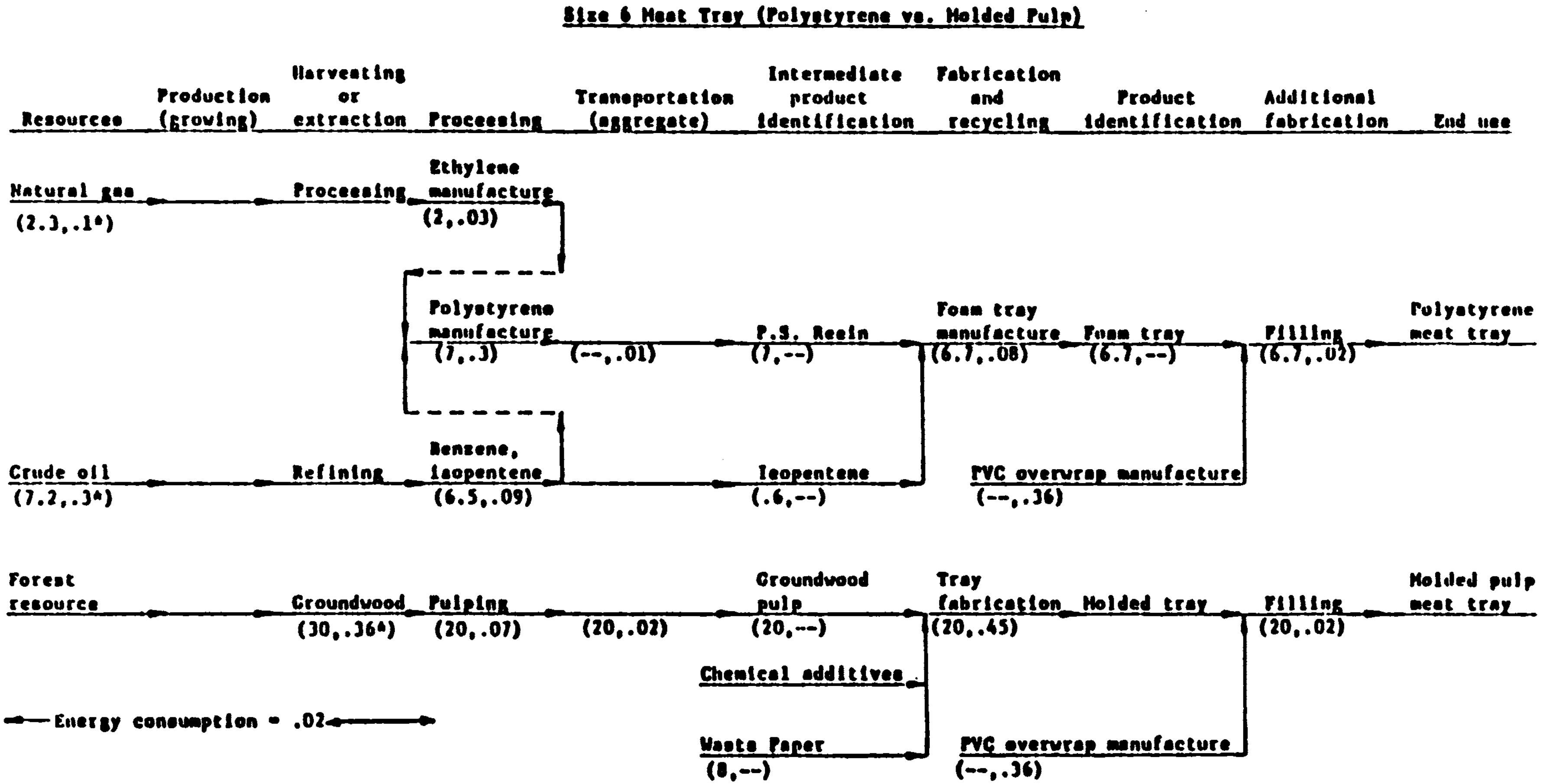
The above comparison is not entirely fair to plastics if there is the possibility of reusing the plastic containers. As an example, to make and fill a half-gallon plastic milk container a single time requires about 8.4×10^6 joules of energy. If it were reused, and the washing and filling costs remained the same with each use ($\approx 3.2 \times 10^6$ joules), then the cost



Note: Numbers in the parentheses below the activity links refer to mass flow in grams and energy requirement in million joules, respectively, for the corresponding steps in the manufacture of one-half gallon milk containers.

* Energy content of resource.

Figure 10. Reference Materials Trajectories



Note: Numbers in the parentheses below the activity links refer to mass flow in grams and energy requirement in million joules, respectively, for the corresponding steps in the manufacture of one Size 6 meat tray.

* Energy content of resource.

Figure 11. Reference Materials Trajectory

would drop to 5.8 x 10^6 joules with one reuse, to 4.9 x 10^6 joules with two reuses, and to 4.5 x 10^6 joules with three reuses. Similarly, although a single use of plastic bags requires more energy than paperbags, the two become comparable if more durable polyethylene bags are reused once. These results are summarized in Table 5. Using this information in conjunction with RMS with sufficient disaggregation in the containers and packaging sector, the peturbation technique can be applied in a rather straightforward manner to assess the full materials system implications in terms of energy and source requirements arising from the substitution measures considered here.

CONCLUSIONS

The Role of Optimization

The optimization of a system, such as the materials system or a subsector, implies a selection of individual processes or material forms among alternatives on the basis of the specific optimization criteria, usually minimum cost. In optimization analyses, the critical factors are the characterization of competing options and the specifications of the objectives to be optimized. Optimization is generally considered to be a normative process indicating the preferred decision or direction--what should be done--given the assumptions that are in the analysis. In some cases, particularly where uncertainties are not dominant and objectives are relatively simple, the optimization process has some simulation capability as well.

In more complex situations, optimal solutions may be tested for "robustness" against the major uncertainties--e.g., how sensitive is the normative solution to changes in the state of the world or the other relevant factors. Techniques of multi-objective analysis have been developed to deal with complex objective functions that include several attributes.

In practical applications, particularly where large uncertainties exist or objectives are complex, there has been considerable criticism of the over simplification necessary to apply optimization techniques. Some of this criticism is deserved; however, alternative techniques really do not handle such problems well either. Solutions determined by optimization techniques can provide useful insights into the feasible limits of, say, cost minimization or profit maximization under idealized circumstances. While such solutions may not be practical, they provide a well characterized benchmark against which practical compromises may be gauged. In this sense, a normative optimization is similar to the idealized Carnot-cycle limits in thermodynamics. While the presence of irreversibility induced by friction and energy exchange complicates the design and analysis of real processes, it has still proven quite useful to consider the ideal Carnot-cycle limits as a benchmark for comparison.

Table 5. Energy Requirements for Typical Containers and Packaging

		Raw material reqirements Per unit Product			Energy of Manufacture		Energy content of of raw materials	Total Energy
Container/packaging (product) type	Unit Weight (grams)	Natural gas (grams)	Crude Oil (grams)	Wood (grams)	per unit Product (10^6 Joules)	10^6 Joules/ Kg. of Product	per unit product (10^6 Joules)	per unit product (10^6 Joules)
Half-gallon Milk Container								
Polyethylene	54	22	55	--	8.4 5.2*	155.0 96.0*	3.5	11.9
Paper	64	--	--	130	6.4 3.0*	100.0 47.0	1.5	7.9
Size 5 Meat Tray								
Polystryene plastic	6.7	2.3	7.2	--	0.9	127.0	0.4	1.30
Wood pulp	20	--	--	30	0.94	47.0	0.36	1.30
Flexible Container (bag or Sack)								
Polyethylene	18	6	16	---	2.9	≈160.0	1.0	3.9
Kraft paper	36	--	--	70	1.7	≈48.0	0.8	2.5

*These values exclude the energy required for filling the containers.

The major optimization techniques that are employed in resource allocation, process selection, and substitution analysis include mathematical programming (linear and nonlinear), dynamic programming, and optimal control theory. Linear programming has been used extensively in view of the ability to solve very large sets of equations and to capture extensive technical detail on the structure of the system being analyzed. A major strength of the technique is to combine technical and economic factors. The structural equations of an optimization model are usually developed in physical terms--e.g., energy and material balances, equipment utilization, and manpower requirements; while the objective functions generally involve such economic parameters as cost minimization or profit maximization. Solutions to resource allocation or process selection problems formulated along these lines reveal the marginal value, or cost, of constrained resources. Examples of this class of model include the work of Pilati (1979) and Sparrow (1978).

In application to the analysis of substitution within the materials system, optimization techniques have proven to have their greatest strength in the analysis of resource inputs and process selection to produce a well characterized process. Optimization techniques have been applied successfully to steel, aluminum, paper, and other materials sectors to analyze the effect of changing costs of resources, energy, and labor on the production plant or process to produce these materials. There has been less success in applying optimization techniques to end-use substitution of materials, due largely to the many attributes that are involved in the materials selection process. Questions of cost, durability, and esthetics are all involved in a very complex decision problem. Optimization is practiced implicitly in the product design as, in the case of the automobile, weight and cost objectives are specified and material substitution options are explored at the component level.

A more fruitful area in materials utilization and substitution for the more formal optimization modeling approaches is the selection of materials to achieve life-cycle cost minimization. The materials system is full of examples where minimum first cost has dominated the materials selection process. This has sometimes led to very high operating and maintenance costs that are often borne by groups other than those benefiting from the initial selection. This is a question of good design practice as well as public policy. Increased emphasis must be placed on life-cycle cost minimization, taking operating, maintenance, and disposal costs into account along with first cost.

REFERENCES

Aluminum Statistical Review. 1977. The Aluminum Association, Inc., 818 Connecticut Avenue, N.W., Washington, D.C. 20006.

Annual Statistical Report. 1978. Iron and Steel Institute, 1000 Sixteenth Street, N.W., Washington, D.C. 20036.

Ayres, R. V., and A. V. Kneese. 1969. Production, Consumption and Externalities: American Economic Review.

Bhagat, N., et. al. 1976. Reference Materials System: A Sourcebook for Material Assessment, Brookhaven National Laboratory, BNL 50609.

Hannon, M. B. 1972. System Energy and Recycling, A Study of the Beverage Industry. University of Illinois, Urbana.

Hoffman, K. C. 1975. Methodology of Technology Analysis with Application to Energy Assessment, American Society of Mechanical Engineers, Paper 75-Wa/TS-8. New York.

Hunt, R. G., and R. O. Welch. November 1974. Resource and Environmental Profile Analysis of Plastic and Nonplastic Containers, Midwest Research Institute, Kansas City, Missouri: MRI Project No. 3719-D.

Kearney, D. 1977. Policy Issues Involved in the Use of Biomass as an Energy Source: Conference on Energy Development. Dublin, Ireland.

Makino, H., and R. S., Berry. June 1973. Consumer Goods, A Thermodynamic Analysis of Packaging, Transport, and Storage: Illinois Institute of Environmental Quality.

Pilati, D. A., W. Marcuse, and F. T. Sparrow. 1979. Process Modeling of Energy-Intensive Industries, paper presented at XXIV International TIMS Meeting, Honolulu, Hawaii.

Sparrow, F. T. 1978. The Iron and Steel Industry Process Model, Brookhaven National Laboratory.

Statistical Abstract of the United States. 1978. U.S. Department of Commerce, Bureau of the Census.

RECENT ADVANCES IN ECONOMETRIC ESTIMATION OF MATERIALS SUBSTITUTION

Margaret E. Slade
Bureau of Economics
Federal Trade Commission
Washington, D.C.

INTRODUCTION

The 18th century classical economists, particularly Malthus and Ricardo, were very aware of the role of land in the production process. To them, the principal inputs to production were land and labor. Much of their focus was on agricultural land, which more broadly defined can be taken to encompass all natural-resource inputs, nonrenewable as well as renewable. As the Western world became more industrialized, however, the interest in natural resources was replaced by an emphasis on capital. The neoclassical production function of the 19th century was a function solely of capital and labor, and, until recently, most aggregate production functions that have been estimated were of the value-added variety, with intermediate inputs such as energy and materials ignored. However, in the 1970's several developments led to a renewed interest in the role of natural resources in production. On the political side, the Arab oil embargo, the Organization of Petroleum Exporting Countries' price increases, and the United States' growing dependence on imports of many minerals focused attention on the possible scarcity of nonrenewable energy and minerals. And on the theoretical side, the derivation of more sophisticated techniques for estimating production and cost functions made it possible to assess factor substitution possibilities with greater accuracy.

Interest in the possibility of substituting materials for other inputs to the production process is motivated by several concerns. Perhaps the most important ones are those pertaining to short-run supply disruptions and to long-run exhaustion of depletable natural resources. In the short run, most industrial nations, as net importers of many minerals, are vulnerable to cartel action and embargoes by producing countries. They therefore need to know what substitutes exist for commodities that are temporarily in short supply and how long it takes to substitute an available commodity for one whose price has increased or that cannot be obtained. In the long run, as the stocks of natural resources are depleted, it becomes essential to assess the possibility of substituting reproducible inputs such as capital and labor for scarce inputs such as nonrenewable energy and minerals. Econometric models have been constructed and estimated to help us understand the process of substitution and to aid in assessing the importance of both these issues.

In this paper, some recent advances in econometric estimation of material substitution possibilities are discussed. The organization of the paper is as follows:

- The relationship between separable production functions and substitutability is developed, flexible functional forms are defined, and the notion of duality is discussed.

- Some of the earlier estimates of elasticities of substitution between materials and other inputs that were obtained from flexible production-function models are reviewed and more recent modifications to the earlier models are described.

- Areas for further research are suggested.

SEPARABILITY, FLEXIBILITY, AND DUALITY

Materials are used primarily as intermediate inputs to the production of final goods. Therefore, the demand for materials is derived, that is, it is induced by the demand for the products that materials produced. To model substitution, we must therefore model the way in which materials are used in production. The econometric approach to modeling production is to estimate production functions that show the quantity of output that can be obtained from given input quantities and how the level of output changes as the quantities of inputs change. With neoclassical production function, substitution possibilities are assumed to be smooth and continuous (i.e., it is assumed that small changes in relative factor prices can lead to small changes in relative factor usage). A neoclassical production function can therefore be thought of as a smooth approximation to an engineering production function with a large number of production processes. Some of the conceptual problems that arise in modeling production functions are discussed briefly below.

In order to estimate production functions, we must aggregate inputs. For example, "labor" is not homogeneous but consists of unskilled, clerical, technical, professional, and many other classes of laborers. Each of these classes is in turn an aggregate of many subclasses. The way in which inputs can be aggregated relates directly to the possibility of substitution. To understand this relationship, some definitions are in order.

The elasticity of substitution between two factors of production is the percent change in factor proportions due to a percent change in their marginal rate of technical substitution (i.e., in the slope of the isoquant relating the two factors). If firms minimize cost, they will equate the marginal rate of technical substitution between factors to the factor price ratio. Therefore, the elasticity of substitution measures the percent change in factor usage ratio due to a percent change in the factor price ratio. If the elasticity of substitution is infinite, the inputs are perfect substitutes, whereas if the elasticity is zero, no substitution is possible.[1] If the elasticity of substitution is negative, the two inputs are said to be complements. That is, instead of being replacements for one another in production, the inputs must be used together and an increase in the price of one will reduce the demand for the other.[2] Most recent econometric work on materials substitution has focused on estimating elasticities of substitution between materials as an aggregate and other input aggregates, or between different materials.

A production function is said to be strongly separable if substitution between inputs in two groups or classes of inputs does not depend on the quantity of inputs in other groups. Separability is an important concept because, if production functions are not separable, we cannot simply look at prices and availability of materials to determine their use but must also know about the capital stock in place and the availability of energy and laborers before we can assess substitutability. If input groups are separable, the estimation of production functions can be sequential. That is, we can determine the optimal (least-cost) mix of material inputs, the optimal mix of energy inputs, and so forth, and, holding these within-group proportions constant, we can determine the optimal mix of materials and energy in an overall production function. If no partitioning of inputs is possible, we must estimate production functions with a very large number of arguments.

Berndt and Christensen (1973) showed that strong separability is equivalent to the assumption that all elasticities of substitution between inputs from different groups are both equal and constant.[3] Strong separability therefore places a priority restrictions on substitution possibilities. It is these restrictions that have generated interest in so-called flexible functional forms.[4]

---

1 With a Leontief (fixed coefficient) production function, the elasticity of substitution is zero, whereas with a Cobb-Douglass function, it is identically one.

2 Complements are usually defined in terms of their cross-elasticity of demand, not their elasticity of substitution. However the Allen partial elasticity of substitution, which measures the response of derived demand to an input price change, holding output and all other input prices fixed, is equal to the cross-elasticity of demand divided by the input's cost share. Because the cost share is always positive, the signs of the cross-elasticity of demand and the Allen partial elasticity of substitution are the same. Therefore, the two definitions are equivalent if the Allen partial elasticity of substitution is used.

3 It is the Allen partial elasticities of substitution that are equal and constant.

4 Flexible forms have not eliminated the separability assumption (i.e., no matter how much we disaggregate, some separability assumption is made in the choice of input groups). However, flexible forms have highlighted the importance of such assumptions.

A functional form is said to be flexible if it satisfies two conditions:

(1.) it places no a priori restrictions on the elasticities of substitution, and

(2.) it provides a local second-order approximation to an arbitrary twice-differential function.[5]

Several commonly used flexible forms are: the transcendental logarithmic (translog) (Christensen, Jorgenson, and Lau, 1971), the generalized Leontif (Diewert, 1971), and the generalized Cobb-Douglass (Diewert, 1973).

Elasticities of substitution are not always estimated from production functions. Often it is easier to estimate the cost function, which relates minimum production cost to input prices and output level, than to estimate the production function. If certain restrictions are placed on the cost function, duality theory assures us that a well-defined production function exists (Shephard, 1953). The dual of a particular cost function such as a translog is not necessarily a translog production function. However, the technologies of input substitution can be determined by analyzing the cost function alone. The use of duality theory in the derivation of production and cost functions has revolutionized production modeling in the last decade.

PRODUCTION-FUNCTION MODELS

First-Generation Models

Most of the earlier estimates of elasticities of substitution between materials and other input groups were obtained using a four-input translog production or cost function. The four-factor cost function is quadratic in the logarithms of the prices of the four inputs--capital (K), labor (L), energy (E), and materials (M).[6]

Generally, cost-minimizing input-demand equations are derived from the cost function (using Shephard's lemma) and estimated jointly with the cost function.[7] The elasticities of substitution, which are not constrained to be constant but vary with the input's cost share (i.e., with the ratio of expenditure on the input to total expenditure), are functionally related to the cost-function parameters and can therefore be calculated from estimates of these parameters. In addition, the own-and cross-price elasticities of demand are analytically related to the elasticities of substitution.

---

5 A second-order approximation is one that ignores higher order terms in a Taylor-series expansion of the function.

6 These inputs are themselves aggregates, often constructed using a Divisia index.

7 For a discussion of Shephard's lemma, see Diewert (1974).

Examples of empirical estimates of elasticities of substitution between materials and other input groups using translog production or cost functions include Hudson and Jorgenson (1975),[8] Berndt and Wood (1975), and Humphrey and Moroney (1975).[9] The results in these papers cast doubt on previous empirical studies of investment demand and capital-labor substitution which used a value-added (capital-labor or KL) specification. For example, Berndt and Wood (1975) tested the hypothesis that capital and labor are separable from energy and materials and found that the parameter restrictions that would imply separability were not satisfied. Therefore, at least a four-input specification is warranted.

These earlier studies, while an improvement on the typical value-added approach, do not help answer some of the most interesting questions about materials substitution. As noted in the introduction, much of the attention given to materials substitution stems from interest in either the short-run problem of supply disruptions (i.e., being cut off from imports of raw materials) or in the long-run problem of running out of exhaustible resources. Because the early capital-labor-energy-materials (KLEM) models were static, they could not tell us how long the substitution process takes (i.e., how quickly we can substitute away from unavailable materials during periods of temporary supply disruptions).[10] Many of the later models described in the next section attempt to remedy this problem by introducing dynamic adjustment to change conditions.

The long-run problem of substituting reproducible inputs for depletable natural resources is much more difficult to analyze in an econometric framework. We need to know the extent of substitution that is possible at very low levels of use of one or more inputs, and existing data does not generally provide such observations. As Dasgupta and Heal (1979) point out, the assumption that the elasticity of substitution is independent of the capital-resource ratio is a treacherous one to make. Therefore, econometric models tell us little about the ease of adjusting to long-run natural-resource scarcity and should not be used to assess the impact of running out of critical materials.

Second-Generation Models

Second-generation models will be classified as those that attempt to remedy two of the problems encountered with the earlier specifications. The problems are those related to dynamic adjustment and to aggregation.

Dynamic Models

There are many reasons why, in the real world, substitution takes time. For example, if the price of an input goes up, it may have to remain at the higher level for a considerable time before purchasers expect the higher price

8 The Hudson-Jorgenson model consists of KLEM submodels for nine industrial sectors.

9 Humphrey and Moroney used a three-input (KLM) model for twelve product groups.

10 A static model describes a system in a state of equilibrium but does not specify the adjustment to equilibrium after the system has been perturbed.

to persist and decide to look for a substitute. Even after the desirability of substitution is perceived, changing factor proportions may involve planning, issuing new contracts, and altering capital equipment and is therefore neither instantaneous nor costless. We can thus expect factor markets to be in disequilibrium (i.e., we can expect delays in adjusting to optimal factor proportions after a change in factor-price ratios) and should attempt to incorporate the adjustment process into models of input choice.

There are many methods of introducing dynamic adjustment into models of factor substitution. For example, Berndt, Fuss, and Waverman (1978) proposed two dynamic approaches. First, they incorporated Koyck adjustment matrices into a static system of factor demand equations, and second, they introduced costs of adjustment into the long-run optimization process. Denny, Fuss, and Waverman (forthcoming) constructed a model where firms minimize the present value of future costs, subject to internal rising marginal costs of adjusting their capital stock. Slade (forthcoming) assumed that firms minimized costs with respect to expected factor prices, and added generalized price-expectations equations to a model of input selection. And Brown and Christensen (forthcoming) employed an ingenious method to obtain long-run elasticities of substitution from a short-run cost function. In the short-run, equilibrium is presumed for the variable factors, conditional on given levels for the fixed factors, whereas in the long run, all inputs (variable and fixed) are in equilibrium. The evidence from these different approaches indicates that the assumption of equilibrium in all factor markets is rarely justified and that static models often yield incorrect estimates of elasticities of substitution. However, one drawback to dynamic models is that the method of introducing the adjustment process affects the estimates of elasticities of substitution obtained. Considerable care should therefore be taken in selecting an appropriate model (i.e., one that closely corresponds to the market being modeled).

Aggregation

It was noted that production functions with less than four inputs may lead to erroneous estimates of substitution possibilities. It remains to be seen, however, if four inputs are sufficient. Questions arise such as:

- Is it meaningful to construct a single aggregate of raw materials used at the economy level? and
- If it is meaningful to construct such an aggregate, how can we measure it?

There are some of the issues that Lau (1979) discussed in a paper dealing with the measurement of raw materials inputs at the plant, firm, industry, and economy level.

When we aggregate inputs into groups, we must choose a measure for the aggregate. A common procedure is to choose one characteristic (e.g., Btu value for fuels) to use as a weight.[11] However, this practice implicitly assumes that all fuels are perfect substitutes in production. The problem is more complex for raw materials because there is no obvious choice of a characteristic to use as a weight. It is therefore preferable to consider a veotor of characteristics of materials and construct hedonic price equations. Additional problems arise when aggregating inputs across plants and across industries. In this case, the hedonic price equation will depend on a vector of plant (or industry) characteristics in addition to the characteristics of materials. Further complications arise when aggregating across end-use sectors. Usually we are forced to make some simplifying assumptions--that prices of all raw materials move together or that quantities of raw materials are used in fixed proportions (Hicks and Leontief aggregation, respectively). Lau (1979) concluded his analysis of materials aggregation with the statement that it appears unlikely that, at the economy level, the cost function is separable in the prices of raw materials so that a meaningful index of aggregate raw materials can be constructed. His conclusion raises questions about the practice of estimating four-input KLEM production functions for the economy as a whole.[12]

Numerous empirical studies have demonstrated that the minimal requirements for consistent input aggregation are rarely found in actual data and that aggregation at any level is thus almost always improper in a theoretical sense.[13] The question then becomes, what sorts of errors do we make when we aggregate? In addition to the question of input aggregation, the question of technology aggregation is very important because it affects the interpretation of the production function as an approximation to a set of interrelated engineering activities. When technology aggregation takes place, even the aggregation of simple subprocesses, there is no guarantee that the engineering character of the subprocess will be reflected in the aggregate. In contrast to input aggregation, with technology aggregation, we have little or no theoretical basis to discriminate between aggregates. Kopp and Smith (forthcoming), in an empirical study of resource substitution under input and technology aggregations, concluded that aggregation of progressively more diverse inputs leads to a diminution of the statistical significance of the estimated elasticities of substitution but not to a reversal of their signs. The loss of statistical significance can be interpreted as a decrease in the power of the neoclassical econometric model to approximate the underlying engineering features. More seriously, they found that aggregation across technologies may lead to the inability of the neoclassical econometric model to discriminate between relationships of substitutability and complementarity (which is intuitively plausible because inputs can be substitutes in one technology and complements in another). In view of these results, it seems imperative that substitution be measured at a much more disaggregate level than has been common practice in the past.

---

11 In constructing indices, we can form a quantity of a price aggregate. The Btu is, or course, a quantity aggregate. Most recent studies using cost functions employ a price aggregate. The price aggregate avoids the need for a consistent quantity index by relying on competitive-market models (an assumption that is often violated in practice).

12 Similar problems arise in measuring a capital aggregate.

13 Minimal requirements for consistent input aggregation are that the input groups be homothetically weakly separable.

One result of the realization that disaggregation is imperative has been the construction of greatly expanded data bases. For example, Jorgenson and Fraumeni (forthcoming) compiled capital, labor, energy, and materials inputs to thirty-six industrial sectors of the U.S. economy, annually for the period 1958-1974. They also constructed indices of the sectoral outputs and all four sectoral inputs for the same period. Finally, they computed indices of sectoral rates of technical change. These data enabled them to characterize substitution possibilities among inputs and changes in these possibilities over time at a much more disaggregated level than had been previously attempted.

SUGGESTIONS FOR FURTHER RESEARCH

There are many directions in which econometric methods of modeling materials substitution can move, only two of which will be mentioned here: the formation of price expectations and rationing.

Rational Expectations

Typically, models of materials substitution either ignore price expectations or assume that expectations are formed in an adaptive fashion (i.e., that expected prices are some function of past observed prices). However, it may be more appropriate to incorporate the notion of rational expectations into dynamic models. The rational expectations hypothesis first formulated by Muth (1961) is that the appropriate forecast of a variable is the expected value of that variable, conditional on all of the information available at the time of the forecast (not just the past history of the variable). If firms are neglecting some information in forming price forecasts, their forecasts will be wrong and speculators will have an incentive to intervene in the market. For example, if firms that purchase copper forecast a price for copper that is too low, speculators will purchase future contracts for copper and the forward price of copper will rise. Adaptive expectations, therefore, unlike rational expectations, are not self-confirming. One contribution to empirical work of the literature on rational expectations is the attempt to disentangle the structural parameters of the (production or cost) function from the (price) expectations equation. Lucas (1976) pointed out that macro models constructed using adaptive expectations cannot be used for policy analysis because the way in which people respond to changed policy variables depends on the source of the change. This idea can also be important in modeling materials markets. For example, it is not realistic to assume that purchasers of copper will react to higher copper prices that result from political disruptions in Zaire (that they perceive as temporary) in the same way as they will react to higher copper prices that result from environmental controls on smelter emmissions (that they perceive as permanent). In addition, purchasers of materials react more rapidly to information that they perceive as accurate than to information that they perceive as fuzzy or uncertain. If the source of a change is a determinant of the speed of adjustment to that change, the distributed lag parameters in a model should be endogenously determined (i.e., they should change with changes in the exogenous variables), especially if the model will be used for policy analysis and forecasting. The incorporation of rational expectations could therefore be a valuable extension to dynamic models of materials substitution.

Rationing

The notion of disequilibrium in factor markets was mentioned. It was assumed that firms were not equating factor-price ratios to their marginal rate of technical substitution because there were costs associated with changing factor proportions. However, there are other ways in which disequilibrium can occur. In many metals markets, particularly those where prices are administered, rationing is a common phenomenon (i.e., producers do not allow prices to rise to the level that would equate supply and demand in the short run but, instead, selectively ration). Rationing is sometimes explained as an attempt to prevent the long-run substitution away from a commodity that might be triggered by a very high price for that commodity (see, for example, McNicol, 1975). When rationing occurs, the assumption that consumers are able to obtain materials at their market price is violated. MacKinnon and Olewiler (1980), in an empirical study of the U.S. copper market, showed that estimated own-price elasticities of demand are significantly different when rationing is considered from those obtained when equilibrium is assumed. Presumably, estimates of elasticities of substitution would also be affected by the consideration of nonprice rationing. Therefore, in markets where rationing is prevalent, disequilibrium models should provide better estimates of substitution possibilities than do the present equilibrium models.

None of the models discussed in this paper can be considered the ultimate econometric model of materials substitution. Although expectations and rationing have been focused on as topics worthy of future research efforts, there is room for improvement and further refinement in many areas. Because materials substitution will remain a topic of vital importance and empirical relevance, much progress in modeling should occur in the next few years and we can expect third-generation models to emerge that will supersede the first-and second-generation efforts that are discussed here.

REFERENCES

Berndt, E. R., and L. R. Christensen. 1973. The Internal Structure of Functional Relationships; Separability, Substitution, and Aggregation, Review of Economic Studies, Vol. 40.

Berndt. E. R., M. A. Fuss and L. Waverman. 1978. Factor Markets in General Disequilibrium: Dynamic Models of the Industrial Demand for Energy. Workshop on Energy Supply and Demand, International Energy Agency. Paris.

Berndt, E. R., and D. O. Wood. 1975. Technology, Prices, and the Derived Demand for Energy. The Review of Economics and Statistics, Vol. 57, No. 3.

Brown, R. S., L. R. Christensen, 1947-1974. An Assessment of Substitution Possibilities Among Factor Inputs in U.S. Agriculture. Modeling and Measuring National Resource Substitution, E.R. Berndt and B. Field, Editors. M.I.T. Press, Cambridge, Massachusetts. Forthcoming.

Christensen, L. R., D. W. Jorgenson, and L. R. Lau. 1971. Conjugate Quality and the Transcental Logarithmic Function, Econometrica, Vol. 39, No. 2.

Dasgupta, P., and G. Heal. 1979. Economics of Exhaustible Resources, Cambridge University Press. Cambridge.

Denny M., M. A. Fuss, and L. Waverman. The Substitution Possibilities for Energy: Evidence from U.S. and Canadian Manufacturing Industries, in Modeling and Measuring Natural Resource Substitution. E. R. Berndt and B. Field, Editors, M.I.T. Press, Cambridge, Massachusetts. Forthcoming.

Diewert, W. E. 1971. An Application of the Shephard Duality Theorem: A Generalized Leontief Production Function. Journal of Political Economy. Vol. 79.

Diewert, W. E., 1973. Separability and the Generalized Cobb-Douglass Utility Function, Department of Manpower and Immigration. Mimeo, Ottawa.

Diewert, W. E. 1974. Applications of Duality Theory, in Frontiers of Quantitative Economics, M. Intrilligator and D. Kendrick, Editors. North-Holland, Amsterdam.

Hudson, E. A., and D. W. Jorgenson. 1975. U.S. Energy Policy and Economic Growth, 1975-2000. The Bell Journal of Economics and Management Science, Vol. 5, No. 2.

Humphrey, D. B., and J. R. Moroney. 1975. Substitution Among Capital, Labor, and Natural Resource Products in American Manufacturing. Journal of Political Economy, Vol. 83, No. 1.

Jorgensen, D. W., and B. M. Fraumeni. Substitution and Technical Change in Production, in Modeling and Measuring Resource Substitution. E. R. Berndt and B. Field Editors, M.I.T. Press, Cambridge, Massachusetts. Forthcoming.

Kopp, R. J., and V. K. Smith. Measuring Prospects for Resource Substitution Under Input and Technology Aggregations, in Modeling and Measuring Natural Resource Substitution. E. R. Berndt and B. Field, Editors. M.I.T. Press, Cambridge, Massachusetts. Forthcoming

Lau, L. J., 1979. On the Measurement of Raw Materials Inputs. Paper prepared for a conference sponsored by Resources for the Future.

Lucas, R. E. 1976. Econometric Policy Evaluation; A Critique, in The Phillips Curve and Labor Markets. Brunner and Meltzer, Editors. Carnegie Rochester Series in Public Policy 1. North-Holland, Amsterdam.

McKinnon, J. R., and N. D. Olewiler. 1980. Disequilibrium Estimation of the Demand for Copper. The Bell Journal or Economics, Vol. 11, No. 1.

McNicol, D. L. 1975. The Two Price Systems for Copper. The Bell Journal of Economics, Vol. 6, No. 1.

Muth, J. R. 1961. Rational Expectations and the Theory of Price Movements, Econometrica, Vol. 29, No. 3.

Shephard, R. W. 1953. Cost and Production Functions. Princeton University Press, Princeton, N.J.

Slade, M. E. An Econometric Model of the U.S. Copper Industry: How Cost Changes Affect Substitution and Recycling, Garland Publishing, Inc., N.Y. Forthcoming.

MODELING THE DYNAMICS OF SUBSTITUTION

James M. Lyneis
Pugh Roberts Associates, Inc.
Cambridge, Massachusetts

The metals industries are emerging from another period of short supply and escalating prices. Some of the latest shortage was no doubt caused by normal industry cycles. But many observers see longer-term trends contributing to the shortages and portending greater future difficulties. These trends include depletion of ore bodies in developed countries and consequent concentration of supply in less developed countries, environmental costs and restrictions on supply, political instability in key supplying countries, and increased competition for available supply because of industrialization in developing countries.

These longer-term trends offer the potential for the formation of effective cartels, supply disruptions, and extortionate price-fixing. Serious economic disruption could occur unless governments, producers, and consumers take appropriate actions to protect themselves from dependence on foreign sources of critical metals.

But the need for action, and the appropriate actions, depend importantly on the potential for substitution. To the extent that demand can be shifted to alternative materials, the effects of supply disruptions will be mitigated. As a result, buffer stocks and research can be focused on those segments of demand for which substitution is impossible or impractical.

How will supply and demand evolve under the likely trends for economic growth, technological change, and other key factors? How quickly will substitution occur in the case of supply disruptions or rapid price escalation? How large a buffer stock is necessary during the transition and for the unsubstitutable demand? What are the impacts on the markets of competiting materials? An understanding of the dynamics of substitution is essential to the development of effective action.

Dynamics of Substitution

Substitution is a dynamic process--that is, it occurs over time as a consequence of changing conditions. As in all dynamic processes, feedback, inertia, and delays play important roles.

Feedback occurs whenever an action generates a response, and the response then elicits further action, as illustrated in Figure 12. In the process of substitution, feedback occurs at the macro- and the micro-levels. At the macro-level, shown in Figure 13, an increase in price encourages substitution, which then reduces demand. The reduction in demand, other things remaining equal, decreases price, thereby slowing or reversing the substitution. Figure 13 illustrates what might be called "short-run substitution," where competing materials are readily available and accepted. Figure 14 expands Figure 13 to incorporate "long-run substitution." An increase in prices causes investment in research and development. Eventually, alternative technologies are developed which in turn increase substitution, reduce demand, and decrease price. The principal characteristic of feedback is that actions today influence what happens in the future. For example, substitution today may reduce substitution in the future because prices are lower. Feedback is therefore an essential element of dynamic behavior.

Feedback also influences substitution on the micro-level, as illustrated in Figure 15. Market share moves over time toward the market share indicated by relative prices. The speed of the movement is governed by resistance to change. When market share, say of aluminum in electrical uses, is low, resistance to change is high and substitution occurs only slowly. But as market share builds, resistance to change falls, thereby increasing the rate of substitution and producing a "bandwagon" effect. This micro-level feedback affects the speed with which substitution occurs. Inertia reflects the tendency to continue doing what you have been doing in the past and is represented by levels (accumulations or stocks) in a model. Figure 16 illustrates how inertia is a part of the micro-level feedback process. Demand for copper is a function of desired materials use and current market share of copper. Current market share is a level which adjusts over time toward market share indicated by relative price. Users will continue to employ copper until inertia is overcome and current market share adjusts to indicated market share.

Delays are the final element in dynamic processes. Delays govern the amount of inertia in a system. Several delays are important in substitution:

1. The length of time over which price changes must persist before substitution is warranted

2. Delays in developing alternative technologies

3. Delays in adopting the competing materials, as governed by resistance to change and minimum times to convert.

The costs of conversion and the ratio of material costs to total product cost influence the length of these delays.

In summary, substitution is a process which occurs over time in response in changing conditions. Some of these conditions may be external to the substitution process itself--for example, supply shortages which cause an increase in price. But others are the result of feedbacks from past changes: prior substitution affects current prices which in turn affect current substitution. Inertia and delays influence the speed with which substitution occurs. An accurate forecast requires that these processes be represented, as they critically affect behavior.

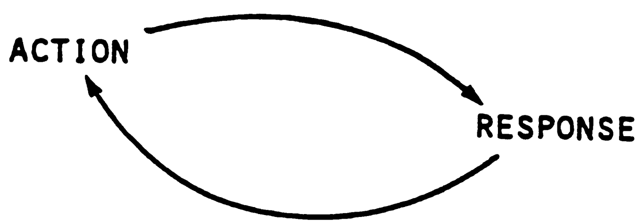


Figure 12. Feedback

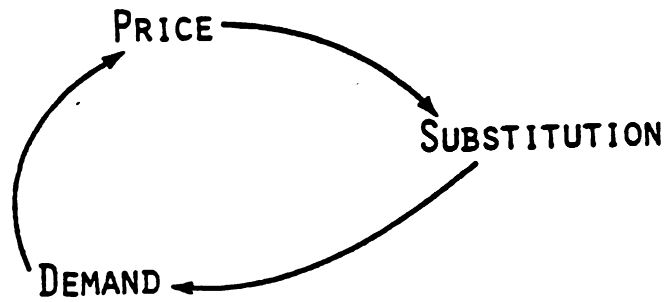


Figure 13. Price Feedback

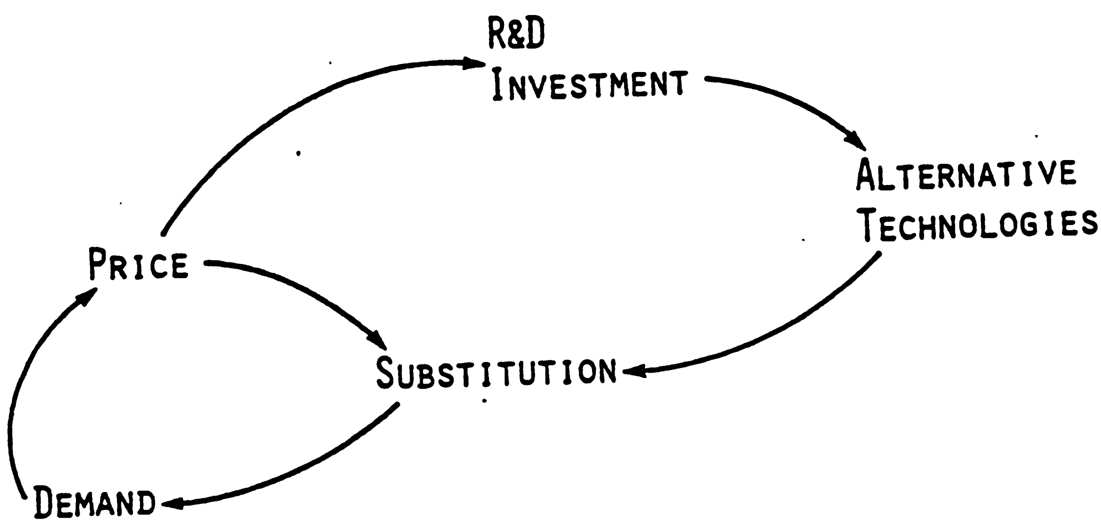


Figure 14. Macro Level Feedback

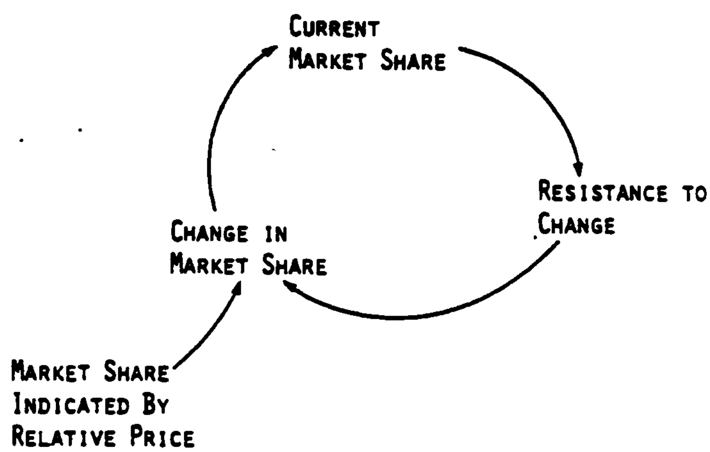


Figure 15. Micro Level Feedback

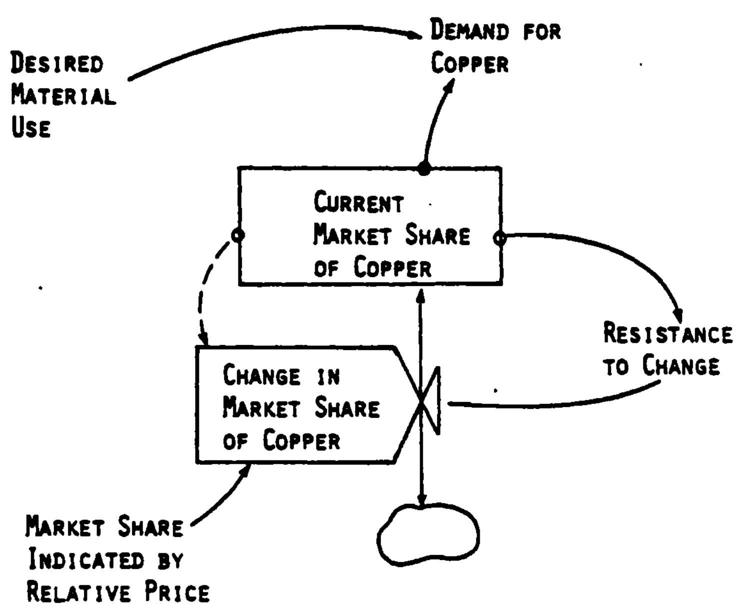


Figure 16. Inertia in Dynamic Processes

A Generic Model of Substitution

The dynamics of substitution cannot generally be separated from the larger dynamics of supply and demand, of which it is an integral part. Therefore, any model of "substitution" must also incorporate some of the macro-feedbacks affecting supply and demand. These issues are discussed in the first subsection below. The detailed process of substitution, generally represented on an end-use basis, is then discussed in the following subsection.

Macro-Feedbacks Influencing Substitution

Figure 17 illustrates the macro-feedbacks which affect the dynamics of substitution. Which of these feedbacks are modeled, and in how much detail, depends on the purpose of the model. Short-term models used for forecasting purposes may take the price of competing materials and technologies as given; mid-term models for policy purposes--for example, determining the size of buffer stocks--may want to consider the effects of disruptions in the supply of one metal on other metals, and hence must consider these feedbacks; longer-term models may want to consider the likely evolution of alternative technologies.

It is therefore impossible to define a "generic" model of these macro-level feedbacks. Figure 17 merely serves to highlight the range of possibilities which must be considered:

(1) short-run effect of price;

(2) effect of substitution on price of competing materials;

(3) price-supply; and

(4) development of alternative technologies, which may reduce the amount of material needed or introduce new competing mateials.

Detailed Substitution Process

Figure 18 shows the factors determining the specific amount of substitution. Because many of these factors are easier to pinpoint for particular end uses of a metal, the structure shown is often duplicated to determine substitution on an end-use basis.

The first characteristic of substitution illustrated in Figure 18 is that demand is split into substitutable and nonsubstitutable components, as governed by legal and technological constraints. This split is important from a policy point of view because, depending on the speed with which substitution occurs, only the nonsubstitutable demand need be covered in any supply disruption.

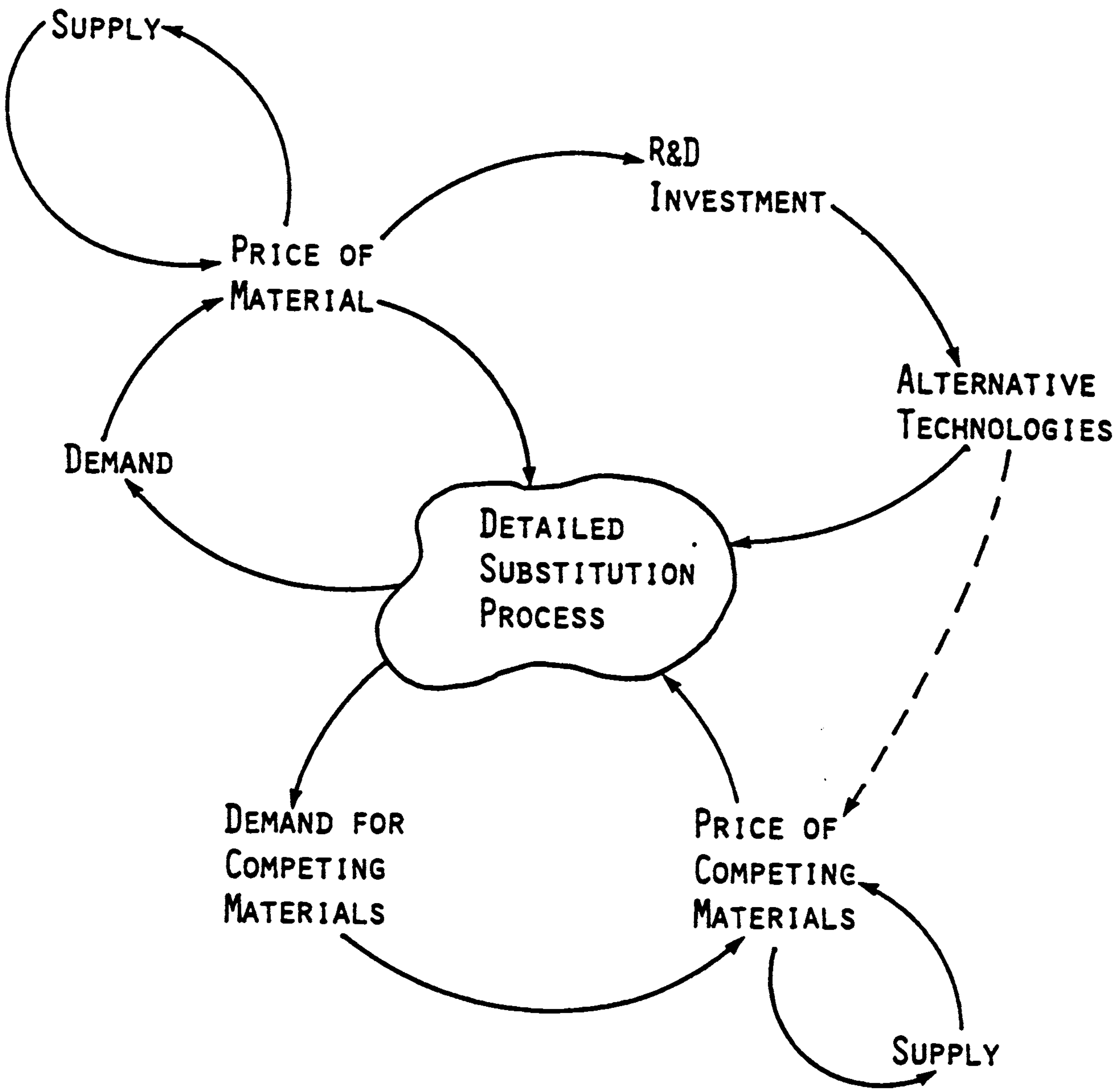


Figure 17. Macro-Feedbacks Influencing Substitution

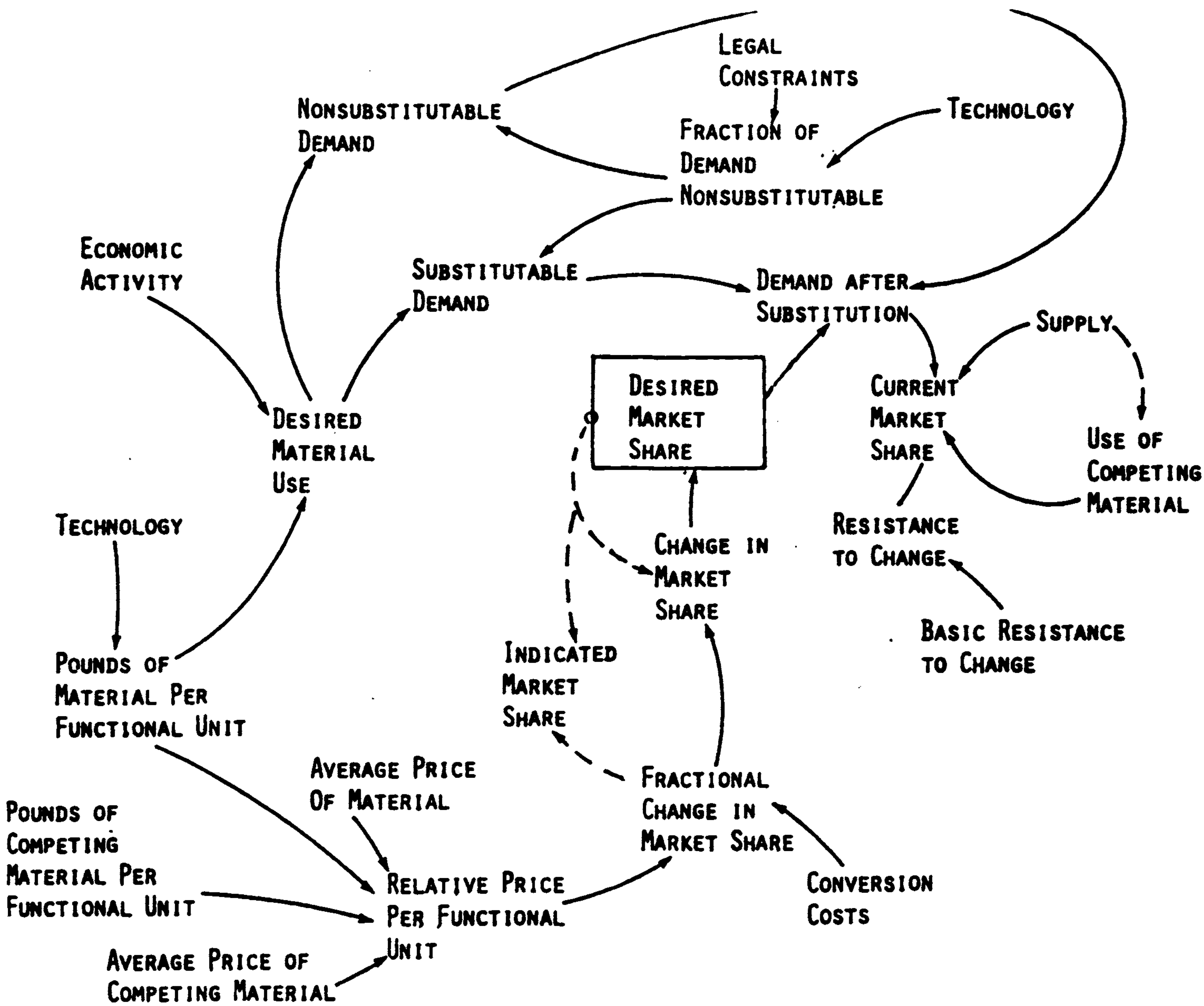


Figure 18. Detailed Substitution Process

Substitutable demand, together with desired market share, determines demand for the material. Supply shortages may, however, constrain substitution and slow actual gains in market share.

Resistance to change affects how quickly desired market share increases or decreases. Resistance to change is expressed in years required to make changes. Mathematically, it equals:

RTC = BRTC * ECMS

RTC--Resistance to Change (Years)

BRTC--Basic Resistance to Change (Years)

ECMS--Effect of Current Market Share (Dimensionless)

The basic resistance to change represents the reluctance to adopt something new over a tried and true product. It is a constant which may be different for different end uses. Effect of current market share acts to reduce the constant as market share increases or as adoption spreads. The effect is illustrated in Figure 19. When the substitution is just beginning and market share is low, consumers are wary of substituting. They generally prefer to wait until the market share grows somewhat and that application is tested before they move. This type of hesitancy leads to the "bandwagon" phenomenon which begins with slow growth, as some innovative users perceive a cost or technical advantage, then goes through a phase of fairly rapid growth as customers perceive a successful application, and finally goes through a leveling-off phase as maximum market share is achieved (also referred to as "S-shaped growth"). A minimum resistance reflects the "time required to make a conversion."

The functional change in market share depends on relative price per functional unit and on conversion costs, as illustrated in Figure 20. When relative prices, determined by the difference between material price per functional unit and competing material price per functional unit, are equal, fractional change in market share equals zero-and-consumers have no incentive to change so they stick with present materials. As prices fall relative to competing materials, fractional change increases until it reaches 1.0--all consumers of the competing material desire to switch to the material in question. Conversely, as prices rise relative to competing materials, fractional change falls until it reaches negative 1.0--all consumers of the material desire to switch to the competing material. (Recall that reluctance to change governs how quickly the change occurs.)

As the driving force behind change, price difference are compared to conversion costs spread over a payback period. When the ratio is 1.0 (or negative 1.0), the price difference is such that conversion costs will be recovered in the payback period. Not everyone switches material at this point, however. Some people are more risk averse than others and therefore require a shorter payback period. Other people may have higher conversion

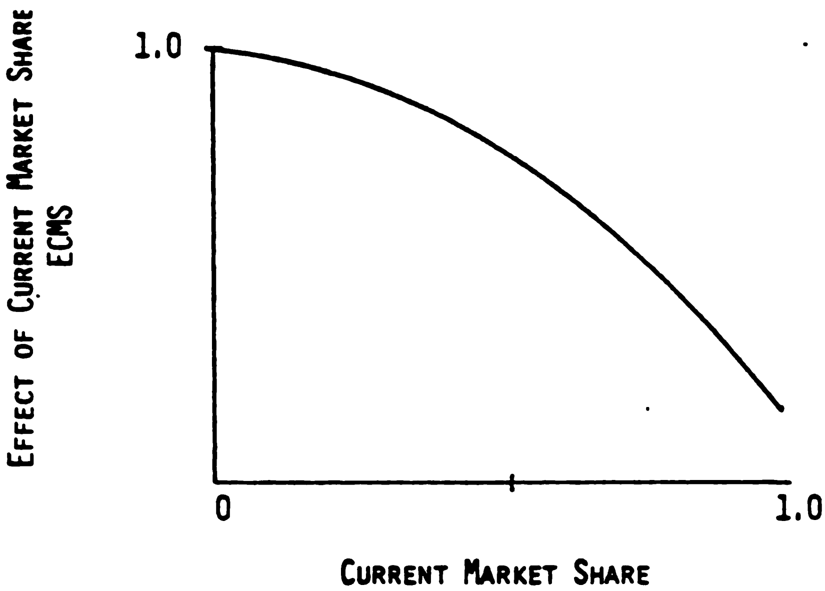


Figure 19. Effect of Current Market Share on Reluctance to Change.

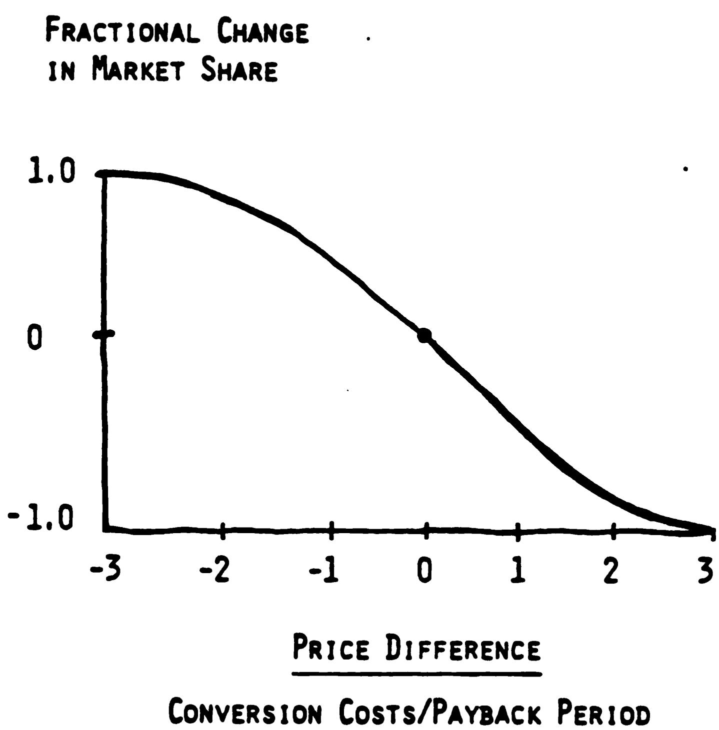


Figure 20. Fractional Change in Market Share

costs than the average used as an index and hence switch at a higher price difference. Finally, to the extent the model aggregates several end uses, conversion costs are likely to be dispersed about an average. For these reasons, complete conversion does not occur until price difference exceeds conversion costs by three times.

The price differential is based not on current price but on an average of prices. Such an averaging reflects the desire to assure that price changes are permanent before making a switch. The length of the averaging time probably depends on the cost of conversion. If costs are low, quick switches between materials can be justified. If costs are high, price changes must last longer in order to pay back conversion costs.

Several complicating factors can be introduced into the model of the substitution process. First, conversion costs might differ depending on the direction of the change. Second, more than one competing material can be considered, although this may not be important if the end-use disaggregation is detailed enough. And finally, availability or security of supply can be introduced as a reason to substitute materials.

In summary, the model of the substitution process captures a number of relevant considerations: legal and technical constraints making some demand unsubstitutable; resistance to change and bandwagon effects; breakeven costs based on a functional unit basis; conversion costs; and the need for a price change to be persistent to elicit substitution. By representing these factors, the model can more accurately portray the time behavior of the substitution process.

Model Applications

Pugh Roberts Associates, Inc., has successfully employed various versions of both the macro- and micro-feedback processes during a number of modeling studies over the past five years, including:

- An analysis of supply and demand for copper in the United States (for the National Science Foundation 1976).

- Forecasting price trends for copper (several proprietary clients, 1976-79).

- Analysis of supply and demand for soda ash (proprietary investment analysis).

- Analysis of worldwide demand for copper, aluminum, steel, nickel, and manganese over the next 20 years (proprietary, 1979-present).

ECONOMETRICS AND DEMAND MODELING

Bernard V. Reddy
James C. Burrows
Charles River Associates, Incorporated
Boston, Massachusetts

An analysis of the applicability of econometrics to issues of materials substitutions must begin with a clear understanding of what econometrics is. The first editor of Econometrica and cowinner of the first Noble prize awarded in economics, Ragnar Frisch (1933), defined econometrics as follows:

> Thus, econometrics is by no means the same as economic statistics. Nor is it identical with what we call general economic theory, although a considerable portion of this theory has a definitely quantitative character. Nor should econometrics be taken as synonomous with the application of mathematics to economics. Experience has shown that each of these three view-points, that of statistics, economic theory, and mathematics, is a necessary but not by itself a sufficient condition for a real understanding of the quantitative relations in modern economic life. It is the unification of all three that is powerful. And it is this unification that constitutes econometrics.

This definition indicates clearly that econometrics is more than trend analysis, more than the blind application of statistical tools to economic data. When estimating a demand equation the econometrician first must take care in specifying the equation so that it is consistent with economic theory and physical reality. Only then can the econometrician use the tools of mathematics and statistics to estimate the equation.

Many statistical techniques are grouped together under the heading of econometrics, but this paper will focus on least squares and its variants, because they are by far the most commonly applied statistical tools used in the estimation of models of demand.

Implicit Assumptions

Many of the assumptions implicit in econometric analysis are discussed in Haavelmo's (1944) seminal article. Constancy of the functional relationship and reversibility both are implicit in the use of regression equations.

Constant Relationships

A regression equation states that a given variable is a function of other variables. The general form of a linear regression equation is:

$$Y_t = b_O = \sum_i X_{it} b_i + e_{it} \qquad (1)$$

where:

Y_t = Dependent variable in period t;

X_{it} = The ith independent variable in period t;

b_i = Coefficient to be estimated for independent variable i;

e_t = The unobservable error in period t.*

*Assumed for convenience, is that observations of variables are made over time, the usual case with demand models for metals. Very few demand models for metals pool observations across states or countries.

The unknown parameters b_i are assumed to remain constant throughout the estimation period and forecast period. If one or more of the b_i are thought to be changing over time, the manner of change should be specified explicitly and incorporated into the equation, leaving an equation with a new set of unknown parameters that, in fact, will remain constant. This will be discussed in greater detail below in the section on equation specification. The implicit assumption of constancy in the function relating the exogenous variables to the endogenous variable means that a regression equation can be used to predict the future only if the future will be similar to the past. Sudden new developments, such as those caused by technological advances, cannot be predicted with econometric methods.

Reversibility

A corollary to the implicit assumption of constancy is reversibility. Suppose an exogenous variable changes but then returns to its original level. The regression equation states that the original change will lead to a change in the endogenous variable, but when the exogenous variable returns to its original level, so must the endogenous variable (assuming that all other variables have remained constant). This assumption is not always valid for metal markets, because some substititions are not reversible. For example, for many years ceramic magnets were cheaper than Alnico magnets, which contain substantial quantities of cobalt. However, many users of Alnico magnets did not bother to substitute ceramics for Alnico's, because the design costs of such substitutions were rather high. When cobalt prices increased to record levels in 1978 and 1979, many of these substitutions finally were made. Even if cobalt prices return to their pre-1978 levels, in constant dollars, these substitutions in many cases will not be reversed. A regression equation estimated for the demand for cobalt in Alnico's would recognize the recent decline in cobalt consumption in this category as a price effect, but the equation would also imply that cobalt would regain that market if the price fell. Blind estimation of such a regression equation would lead to misleading results.

Estimation Methods

Ordinary least squares generally is the estimation method chosen for regression equations, but in certain circumstances other methods are appropriate. Ordinary least squares is the most appropriate estimator, in a certain well-defined manner, if the error term in Equation 1 satisfies the follow conditions:

- The error terms are not correlated over time.
- The error terms over time have the same variance.
- The error terms are uncorrelated with the variables on the right side of the regression equation.

The causes of failure of one or more of these assumptions, and the appropriate methods correcting for them are discussed below.

Autocorrelation

If the error terms are correlated over time, autocorrelation is said to be present. Autocorrelation in and of itself is not a sufficient reason to avoid using ordinary least squares, although it can be a source of major difficulties if certain other problems accompany it. Without such other problems, autocorrelation will cause the parameter variances of ordinary least squares to be higher than the parameter variances of the most efficient estimation methods. Furthermore, the estimates of the parameter variances printed by a regression program will understate the true parameter variances. However, least squares will still yield unbiased estimates--that is, on the average (in the expectational sense) the coefficient estimates will be correct. In addition, the most efficient methods are known to be superior to least squares only in large samples. It is not clear that they are superior in an equation estimated with only 20 observations, as is often the case in annual metal models.

The appropriate estimation method in the presence of autocorrelation calls for transforming the data in a particular manner and applying least squares. The transformation may have to be repeated several times. If one of the variables appearing on the right side of the equation is the dependent variable in an earlier period, the existence of autocorrelation is a serious problem, and more sophisticated methods are called for in handling it.

Heteroscedasticity

Heteroscedasticity exists when the error terms have different variance in different time periods. This is likely to happen in many equations that describe market behavior over time, if the market is growing. The error terms may appear to have multiplicative, rather than additive, effects in a growing demand sector. Fortunately, heteroscedasticity is a problem that generally can be ignored. As with autocorrelation, heteroscedasticity causes least squares to have higher parameter variances than do more efficient methods, but least squares still is unbiased. Simple transformations of the equations can correct for heteroscedasticity. Estimating an equation in logs often serves to correct for it, as the logarithmic form assumes that the error terms have multiplicative effects.

Simultaneity and Related Problems

The most serious failures of the assumptions generally made for least squares concern the correlation between the error terms and the right-hand variables in an equation. If one or more variables are correlated with the error term, then all of the parameter estimates will be biased and inconsistent.

This assumption of no correlation usually fails because the variable in question is an endogenous variable in the same simultaneous model.

Consider a simple, two-equation model of a hypothetical market:

$$S + a_0 + a_1 P + a_2 CAP + e_1 \qquad (2)$$

and

$$D = b_0 - b_1 P + b_2 GNP + e_2, \qquad (3)$$

with the market-clearing identity D = S closing the model, where:

S = Supply.

D = Demand.

P = Price.

CAP = Production capacity.

The price is determined implicity by the market-clearing identity. Because of the market-clearing identity, we can set the supply equation equal to the demand equation and solve for price, which will be a function of both error terms, e_1 and e_2. As a result, price will be correlated with the error terms in both equations. The use of ordinary least squares will lead to biased, inconsistent parameter estimates. Simultaneous equation methods generally should be used to "purge" the correlation from the price variable in estimating the supply and demand equations.

However, simultaneous equation methods are demonstrably superior to ordinary least squares only if large samples are available. They are biased in small samples, although the bias disappears as the sample size increases. As a result, it is not clear that simultaneous equation methods are superior to least squares in samples with 30 or fewer observations.

Multicollinearity

Multicollinearity does not result from the failure of any of the error term assumptions in least squares, but many econometricians attempt to correct for it. Multicollinearity arises when two or more of the exogenous variables in an equation are so highly correlated over time that the method of least squares cannot distinguish the independent effects of the variables. Parameter estimates are imprecise when multicollinearity is present, although the equation might fit very well. Multicollinearity essentially is a data problem: the exogenous variables are not different enough. As a result, attempts to correct for multicollinearity without adding additional information are misguided. Sometimes it is possible to know in advance the relative magnitudes of the coefficients of the variables that are causing the problem. If so, it is a simple matter to impose constraints on the coefficients in estimation and improve the parameter estimates considerably.

Specification

In many ways equation specification is the most difficult task facing the applied econometrician. Economic theory and technological relationships sometimes dictate that certain variables should enter an equation in a particular manner. In most cases considerable judgment must be applied in specifying an equation, even if the variables that belong in it are known.

Functional Form

The first task in specifying an equation is to select a functional form. The most common forms are linear and logarithmic, but semilog and other forms are used occasionally. All simple functional forms should be considered as approximations to what are undoubtedly far more complicated relationships.

Linear Form

The linear form as in Equation 1 is probably the most commonly specified. A linear specification carries many implications. First, the error term has additive, not multiplicative effects. Second, the elasticities are of the form: $E_i = b_i X_i/Y$. This means that the elasticity estimates are not constant (which is not necessarily a drawback). The price elasticity of a linear demand equation, assuming that price has the expected negative coefficient, approaches negative infinity as price approaches some positive yet finite level. The elasticity of demand with respect to a measure of economic activity approaches unity as economic activity increases, as does the elasticity of supply with respect to price. Third, the marginal effects of each exogenous variable on the endogenous variable are independant. For example, the effect on demand of an increase in price do not depend on the level of economic activity.

Logarithmic Form

The logarithmic form looks identical to the linear form, except that all of the economic variables are replaced with their natural logs:

$$\log (Y_t) = b_o + \sum_i b_i \log (X_{it}) + e_t \qquad (4)$$

The elasticities in the logarithmic equation are simply the coefficients, the b_1, so they are constant. This means that as price goes up, larger increases in price are required to elicit constant declines in demand. The error terms have multiplicative effects, as can be seen by exponentiating both sides of Equation 4. The marginal effects of exogenous variables on the endogenous variable depend on the levels of the other exogenous variables. This can be desirable property if it is thought that price influences the unit input of a good into a production process. However, it can lead to problems if two measures of economic activity are included in a demand equation. For example, suppose that a demand equation is estimated for ferrochromium consumption, and the the production of stainless steel and the production of alloy steel are included in the equation. Why should the marginal effect of stainless steel production on ferrochromium consumption depend on the level of alloy steel production? This property generally makes the logarithmic form undersirable if two or more measures of economic activity must be included in a demand equation.

Other Forms

Although econometric equations can be specified in other forms, they seldom are. Probably the most common other form is a mixture of the linear and logarithmic. For example, the log of the dependent variable can be regressed on a linear function of the independent variables:

$$\log (Y_t) = b_o + \sum_i b_i X_{it} + e_t \qquad (5)$$

In this form the elasticities can be expressed as b_1X_1, which means that all elasticities approach infinity (or negative infinity, if b_1 is less than zero) as the independent variables approach infinity. A price elasticity of demand specified in semilog form, therefore, is a compromise between the fixed price elasticity of the logarithmic model and the rapidly changing elasticity of the linear model.

This semilog specification has one major drawback in demand modeling. The elasticity of demand with respect to an economic activity variable, like chromium consumption with respect to stainless steel production, should not increase without bound. This suggests that a mixed model might be appealing in some cases:

$$\log (Y_t) = b_o + b_1 \log (X_t) - b_2 P_t + e_t. \qquad (6)$$

This specification, to which other variables might be added, has a constant elasticity for X and a changing elasticity for P.

Interactive Effects

Variables can be entered into an equation in almost any form, provided that the form accords with economic theory and technological constraints. Powers, inverses, logs, and other forms can be used. Of potentially greater interest, however, is the ease with which interactive effects can be incorporated into an equation. For example, suppose that a certain price effect is expected to be observed only if the price is above a specified level. In that case two price variables can be defined, one equal to the minimum of price and the specified level, and the other equal to the difference between price and the specified level, if positive, and zero otherwise. That is, suppose that P^* is the point at which the price effect changes. Then define $P_{1t} = \min(P_t, P^*)$ and $P_{2t} = \max(0, P_t - P^*)$. The equation specification becomes:

$$Y_t = b_o + b_1 P_{1t} + b_2 P_{2t} + \text{other terms} \qquad (7)$$

Consider again the case of nonreversible substitution. Suppose that demand in a particular sector includes two uses, one of which is subject to irreversible substitution if the price exceeds some level P^*. Then define the variable.

$$Z_t = \max\left(0, \max_{m \leq t} (P_n - P^*)\right).$$

That is, Z_t equals the largest spread observed between P and P^*, if positive; if P has never exceeded P^*, then Z_t is defined as zero. The demand equation might be specified as:

$$y_t = b_o + {}_i\, b_i X_{it} - a_1 P_t - a_2 Z_t + e_t \qquad (8)$$

As long as the price remains below P^*, no irreversible substitution occurs. If the price increases above P, a sort of ratchet effect occurs. The higher the price goes, the greater the permanent decline in demand.

Many other kinds of interactive effects are possible. They can involve two or more economic variables, or they can involve economic variables and dummy variables (indicator variables that generally take on values of either zero or one), or they can involve time trends. The flexibility provided by interactive effects greatly expands the usefulness of econometric analyses of metal demands.

Dummy Variables

Dummy variables have many other uses. Suppose that a price effect or an economic activity effect is thought to have changed after a certain year, due to technological change or other reasons. Then a dummy variable defined as zero before the change and one after can be used to measure the change in one of several ways. In the case of a price effect the dummy can be entered in an interactive manner with the price variable. For an economic activity effect, such as a sudden decline in the unit consumption of metal in a particular use, the dummy could be entered directly into a logarithmic equation or interacted with the variable in question in a linear equation. In this way the data can be used to measure the effects of the technological change.

Lags

In many cases, the effects of a change in price or economic activity are not felt immediately but are spread over a number of years. In such circumstances the inclusion of only the current price or a current measure of economic activity inadequately reflects real world behavior. Lagged values must be entered into the equation in some manner.

Lag structures can take many shapes, some of which are depicted in Figure 21. Lag structures usually are imposed, within broad limits, on variables because simply including many lagged variables in an equation usually leads to a high degree of multicollinearity. Imposing a lag structure can be viewed as using prior information to eliminate the multicollinearity.

Most of the lag structures depicted in Figure 21 are imposed by constructing one or more new variables and using them in the regression equation. For example, suppose that the linearly declining lag structure is appropriate for price effects in a given equation, with a maximum lag of T. This lag structure can be imposed by constructing:

$$Z_t = \sum_{m=0}^{T} (T+1-m)\ P_{t-m} / \ ((T+1)\ (T+2)/2) \tag{9}$$

and replacing the price variables in the equation with Z_t.

The major exception to this approach is the geometric lag and its variants. The geometric lag is imposed by including the lagged dependent variable in the right side of the regression equation:

$$Y_t = b_o + \sum_i b_i\ X_{it} + hY_{t-1} + e_t \tag{10}$$

The geometric lag is attractive because of the simplicity it offers in estimation, but the simplicity can be deceiving. The major problem is that it imposes the same lag structure on all of the exogenous variables in the equation. Special techniques can be used to impose the lag structure on only one variable, but then the simplicity disappears. Furthermore, the problem of autocorrelation with a lagged endogenous variable can arise. The linearity declining lag probably is preferable to the geometric lag in most demand equations.

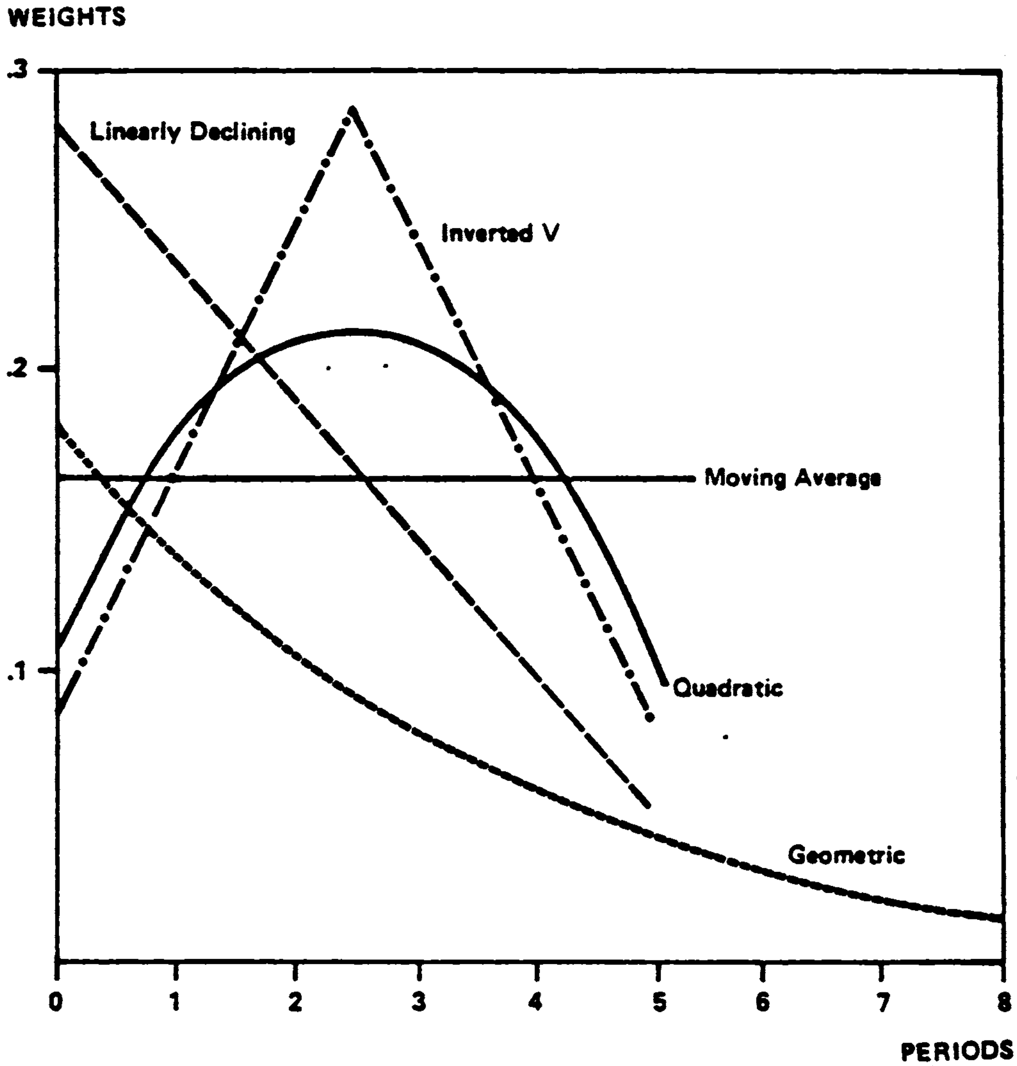


Figure 21. Examples of Lag Structures with Means of Lag of 2.5 Periods

Long lag structures seldom are imposed on economic activity variables in demand equations, because the effects of such variables generally are more immediate. An annual model seldom will contain more than the current and one lagged economic activity variable. The coefficient of the current variable generally should be positive, but the coefficient of the lagged variable often can be negative, depending on the particular specification. Suppose that a demand equation is being estimated for a product generally used in capital goods, but that no good measure exists of investment in the appropriate sector of the economy. In this case consumption of the metal is likely to be related positively to both output and the increase in output in that sector. This implies a negative coefficient for the lagged economic activity variable.

Deflation of Variables

All variables measured in dollars, whether they be prices, shipments, inventories, or other relevant variables, should be deflated by appropriate price deflators before being included in demand equations.

Demands for materials are real, physical demands, and they respond to real movements in the economy. Mixing real and nomial variables in a regression equation is seldom if ever justified.

Combined Methods

Economectric methods need not depend solely on statistical analyses of data. If information is available that is not fully reflected in the data, it should be used to improve the accuracy and applicability of the econometric results. This procedure often combnines econometric and engineering methods.

In a project for the Experimental Technology Incentives Program of the National Bureau of Standards, researchers at Charles River Associates followed such an approach in constructing demand equations for both chromium and manganese. We were interested in finding how demand would respond to movements in price far beyond this historical range, by a factor of approximately 10. Econometric methods alone cannot be trusted to yield accurate answers to such questions, because they only reflect past behavior and future behavior is likely to be quite different. We gathered engineering information on the technological possibilities for substitution and conservation over time, at various price levels. On a judgmental basis we then derived engineering estimates of the price elasticity of demand at prices far outside the historical range. We then used this information in estimating demand equation, imposing the engineering elasticities on the regression equation. The result, for both chromium and manganese, was a demand curve that reflected the engineering realities that consumers would face at vastly higher prices.

A somewhat different approach was used in a recent study of the world cobalt market that Charles River Associates performed. Economists and engineers worked together to determine the economic and technical feasibility of substituting other materials for cobalt in its major applications, at three different price levels. Forecasts of economic activity then were made, concentrating on those variables that are most important in determining the level of cobalt. Judgmental forecasts of demand in each of the important uses were made for a 10-year period for each of the three price levels. These judgmental forecasts were pooled across price levels and regressed on the measures of economic activity and prices, to derive engineering economic demand curves for the future. The equations were used in a model of the cobalt market to analyze the likely path of future cobalt prices. This approach enabled us to predict demand at price levels other than those assumed in the engineering analysis, and to change the assumptions concerning economic growth without being forced to redo the judgmental forecasts.

Usefulness and Limitations

Econometric analysis of materials demand is particularly useful when applied in situations that can be parametrized to resemble the past. If a major shift in technology will occur in the future, econometrics can be used to summarize the extent of the shift, as Charles River Associates did for cobalt, but econometrics cannot predict that such a shift will occur. For example, econometrics could not have been used in the late 1960's to predict the shift from low carbon ferrochromium to high carbon ferrochromium that would take place in the following decade after the introduction of the argon-oxygen decarburization (AOD) process in stainless steel production. Common sense said that such a substitution would occur, but the statistical data available would have been of no value in assessing the likely extent of the shift. The data available in the last decade, however, probably could be used to predict the extent of a reverse substitution, should the AOD process suddenly lose its economic advantages (a highly unlikely event).

One of the problems encountered when applying an econometric demand equation to a suddenly different world can be seen in Figure 22. This figure depicts three demand curves--linear, logarithmic, and semilog--with the same elasticity (-1.0) and demand levels at P = 2. It is immediately apparent that the curves differ very little for prices between 1.5 and 2.5 Given historical data in such a range, the curves may be virtually indistinguishable on a statistical basis. That is, all three specifications might fit the data equally well. For forecasting within the historical range, it should matter little which demand curve is used. If a sudden shortage develops and price increases drastically, however, the curves are very different. Consider price increases from a hypothetical historical average of P = 2. In order to realize a 50 percent decline in demand, the linear curve would call for a price increase on only 50 percent, the semilog curve for an increase of 70 percent, and the logarithmic curve for an increase of 100 percent. To realize a decline in demand of 75 percent, the linear curve calls for a 75 percent increase in price, the semilog curve for a 120 percent increase, and the logarithmic curve for a 300 percent increase. The functional form chosen is very important to the analysis, although historically the choice of functional form mattered little.

The lesson is clear. Great care should be taken if an economic model must be used to forecast outside the historical range of the data. Incorporating engineering and other information directly into the econometric analysis can greatly improve the usefulness of econometrics in such applications.

The major conclusion to be drawn from this paper is that econometrics, if properly applied, can be a very useful tool in studying past materials substitution, in forecasting the future if it will sufficiently resemble the present, and in incorporating certain kinds of engineering information into a consistent market analysis. Great care must be taken at each step of the analysis, from equation specification, through estimation, and into use. As Frisch said:

> Economectrics is a powerful tool, but also a dangerous one. There are so many chances of abusing it, that it should only be put into the hands of really first-rate men. Others should be absolutely discouraged from taking up econometrics. (Frisch, 1946)

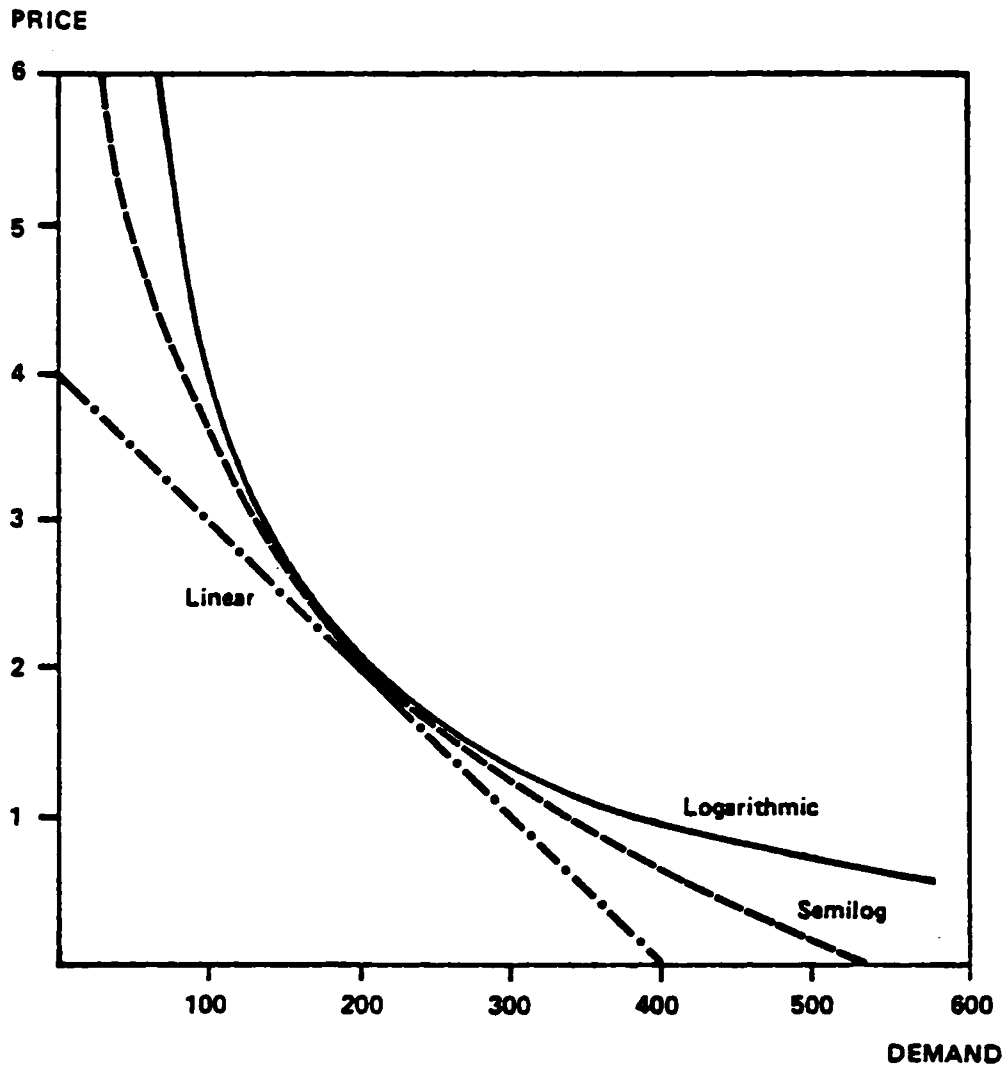


Figure 22. Comparison of Functional Forms for Demand Curves

REFERENCES

Frisch, Ragnar. 1933. Editorial, Ecometrica 1, 2.

Frisch, Ragnar. 1946. The Responsibility of the Econometrician. Econometrica 14:4.

Haavelmo, Grygve. 1944. The Probability Approach in Econometrics. Econometrica Supplement, 118.

MATERIALS SHORTAGES AND/OR MATERIALS SUBSTITUTION: ALTERNATIVE PATHS TO THE YEAR 2000

Ira Sohn
Institute for Economic Analysis
New York University
New York

John Maynard Keynes, the celebrated British economist and government adviser, once remarked that the role of government policies should be directed to short-term disturbances rather than long-term considerations because, "in the long-run we are all dead." However, it is unclear if the statement was made by Keynes the Economist or Keynes the Government Adviser. If it was the latter, the long-run probably is the day after Election Day, while if it was the former, the short-run could be years, or however long is required to change the amount of a previously fixed primary input--i.e., labor, capital, natural resources, or technology.

One would have hoped that it was Keynes the Economist speaking but judging by his classic treatise, (1936), around which most western economic policy decisions have been framed since the end of the last World War, unfortunately policy decisions based on short-term political expediencies have all too often supplanted considerations based on an integrated, comprehensive, and internally consistent long-term approach. Of all the western countries, guided by the Keynesian approach, it is in the United States that his obsession with the politician's myopia is most exemplified. Many of our current economic and political problems, domestically and internationally, can be attributed to this fixation of focusing on the immediate short-term without regard to the wider long-term approach.

Nowhere is our prevailing short-sighted approach to economic problems more acute than in our treatment of the so called "primary factors of production," mentioned above. A shortage of engineers cannot be relieved in one, two, or even three years. On the other hand, a surplus of teachers cannot be removed in one year. New capacity for extraction and processing of minerals, both fuel and nonfuel, cannot be brought on-stream, judging by today's experiences, even in five years. The development and installation of new technologies involves massive research expenditures in terms of human, material, and financial resources. Even the comfort of a full four-year Presidential term of office could not effectively address these problems.

Recognizing that the lead time needed to supply additional energy capacity is anywhere from five to 12 years, depending on the energy source, or as we are painfully witnessing, at least three to five years to retool the automobile industry, how can we incorporate these "technical" facts into our methodological framework? If the nation is to make a serious attempt to limit its dependence on increasingly unreliable sources of foreign oil by generating energy from our ample domestic coal sources, who is to insure that the rolling stock and railroad track needed to transport the coal, not to say the additional steel capacity required to build the cars and track, will be forthcoming? Can we at least provide the signals for this change by describing the implications of this policy on the whole structure of industrial output? How will employment opportunities be affected by the adoption of this policy or any combination of other alternative energy policies?

Generally speaking, our economy is required to turn out, on an annual basis, a high "standard of living" for approximately 220 million people. That "standard of living" is shaped by three forces: a culture, which fixes what economists refer to as the "tastes and preferences" of consumers; technology, which prescribes the efficient means for producing the goods and services; and the endowments of primary resources--labor, capital, and natural resources, the stock of which influences on the one hand what is produced domestically and on the other which goods are imported in order to secure the above "standard of living."

In addition to the physical constraints of factor endowments and the technological constraints imposed on production, social, institutional, and environmental limitations are designed by the government to influence or conform with the consumer's "tastes and preferences."

Any comprehensive overview of our increasingly complex economy has to describe the structure of consumption, the technological structure, and the factor endowments in addition to the political, demographic, social, environmental, and institutional phenomena around which the accommodating, and in many cases, competing economic relationships are woven.

Over more than four decades since Keynes wrote, our understanding about the way in which our economy, so to say, "hangs together" has become both wider and deeper. Our methodological approach to economic analysis, what Schumpeter referred to as the "economist's analytical tool kit," now incorporates the input-output technique, which has facilitated a structural description of the economy based on an empirical investigation of its component parts.

The input-output technique of economic analysis, first introduced some 40 years ago, is specifically designed as a tool for the systematic analysis of the mutual interdependence between the different parts of an economy. It is a method of analysis which describes the economy as a system of interacting activities--both in a direct and indirect way. Hence, it is particularly well suited not only to the preparation of internally consistent multisectoral projections of prevailing economic trends but also for a detailed quantitative assessment of both the direct and indirect secondary effects of any single policy action or any combination of policy actions.

An input-output table provides a systematic picture of the flow of goods and services among all the producing and consuming sectors of a given economy; that is, among all the various branches of business, households, and government. It could also register the flow of goods and services out of a given region and the flow of goods and services received from the outside. The input structure of each producing sector is specified in terms of what might be called its technology or its "cooking recipe," that is, a set of technical coefficients specifying the amount of goods and services, including labor, that a sector has to absorb to produce a unit of its own output. A separate set of capital coefficients describes the stocks of buildings and equipment as well as of all kinds of working inventories that each producing sector maintains in the process of transforming the proper combination of its inputs into its final output of goods and services. The inputs of primary natural resources, such as agricultural land, water, and ores, the output of pollutants by all producing sectors of the economy, as well as households (i.e., consumption patterns), are depicted and analyzed in modern input-output analysis along with the production and consumption of ordinary goods. Prices are also determined in an open input-output system from a set of equations which states that the price which each productive sector of the economy receives must equal the total payments made by it, per unit of its product, for inputs purchased from itself and from the other industries, plus a "value-added," which essentially represents payment for labor, capital, taxes, etc.

The input-output approach is particularly well suited to address questions of public policy debate. For example, if the current rate of foreign imports as a percentage of total U.S. sales persists or increases, what will be the direct and, perhaps more important, the indirect effects of this on the demand for labor in the U.S.? What are the direct and indirect costs imposed on domestic manufacturers that can be attributed to current or future EPA abatement standards? Or, what are the direct and indirect implications for the domestic economy, if high-strength plastics are substituted for metallic materials in automobiles?

In addition to the concern currently being voiced with regard to our dependence on foreign sources of supply to satisfy our energy requirements, we are also examining our position with regard to long-term, nonfuel mineral resources--both metallic and nonmetallic.

THE APPLICATION OF INPUT-OUTPUT ANALYSIS TO MODEL THE MINERALS SECTOR*

Concern about the adequacy of nonfuel mineral supplies, import dependence, trade restrictions, environmental and safety regulation, rising prices, sluggish technological innovation, and failure to expand capacity, as well as environmental consequences of more intensive mineral exploitation, is

* This section first appeared in the "The Future of the U.S. Minerals Industry Within the Changing Structure of the U.S. and World Economy," W. Leontief, S. Nasar-O'Brien, I. Sohn, C. Varsavsky, AIME Annual Meeting, Las Vegas, Nevada, February 1980.

clearly growing in the United States. As yet, there is no comprehensive or consistent program for nonfuel minerals, nor has one been proposed. In fact, the collection and assembly of elementary information and the application of systematic analysis are still in their very early stages with respect to nonfuel minerals.

A research program now in its third and final year is a study of the production and consumption of 26 nonfuel minerals in the United States to the year 2000. The study being funded by the RANN division of the National Science Foundation and the Bureau of Mines. The model which is being constructed is not a minerals model per se, but rather a general input-output model whose formulation and data base permit its use for analyzing the present and future position of nonfuel minerals in the United States economy.

A model of this type can be helpful in answering certain types of questions. For example,

- What are the economic and (to a limited extend) environmental consequences of technological change in minerals producing and consuming sectors?
- What are the implications of scarcity due to insufficient mining and processing capacity?
- What are the economic implications of specific environmental, trade, subsidy, or tax policies?
- What are the consequences of increased national income and population growth on minerals demand?
- What are the consequences of changes in the composition of the final bill of goods on minerals demand?

In short, the model is designed to address questions concerning mineral requirements under alternative hypothetical scenarios, rather than with the determinants of supply or final demand. Input-output models depart from "mainstream" economics modeling approaches to minerals supply and demand in that they represent intermediate demand explicitly, and focus on changes in the intermediate structure of demand stemming from substitution, and/or adoption of new technologies. Furthermore, the focus on technological underpinnings permits the introduction, as in this particular model, of physical quantities (i.e., tons) and the expression of physical characteristics in the formulation of the equations. The corollary to the introduction of physical, rather than value, units is greater and greater disaggregation of producing and consuming sectors. The fact that not only the sectors under study at a particular time are represented, but that an attempt is made to describe all economic and at least some noneconomic physical flows and stocks (emissions, for instance), implies both consistency and the ability to explore interdependence and second-order effects.

A feature of the particular strategy employed in the current study which should be emphasized--and the ease with which this type of model lends itself to the strategy--is the use of alternative scenarios as opposed to a "best guess." The model, deliberately, has many more variables than equations. Additional degrees of freedom reflect, essentially, two underlying attitudes held by the model builders. First, that the future is determined not simply by today's structure, but also by deliberate policy. Second, that most basic information about possible future income levels, consumption patterns, technological change and substitution, population growth, estimates of resource stocks, policies, and contingencies is fairly primitive, biased toward the present and past, and all projections based on this information are subject to a very great range of uncertainty. The use of scenarios--based on alternative estimates for various combinations of variables specified as exogenous--allows the analysts to make use of currently available information to explore the implications of alternative facts and policies without, however, obscuring from the user of these results the range of uncertainty which is likely to be the case.

SPECIFIC FEATURES OF MINERALS INDUSTRIES

Classification of Minerals

One of the basic requirements for a model which aspires to realism (and application to empirical problems) is that it captures the "major" specific characteristics of the industries or commodities which will be studied; in the case of minerals, these characteristics have generally to do with the underlying physical and technological properties and relationships.

The "essential" description is that of the profile of production and consumption of each nonfuel mineral. Demand for minerals is derived from the demand for products which embody minerals or require them for their production. Changes in the requirements for minerals result from technological changes, substitution between materials, and changes in the level and composition of final demand. A matrix of input-output coefficients permits the tracing of demand for final goods and services to the demand for specific minerals. There are input-output coefficients in official tables that are derived from actual transactions between different sectors in a particular year, balancing all purchases of inputs with sales of outputs, and dividing the amount of each sector's inputs by the sector's output. Each column vector of the matrix, therefore, is a sort of "cooking recipe" or list of needed ingredients which represents the "average" technology of production at one point in time.

Since statistical data on interindustry purchases and sales are limited, very often, to relatively large aggregates of minerals (for instance, nonferrous metals, ferrous metals, non metallic minerals), which lump together, from the point of view of a study of this nature, rather heterogeneous commodities, disaggregation of existing official input-output sectors requires the use of a judicious mix of engineering information, common sense, and practical considerations.

By-Products

The mineral extraction and processing sectors of the economy very often in addition of their principal output, generate varying amounts of by-products or co-products. For example, significant amounts of molybdenum are recovered in the mining of copper in southern Arizona. On the other hand, current environmental policy has required smelters and refineries to control their sulfur dioxide emissions below certain prescribed levels. One consequence of the legislation, according to experts are the Bureau of Mines, is that if current abatement policies are maintained, the amount of by-product sulfuric acid produced by the domestic copper, steel, and petroleum industries should be sufficient to satisfy total annual U.S. demand for sulfuric acid by the year 1985, thus eliminating the need for a domestic elemental sulfur industry. Therefore, any model which attempts to describe the minerals sector within the wider context of a social economic framework should be able to incorporate both of these by-products, other minerals, and pollution into the model.

Recycling

Domestic supplies of minerals can be augmented not only by imports but also by recycling. Of the 26 minerals included in the study, 13 are currently recycled to one degree or another in the United States. Recycling of both home and secondary scrap should be taken into account in projecting the required levels of mining and smelting activity needed to satisfy future levels of final demand. It would also be of interest to calculate the implications for the demand for minerals resulting from alternative rates of recycling in selected minerals. Therefore, recycling activities, and the flexibility of varying their rates, is yet another characteristic of the minerals sector that is being incorporated into the model.

Capital Requirements

The minerals sector, including both the mining and processing stages, is characterized as being highly capital intensive as well as requiring relatively long lead times to bring new industrial capacity into operation. Concern has been voiced both in the private and public sectors regarding the insufficient capacity of domestic suppliers of metallic and nonmetallic nonfuel minerals to meet the current and projected levels of demand for minerals. Increasingly, we have been forced to look abroad for supplementary and in some cases, competitive sources of supply.

Before promoting policies conducive to capital formation in the minerals sector, it is necessary to have a blueprint of the capital structure of this sector--i.e., a detailed bill of plant and equipment that must be in place, and then a detailed account of the capital requirements needed to expand capacity. With this information in hand, internally consistent and well coordinated policies can be formulated, directed at individual sectors, to insure the smooth expansion of required capacity.

With cooperation from major metal producers, such as ASARCO, Inc., AMAX, Inc., Phelps-Dodge, Inc., and INCO, Inc., we have been obtaining detailed engineering data describing the plant and equipment requirements of smelters and mines, disaggregated to the level of individual pieces of equipment and construction materials for each section of the installation. The data are then consigned to about 50 SIC four-digit categories, thus making them compatible with the format of the input-output framework.

Structure of the Minerals Model

The foundation upon which the minerals model is being constructed is the recently released 1972 U.S. Input-Output Table. (This table, prepared by the Bureau of Economic Analysis, Department of Commerce, divides the economy into 496 sectors. A version of the table aggregated to 85 sectors is published in the Survey of Current Business, February and April 1979.) Since this table, even at its most detailed level of disaggregation, has only four nonfuel mineral mining sectors and only seven primary metal sectors, considerable effort is being expended in disaggregating these sectors in order to incorporate a detailed description of the mining and processing of the following minerals specified in our study:

<u>Steel Industry</u>

Iron
Nickel
Manganese
Chromium
Silicon
Tungsten
Molybdenum
Vanadium

<u>Miscellaneous Chemicals</u>

Sulphur
Fluorine
Chlorine
Sodium
Boron

<u>Non-Ferrous Metals</u>

Bauxite and Other aluminum bearing ores
Copper
Titanium
Lead
Precious Metals (gold, silver, platinum)
Zinc
Magnesium
Tin
Mercury

<u>Fertilizer Industry</u>

Potassium
Phosphorus

Both mining and processing sectors are being developed for each of the above 26 minerals. In this way, it becomes possible to distribute the output of each of these sectors to its users and to account for all the inputs that enter into the production of each sectors' output. The input-output techniques can accommodate the use of different physical values (i.e., tons, ounces, etc.) or dollar values to serve as a unit of account for the output of each sector.

Consequently, where appropriate, we are exploiting this special feature of input-output analysis. Although not shown here because of its detail, in aggregate representation of the structural matrix for the minerals model, specially tailored for describing the activities of the minerals sector of the economy in detail, both with respect to its inputs and the use of its output, would show 173 rows and 153 columns of classification. In some cases, the 85 sector level of aggregation of the Department of Commerce Input-Output table is used to represent economic sectors, such as Office, Computing and Accounting Machines, or Radio, Television and Communication Equipment. In others, the more detailed 496-sector classification is used--i.e., Blast Furnace and Steel Mills, or Copper Rolling and Drawing. Finally, in the case of a majority of the minerals, specially developed sectors are incorporated and made compatible with the 85 and 496 sectors. The so-called "peculiarities" of the minerals sector that were discussed above, such as by-products and recycling, are explicitly taken into account in describing the minerals sector within the broader framework of the economy.

The United Nations World Input-Output Model

Since minerals are international in character, projecting the future consumption of minerals necessitates global projections. The United Nations (UN) World Input-Output Model is a newly forged tool capable of tracking many economic and social interdependencies throughout the world. Depite its global scope, the model displays an unusual degree of detail.

In order that it might accommodate the largest possible amounts of data, the model divides the world's economy into 15 regions which fall into three main groups: the developed regions, characterized by considerable industry and relatively high per capita income (North America, Europe, Soviet Union, Australia, South Africa, and Japan); the less developed regions rich in natural resources (the Middle East, Venezuela, some of the Andean countries, and some countries in arid and tropical Africa); and the less developed countries with few resources. The model describes each region in terms of 45 sectors of economic activity, including various types of agriculture, mining manufacturing, utilities, construction, services, transportation, communication, and pollution abatement. Though each region is initially treated separately, the model provides a complex linkage mechanism which analyzes their interconnections through trade, foreign investment, loans and interest payments, and foreign aid.

The UN model has yielded several different projections of the future world economy, depending on which of several assumptions about the rates of growth of population and of gross product per capita one choses to use. For each region, for example, we did not predict a single rate of population growth, but made alternative projections based on a high, a medium, and a low rate. The degree of detail in our model permitted the use of very specialized data--for example, about specific industries in specific regions--and resulted in relatively specific conclusions. The conclusions, however, are not based solely on the alternative assumptions about the growth of population and per capita gross product; they are also subject to modification by other variables.

Estimates of available reserves of various mineral resources differ widely. We did not choose among them, but made alternative computations on the basis of different estimates. We did not attempt, however, to predict possible discoveries of mineral resources. Ultimately, we constructed a set of scenarios, or hypothetical pictures of the world economy in coming decades, by introducing alternative assumptions about some of its components into the model and measuring their effect on other components.

Embedded in the structure of the World Model, which is being modified to incorporate a more detailed minerals classification, are the latest official endowments estimates, by region of reserves and subeconomic resources for most of the minerals in our study. The model will track the depletion of these stocks, by mineral and region, to the year 2000.

As a result of the increasing likelihood by 1990 of deep-sea mining of manganese nodules, with appreciable amounts of cobalt, copper, and nickel as a co-product, we are attempting to build this new source of supply into the framework.

Hence, the need for integrating both the production (endowment) side with the consumption side to insure internal consistency in the world economy. The gradual depletion first on a national and later on a global basis of estimated proven reserves will result in some combination of the following choices:

1. A substitution away from the "depleted" mineral to more abundant minerals or other materials, such as plastics, without violating the technological, environmental and other constraints discussed above.

2. Increased recycling of the "depleted" mineral in order to reduce the rate of extraction of the mineral, again carried out within the context of the above constraints.

3. An increase in the share of imports out of total U.S. domestic supply, with a simultaneous shift in export shares in favor of those regions endowed with both fuel and nonfuel resources (i.e., Australia, Mexico, Columbia).

4. Better "maintenance" of existing goods and structures in which minerals are embedded--i.e., a deliberate policy of protecting new structures and equipment against corrosion.

A recent study on the economic effects of metallic corrosion (Battelle 1978) completed for the Bureau of Standards by the Battelle Columbus Laboratories, carried out a series of projections of the use of metals in the U.S. economy based on alternative assumptions regarding corrosion protection of plant and equipment.

A series of projections based on various combinations of assumptions from the first three groups are being formulated within the context of the U.S. and world models incorporating the expanded mineral sectors described above.

For example, scenarios are being formulated to introduce a gradual increase in recycling rates of the 13 metals presently being recycled. Current recycling rates of these selected metals are presented in Table 6. The present study, however, stops short of considering the whole complex question of the generation of old scrap from discarded goods and structures, which would require linking the final demand sectors back to the production sectors via estimated "lifetimes" of commodities. In order to close the system, modeling the generation of old scrap would, of course, be the next logical step.

Table 6. Current Recycling Rates for Selected Metals[1]

Iron & Steel (m.s.t)	33.36%
Copper (t.s.t)	19.01%
Lead (t.s.t)	34.30%
Zinc (t.s.t)	5.20%
Aluminum (t.s.t)	4.50%
Nickel (t.s.t)	30.10%
Chromium (t.s.t)	5.90%
Gold (t.t.o)	8.50%
Silver (m.t.o)	20.30%
Tungsten (t.lb)	3.5%
Mercury (flasks)	22.90%
Tin (l.t)	18.60%
Magnesium (t.s.t)	0.54%

[1] Ratio of old scrap/total domestic demand, except for iron and steel, ratio of purchased scrap/total domestic demand. All ratios are for 1972. Source: Mineral Facts and Problems, Bicentennial Edition, Bureau of Mines, Bulletin 667, U.S. Department of the Interior 1975.

Scenarios are also being developed within the context of the minerals model described above to accommodate materials substitution over the next 20 years in certain key sectors of the economy. For example, in the 1980's high-strength plastics and/or nickel and titanium based superalloys are expected to be used at the expense of steel in the automobile, aerospace, and construction sectors. The input-output methodology is in a particularly advantageous position to incorporate technological change in materials use into a detailed profile of our complex economy while assessing the implications for the primary factors of production--labor and energy--resulting from the substitution.

A third "class" of scenarios being formulated is the future position of the U.S. economy with respect to trade in nonfuel minerals. This entails projecting the future share of imports as a percentage of total U.S. consumption of the 26 minerals in our study as well as the future export potential of the other regions of the world, as represented in the World Model.

Table 7 presents the trade coefficients which are being used for the baseline scenario, projecting the demand for the 26 nonfuel minerals to the year 2000. The baseline scenario is formulated under the following assumptions:

1. No technological change.

2. No change in recycling rates.

3. No changes in the import-export shares of the U.S. minerals sectors from the 1972 base year.

4. Final demand projections supplied by the Bureau of Labor Statistics, U.S. Department of Labor.

Table 7. 1972 Trade Coefficients for Metallic Minerals Used in Baseline Scenarios*

Iron	-0.44
Molybdenum	+0.37
Nickel	-0.90
Tungsten	-0.43
Manganese	-0.96
Chromium	-0.94
Copper	-0.10
Lead	-0.35
Zinc	-0.62
Gold	-0.78
Silver	-0.35
Bauxite	-0.78
Mercury	-0.77
Vanadium	-0.22
Platinum	-0.66
Titanium	-0.54
Tin	-0.98
Other Nonferrous Metals	-0.0003

*Trade Coefficients = $\frac{\text{Exports - Imports}}{\text{Domestic Consumption}}$

With baseline projections of mineral consumption on hand, we are formulating alternative sets of assumptions with regard to the future rate of world economic development, U.S. economic growth, conservation programs through increased recycling, changing patterns of world trade in minerals, and technological change. Each of the alternative scenarios will produce, as expected, different projections for the U.S. and world minerals output levels to the year 2000.

The question of whether or not sufficient capacity to extract, process, and transport the projected level of future U.S. minerals output will be forthcoming, is not a question that this study chooses to consider. In other words, utilizing the tool of input-output analysis, this research program is projecting the minerals, and all other sectoral, output levels of the U.S. economy under alternative institutional, economic, technological, and social settings. Favorable public and private policy decisions will, no doubt, be needed to insure a smooth and continuous increase in capacity to supply the projected needs of the U.S. and world economy to the year 2000. The implementation of these policies is, however, beyond the scope of this project.

In any systematic approach to the question of materials substitution, two relevant considerations occur: One, can a substitution be made? That is, within the context of technical feasibility and economic efficiency, which materials are admissible candidates as substitutes for each other, and to the exclusion of others? Two, assuming a substitution of one material for another has been implemented, what are the implications of that substitution for the system at large, within the context of institutional, economic, social, environmental, international, and physical constraints. Due to the complex interrelationships that characterize modern developed economies, it is impossible to know the direct and indirect consequences resulting from a single change in the economy--i.e., the substitution of one material for another, without a firm understanding of the underlying structural relationships that constitute a modern economy.

The input-output technique is <u>not</u> designed to address considerations of the first kind, that is, questions that center on what techniques of production should be used, when should the switch to a new technique be executed, what materials are likely candidates for substitution.

On the other hand, the input-output technique is in an advantageous position to consider questions of the second kind: what are the direct and indirect implications for the system as a whole resulting from materials substitution. In addition, since input-output tables can be very finely disaggregated, material by material, by observing both flows and coefficients, the approach does shed light on where to center research for possible future materials substitutions, that is by identifying the sectors and materials in which substitution, at least from a technical if not economic point of view, may be feasible.

With increasingly greater lead times required to bring new capacity on stream, due to technical as well as environmental reasons, the technique of input-output analysis is in a position to provide the necessary signals to private and public policy decision makers. The need for creating an institution equipped with a detailed and accurate profile of our increasingly complex economy cannot be underestimated at a time when we are reminded, almost daily, of the fragility of the entire network of relations that constitute the world economy. This institution would be mandated to oversee and coordinate the projected long-term changes in our economy--whether these changes are manifested in trade policy, environmental policy, social policy, or technological change--in a systematic and internally consistent way. The well documented success of the Norwegian and Japanese economies is, at least in part, due to a concerted and deliberate attempt to seek institutionalized, rational, long-term coordination of the economy in the face of changing economic assumptions, while keeping short-term political whims at bay.

REFERENCES

Keynes, J. M. 1936. The General Theory of Employment, Interest and Money. Harcourt, Brace and World. New York.

Economic Effects of Metallic Corrosion in the United States. 1978. A report to the National Bureau of Standards by Battelle Columbus Laboratories, Columbus, Ohio.

PROJECTIONS AND FORECASTS OF U.S. MINERAL DEMAND BY THE U.S. BUREAU OF MINES*

V. Anthony Cammarota, Jr.
William Y. Mo
Barry W. Klein
U.S. Department of the Interior, Bureau of Mines

Introduction

The Bureau of Mines forecasting system consists of two important components: statistical projections and contingency forecasts. The system is currently used to derive U.S. demand projections and forecasts for mineral commodities by end-use categories in the year 2000.

Both end-use demand projections and forecasts are published in the Bureau of Mines "Mineral Commodity Profiles" series. These reports provide the general public and decision makers with an overview of present and probable future supply-demand relationships for individual mineral industries.

In many cases in the decision-making process, it is asked how the Bureau of Mines projections and forecasts of future U.S. demand by end-use categories were derived. The primary purpose of this paper is to describe the methodology involved in deriving these projections and forecasts so that the user can make intelligent use of them.

The derivation of statistical projections is a more standardized and well-known procedure than making contingency forecasts. Therefore, the statistical projections will be described in the first part, and contingency forecasts based on the statistical projections as a guide will be presented in the second part. Zinc will be used as an example in both cases.

* Presented at the 109th Annual Meeting of the American Institute of Mining Metallurgical and Petroleum Engineers, Inc., Las Vegas, Nevada, February 24-29, 1980.

Statistical Projections

In the Bureau of Mines statistical demand projections system, it is assumed that the end-use consumption of a mineral commodity can be approximated by a simple linear regression equation, as follows:

$$Y_t = a + b\ X_t + e_t (\text{for } t=1,2,3,\ldots\ldots,n) \quad (1)$$

where Y_t = end-use consumption of a mineral commodity at time t,

X_t = a macroeconomic variable at time t,

e_t = a disturbance term,

and a and b are parameters.

Equation (1) specifies that end-use consumption of a mineral commodity is related primarily to a macroeconomic variable. Because mineral commodities are basic raw materials used in our economy, it is not unreasonable to expect that the demand for a mineral commodity will depend upon general economic conditions, and that the demand for the mineral will change as economic conditions change. Furthermore, mineral commodities may be used by different sectors of the economy in a variety of ways. Therefore, 38 macroeconomic variables are selected as possible explanatory variables. The list of these variables is given in the appendix.

Equation (1) is defined basically by two parameters, a and b, but the relationship is inexact because it also contains a disturbance (or error) term e_t. The inclusion of an error term in equation (1) is to take into account the influence of omitted variables, the error of approximation of the functional form, and other unpredictable random effects.

The unknown parameters a and b must be estimated by reference to observable data for some particular historical period. The data period selected for most commodities is 1960-77; in some cases, the estimated equations were based on a shorter period because end-use consumption data were not available as far back as 1960. Basically, the estimation method attempts to find the numerical estimates for the parameters a and b that, when used in Equation (1), best explain the known historical data. In other words, the objective of the method is to estimate the parameter values that make the accumulated squares of the sample period prediction errors as small as possible. This is commonly known as the "least-squares" criterion.

With additional assumptions about the disturbance term e_t in (1), the "least squares" estimators have additional desirable properties such as the "best linear unbiased" and "maximum likelihood" properties.

After several estimated equations have been obtained for an end-use consumption of a mineral commodity, it becomes necessary to determine which estimated equation should be selected as a basis for deriving the statistical projection. Obviously, among the estimated equations, the equation that has the smallest sum of the squared residuals (or predicted errors) should be selected. This is equivalent to selecting the estimated equation that has the highest coefficient of determination (R^2 - value).

The maximum value of R^2 is 1 which only occurs when all historical data points lie exactly on the estimated regression line; the observed values of the dependent variable U_t would be equal to the corresponding predicted values calculated from the estimated regression line. As the independent (or explanatory) variable explains less and less of the variation in the dependent variable, the value of R^2 falls closer and closer to zero. Hence, the R_2 - value provides a useful measure that can be used to determine which macroeconomic variable best describes the historical pattern of each end-use consumption and, therefore, which macroeconomic variable should be used as an explanatory factor for deriving each end-use statistical projection. Furthermore, if we assume that the relationship shown by the selected estimated equation will continue into the future, then projection can be easily obtained by solving this estimated equation for the variable to be projected by substituting the appropriate period's value for the macroeconomic explanatory variable in the estimated equation. Following this procedure, the statistical projection of zinc used for transportation, for example, is estimated to be 487,000 short tons in the year 2000, which is equivalent to 440,000 metric tons. This figure is obtained from the equation having the highest R_2 - value among the estimated equations:

$$Y_t = 120.2968 + 1.8423 \text{ FRBAUTO}_t \quad R^2 = 0.5660 \qquad (2)$$

where Y_t = zinc used for transportation at time t, in thousand short tons, and $FRBAUTO_t$ = Federal Reserve Board U.S. Industrial Production Index for automobiles at time t (1967=1000).

Other U.S. macroeconomic explanatory variables that were used to estimate zinc consumption equations for the transportation end use included gross national product, population, gross private domestic investment, and the Federal Reserve Board Industrial Production indexes for total production; transportation equipment; motor vehicles and parts; automobiles; trucks, buses, trailers; aircraft and parts; ships and boats; tires; fabricated metal products; and basic steel and mill products. The R^2 - values associated with these estimated equations range from 0.0443 to 0.5534.

The value of FRBAUTO in the above equation for the year 2000 is 199, based on a macroeconomic forecasting model from Data Resources, Inc. (DRI).

In the context of using the estimated Equation (2) for projection purposes, it is assumed that the past empirical relationship will continue in the future. Of course, for many reasons this relationship might not hold in the future; therefore contingency analyses to derive a set of forecasts are necessary.

Contingency Forecasts

In-depth contingency forecasts have been made for about 90 mineral commodities at approximately five-year intervals. The purpose of contingency analysis is to try to identify those problems or opportunities that could cause demand for a particular commodity to deviate markedly from its historical trend line. For major commodities or those commodities that had significant changes in supply, demand, economic, or environmental aspects, forecasts were updated when appropriate--usually once but sometimes twice during the past five-years. Because changes drastic enough to change significantly a 20-plus-year forecast seldom occur, the interim forecasts were not as detailed as those made every five years.

Three to five specialists on related commodities participate in making the forecast every five years. For example, in the case of zinc, the zinc commodity specialist, specialists for other base metals, and specialists for metals coproduced with zinc served as a forecasting group directed by the branch or division chief responsible for ensuring consistency among commodity groups.

The commodity specialist discussed the major end use of his or her commodity and described to the panel the economic or industrial indicators that historically related most closely to each end use as determined by the regression analysis. Using the statistical projection as a guide, the specialist arrived at the low, high, and most probable demand. Other specialists on the panel then critiqued the specialist's decisions, drawing on both their general knowledge and their knowledge of their commodities as it related to the commodity under discussion. Information on the properties of, and the present and future supply-demand situation for, the various commodities was especially useful in determining the potential for substitution.

The specialist then recorded the suggestions for the panel concerning the impact of various contingencies on the low, high, and probable demand for each end use. If growth for that commodity was expected to be affected by, or affect the growth of, another commodity, then when the second commodity was forecast this information was taken into account by the second panel. Continuity between panels was achieved by the presence of the appropriate branch or division chief on all panels. Owing to the interrelationships among commodities, and the number of commodities and specialists, most specialists served on several panels, resulting in further interchange of related facts and ideas.

Interim forecasts made between the five-year major forecasts are generally made by the specialists, who calls on other specialists with related commodities as the need arises. His or her contingency analysis and forecasts are reviewed by the branch or division chief.

When there is no indication since the previous forecast of a significant change in the outlook for the commodity, the demand quantities in 2000 are not changed, but the average annual growth rate may change as a result of using a different base year.

Returning to zinc, as a specific example, contingency forecasts represent the qualitative judgments of the Bureau of Mines specialists applied to the quantitative statistical projections of end use. The forecast of total zinc demand to the year 2000, 2 million metric tons, is the sum of the forecast demands for each major end-use category. The end-use forecasts to the year 2000 are made on the basis of the specialists' judgment about the impact on the projected demand for zinc of possible technological developments and market trends that might occur. The commodity specialist determines a forecast range delineated by a high and low that are based on assumptions related to such factors as technology, substitution, and changing demand patterns. The probable forecast is based on an appraisal of those developments that are most likely to occur based on present knowledge.

To illustrate the use of contingency forecasting, we can look at zinc and one of its major end uses, transportation, which in 1977 accounted for almost one-quarter of total demand. It should be pointed out that the use of zinc oxide in tires is included under another end-use category, rubber products. The amount of zinc used historically in transportation was estimated by correlating the quantities of zinc used in galvanizing, diecasting, and brass and bronze, as reported to the Bureau of Mines, with industry shipments of galvanized sheet and strip, brass, and diecasting to the transportation segment of the economy. The American Iron and Steel Institute, the Cooper Development Association, and the Zinc Institute, Inc. conduct surveys on shipments of these products to various industries. This information serves as the guide in making the estimates. Using this technique we estimated that about 270,000 metric tons of zinc was used in transportation in 1977. Of this total, about 178,000 tons was in the form of diecastings, 22,000 tons in brass, and 70,000 tons in zinc-coated steel such as galvanized sheet and Zincro-metal. The demand in 1977 and the low, high, and probable demand in 2000 are shown in Table 8.

Table 8. U.S. Zinc Demand for Transportation, 1977 and 2000 (Thousand metric tons)

	1977	Contingency forecast to 2000		
		Low	High	Probable
Zinc-coated steel	70	90	100	90
Brass	22	15	10	10
Diecasting	178	170	180	180
Batteries	---	95	530	180
Quantity total	270	370	820	460

All of the contingency forecasts utilize the statistical projections as their point of departure. The statistical projection for zinc used in transportation is 440,000 metric tons in the year 2000, as shown in the previous section. To obtain the low forecast, we assumed that a small portion of the forecast production of 19 million vehicles in 2000 would be electric vehicles (EV) powered by zinc-based batteries such as zinc-nickel or zinc-chlorine. In conjunction with this development, less brass would be used in radiators, and less zinc in diecastings would be used in smaller cars that would make up the major model series. Zinc-coated steel would play a major role in preventing corrosion of vehicle bodies, as Government regulations requiring minimum service life of auto bodies are implemented. Based on these contingencies, an estimate was made for the use of zinc for brass, diecasting, coatings, and batteries, as shown in the table, to five total low zinc demand of 370,000 tons for the transportation segment.

For the high of the range, we assumed a higher than forecast growth rate of vehicle production, leading to production of 20 million vehicles, of which about one-half would be EV's using zinc-based batteries. Zinc-coated steel would still be a major factor, and zinc diecasting would continue to play a part, as plastics become expensive and possibly in short supply. Both of these technical end uses show small gains over the low estimate in the table. The use of zinc in brass shows a decline in the high forecast because much less brass would be used in radiators with the high number of EV's being produced. It is evident that the high forecast of 820,000 tons of zinc envisions tremendous growth in zinc-based batteries; zinc demand in transportation would be up only slightly over that of 1977 unless EV's become a dominant factor.

For the probable estimate, we assume a lower than forecast growth rate in vehicle production, as production shifts to more mass transit vehicles, fewer cars and trucks, and smaller cars. A moderate number of EV's could be produced. Using the same relationship between vehicle production and zinc demand as we used for the low and high forecasts, total probable demand was set at 460,000 tons, near the low side of the range.

In summary, the Bureau of Mines forecasting method uses historical data and regression analyses, coupled with projected economic indicators, to reveal a future trend as a starting point. Contingency analysis takes into account many other factors that will influence demand and arrive at a range of possibilities based on expert judgment.

APPENDIX

Macroeconomic Variables Used in Bureau of Mines Statistical Demand Projection System:

- U.S. gross national product.
- U.S. population.
- U.S. gross private domestic investment.
- U.S. new construction activity.

Federal Reserve Board indexes of U.S. industrial production:

- Total production
- Textile mill products
- Paper and products
- Chemical and products
    - Basic chemicals
    - Synthetic materials
    - Paints

- Petroleum products
- Rubber and plastic products
    - Tires
    - Rubber excluding tires.
    - Plastic products

- Stone, clay and glass
- Iron and steel
    - Basic steel and mill products

- Fabricated metal products
    - Metal cans
    - Hardware, plumbing, structural metal

- Nonelectrical machinery
    - Construction and allied equipment
    - Metalworking machinery
    - Special and general industrial equipment

- Electrical machinery
    - Major electrical equipment and parts
    - Househould appliances
    - Communication equipment

- Transportation equipment
    - Motor vehicles and parts
    - Automobiles
    - Trucks, buses, and trailers
    - Aircraft, and parts
    - Ships and boats
    - Food and products
    - Ordnance

A DESIGN/MANUFACTURING INTERACTION TOOL FOR MATERIAL SUBSTITUTION TRADE-OFFS

Bryan R. Noton
Battelle Columbus Laboratories
Columbus, Ohio

Abstract

Constraints frequently exist in achieving cost-effective materials substitution and include lack of manufacturing cost-data to conduct trade-studies, cost of redesign, additional testing and tooling, and requirements to maintain form, fit and function of parts to production. Additional design tools are necessary for the conceptual, preliminary and detailed design phases. Such guides must, for example, enable designers to evaluate alternative concepts and the manufacturing technologies needed for emerging materials. At the conceptual and preliminary design phases, changes will have minimal cost impact. Current methods of estimating maunfacturing costs are discussed. The application of the Air Force ICAM Manufacturing Cost/Design Guide: (MC/DG) to the materials substitution problem, assisting designers and manufacturing engineers to avoid cost-drivers due to shortages, long lead-times, etc., is reviewed.

Materials Substitution Background

We are now in an age for which we are unprepared, an age of material shortages.

We are all acutely aware of the energy dilemma; many solutions are proposed and just as many controversial objections and constraints toward progress are advanced. However important the energy crisis is, of equal importance is the increasing awareness of materials shortages and the immediate need to find alternative materials.

Some of the contributing factors to material shortages are:

- Up to 90 percent of the columbium, manganese, tantalum, cobalt, chromium, bauxite, and alumina used in the United States are imported. Up to 50 percent of many other important materials are imported.

- Political upheavals in some of the exporting countries have disrupted supplies resulting in significant price increases.

- Lead-times for many materials and components have increased by a factor of from 2 to 4 times since 1977. Examples are:

  - Aluminum sheet; 18 weeks in 1977 to 72 weeks in 1980.

  - Precision forgings; 27 weeks in 1977 to 80 weeks in 1980.

  - Purchased parts (built-to-print); 35 weeks in 1977 to 78 weeks in 1980.

Another important factor which contributes to material substitution and the quest for alternative materials is the necessity for increased performance of complex systems such as jet-engines. The design engineer is driven and assessed by his ability to provide the marketplace with an ever-improved product with increased performance. The complex external forces and internal interactions in the design-to-cost process are shown in Figure 23. To accomplish these objectives, the engineer, using advanced technology, has not only improved existing materials, but has developed new "man-made" materials, new products, and new manufacturing technologies. A few examples in the aerospace industry are:

- Transition of the air frame structure from "wood-wire-cloth" to high-strength sheet metal structures utilizing aluminum and titanium.

- High-strength aluminum, steel and titanium alloys.

- Development of high-strength, high-modulus composite materials.

- Advanced welding technology; e.g., electron-beam, laser, and plasma.

- Environmental protection of surfaces.

- High-strength fasteners.

- High-strength castings and forgings.

- Powdered metal technology (shape technology).

- Improved metal removal cutting tools.

- Numerous applications of computer technology (CAM and CAD).

- Development of solid-state electronics.

The aerospace design team priorities are shown in Figure 24. To continue to meet the requirements for increased performance in face of material shortages and lead-time problems, engineers must have a working knowledge of the cost impact on manufacturing when selecting materials for new designs or when substituting alternative materials to improve performance or to overcome a material shortage. The Air Force ICAM Manufacturing Cost/Design Guide will provide this capability in design-to-cost efforts in all phases of the evolution of the design, as well as in the production phase of the product.

Constraints to Material Substitution

Just as constraints exist in the energy problem, such is also the case with the materials shortages problem. Some of these constraints are:

- Lack of engineering design data.
- Lack of adequate manufacturing cost data to conduct trade studies.
- Lack of in-service experience with some new materials.
- Initial "start-up"; e.g., tooling, new skills, learning curves, and high cost of some materials, such as advanced composites, titanium, and steel alloys.
- Increased logistic inventory required for more than one spare replacement.
- Reluctance to change, e.g., "let someone else take the risk" or the "not invented here" syndrome.
- Cost of redesign, additional testing, tooling, and also retraining.
- Necessity to maintain the "form, fit, and function" of parts designed and already committed to production.
- Investment requirements in new facilities and equipment.

Conceptual Design Phase

In today's competitive environment, the survival of a company depends on the ability of its design engineers to anticipate the needs and requirements of the customer, and to perceive how to meet those requirements at a lower cost and in a shorter time-span than competitors.

Top priority must be given to performance. Performance is achieved not only by superior knowledge and skill in the application of design theory and engineering disciplines related to the product, but also of importance is the creative and inventive mind. However, the design engineer must also recognize that performance must be achieved at an affordable cost and a major constraint to this necessary goal, among others, is the shortage of strategic materials.

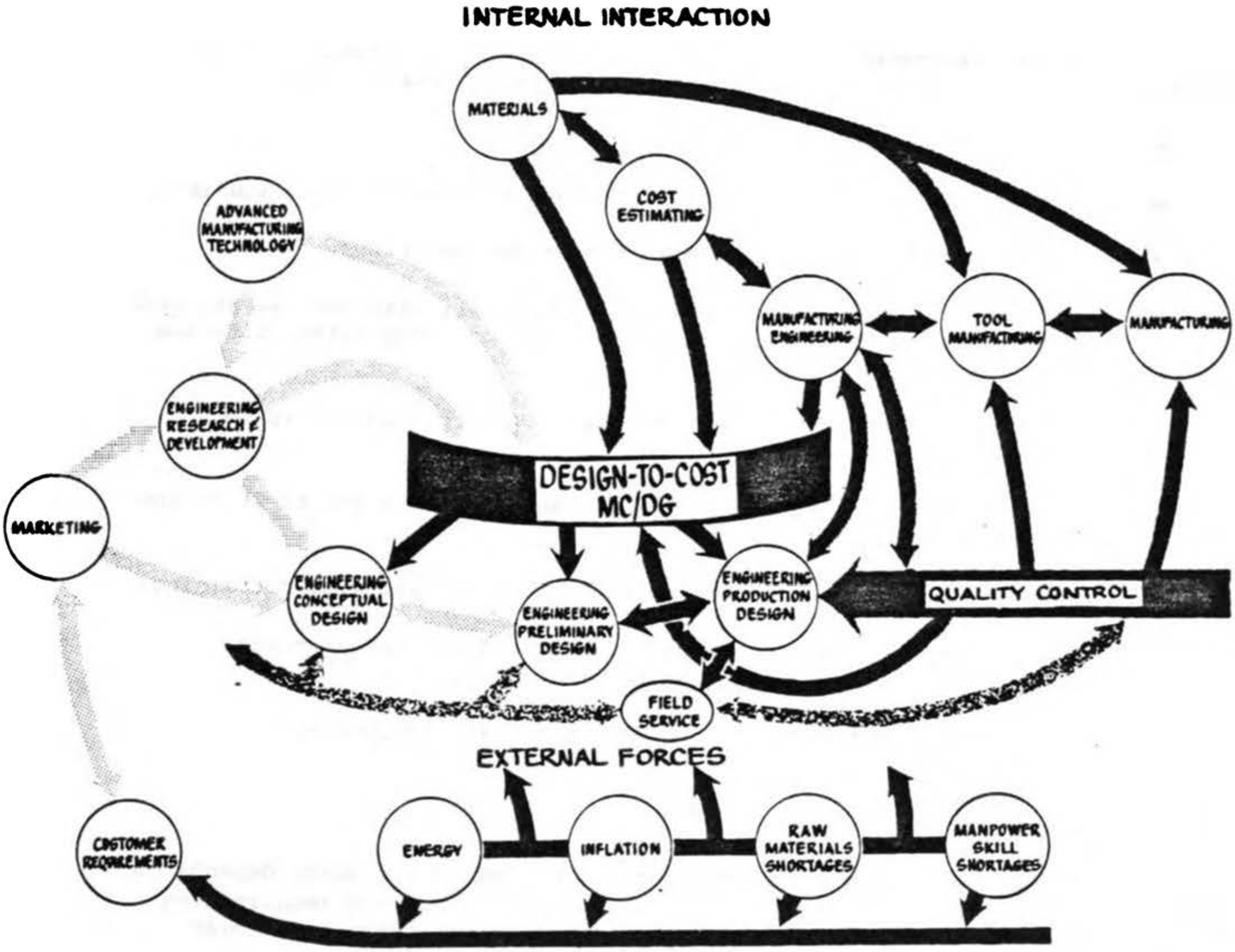


Figure 23. Importance of Recognition of External Forces and Internal Interactions in Design-to-Cost.

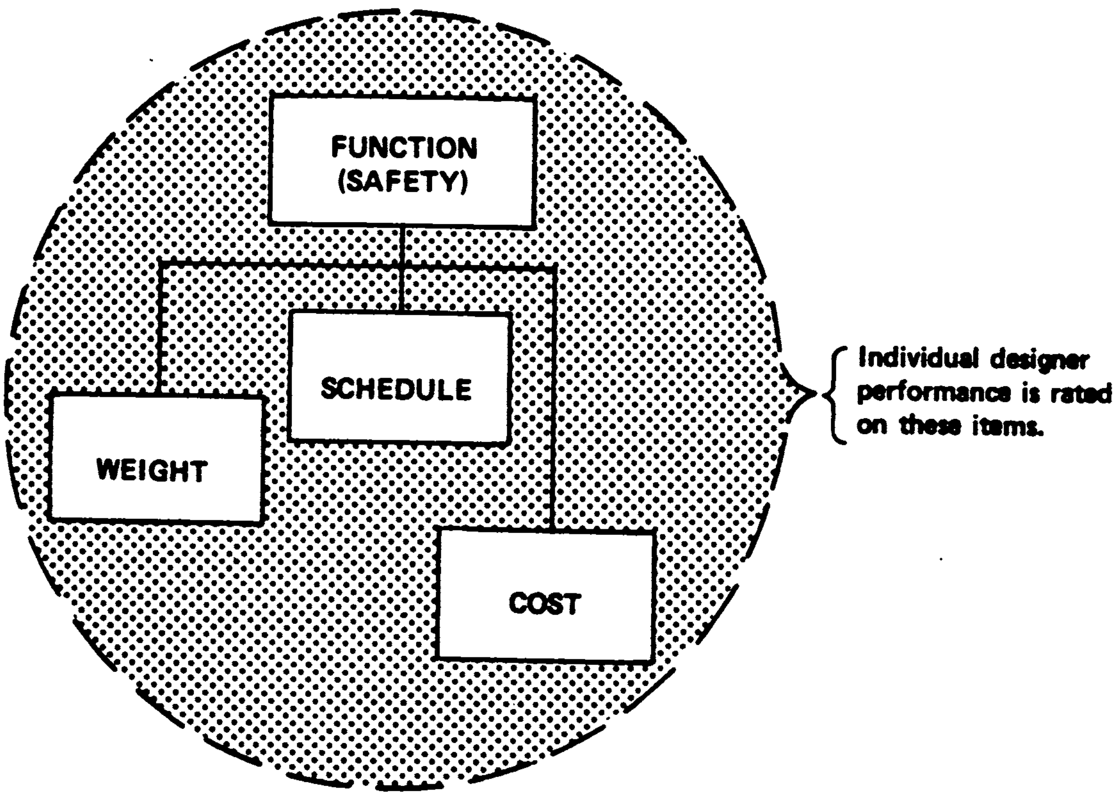


Figure 24. Present Aircraft Design Team Priorities

At the preproposal or conceptual design phase, decisions must be made which will ensure that the materials selected, when required will be readily available and cost-effective to use. The ability to accurately predict technology advances and materials requirements is an important factor in the conceptual design phase. Examples of these factors are:

- Long-range predictions for materials, including availability and cost projections must be utilized to avoid potential problems.
- Time is available to evaluate alternative design approaches.
- Several alternative design concepts or options can be evaluated.
- Manufacturing technology can be applied to "productionize" any new process/method required to utilize advanced materials.
- Material property data can be generated and made available to designers.

- Decisions made at the conceptual design phase are more readily changed with little or no cost impact.
- Degree of commonality with existing designs can be studied to reduce tooling and manufacturing costs.
- Provides the lead-time for the procurement of new facilities or equipment.
- Projected customer performance requirements, funding constraints, and schedule requirements are being generated.
- Provides management with the opportunity to evaluate long-term commitments and improve the competitive position of the company.

Importance of MC/DG at the Conceptual

Design Phase

Although decisions made by the design engineer at this early phase are subject to change, they normally have a major impact on the total life-cycle costs of the system to be developed. The decreasing leverage to achieve minimum cost, as the investments in a system increase, is shown in Figure 25. Examples of the decisions made throughout the design process and their cost impact are illustrated in Figures 26 and 27. It is most important that the design engineer is provided with qualitative and quantitative manufacturing cost data, besides mechanical property data, etc., if a design is to evolve that not only meets or exceeds performance goals, but does so at an affordable cost and ahead of the competition.

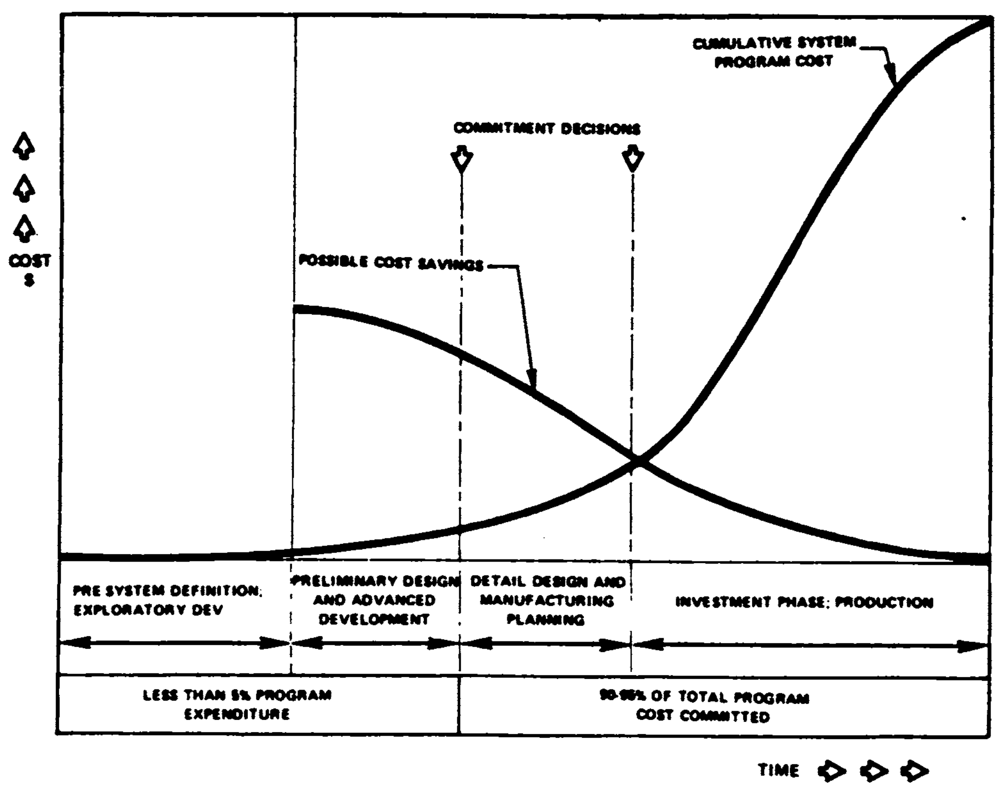


Figure 25. Decreasing Leverage to Achieve Cost Savings as Development Progresses.

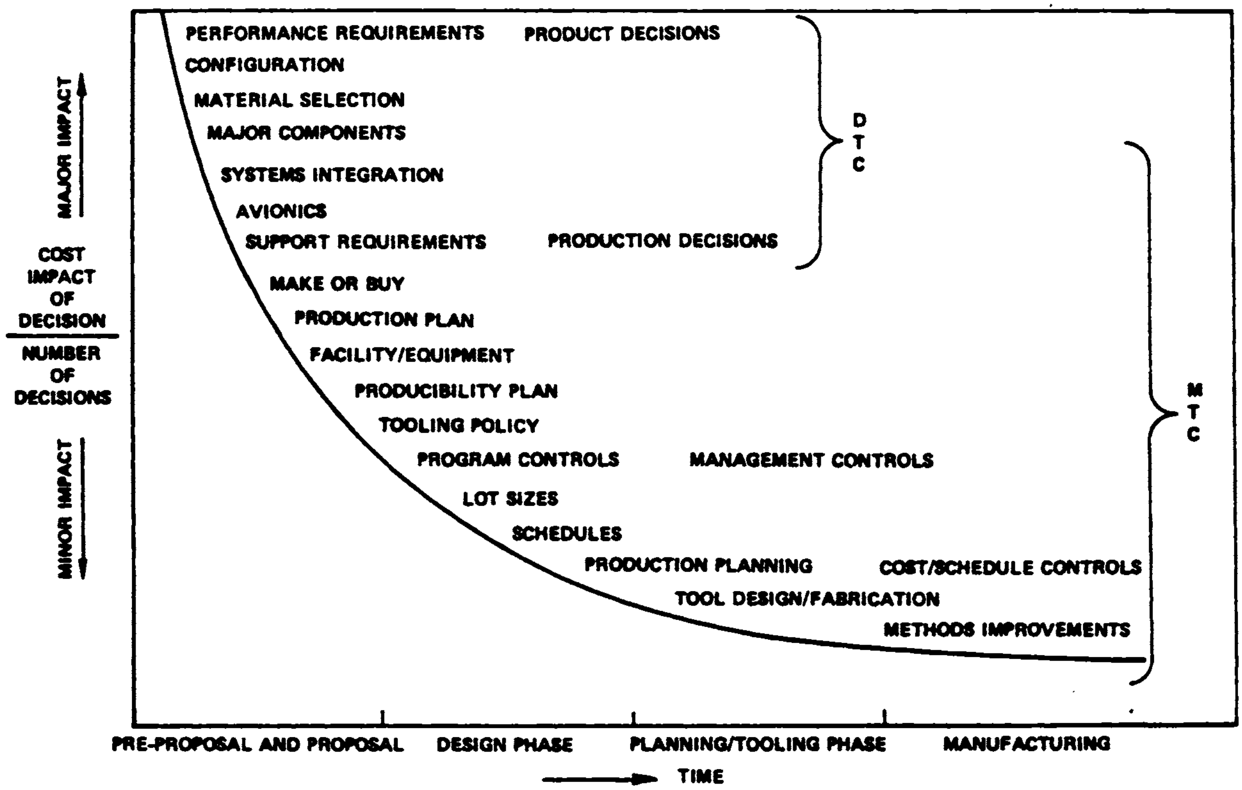


Figure 26. Major Decisions and Decreasing Cost Impact With Time

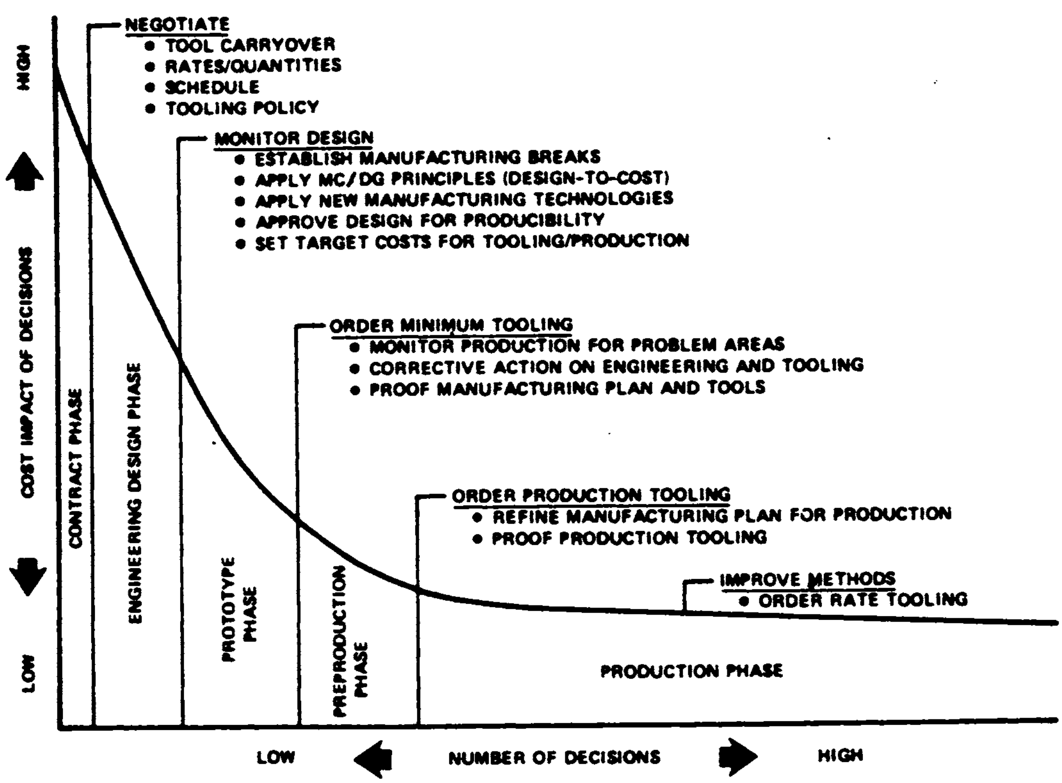


Figure 27. Manufacturing Decisions During Various Design Phases

The MC/DG is a source of manufacturing man-hour data for both metallic and nonmetallic discrete parts and assemblies. It allows the design engineer to evaluate "on paper" the various design approaches envisioned and their relative overall manufacturing cost. The contents of the MC/DG, when fully developed, are shown in Figure 28. Examples of how the MC/DG will be able to aid the design engineer at the conceptual design phase, are:

- Based on long-range material costs and projections of availability, evaluate alternative solutions, such as the application of composites vs. metals.

- Avoid cost-drivers by determining the lowest cost manufacturing methods/processes with available materials.

- Highlight the potential cost impact resulting from designs utilizing strategic metals and alloys such as titanium, cobalt, and chromium. Study design alternatives to identify cost impact early and avoid "built-in" cost escalation.

- Provide management with realistic, timely estimates of projected manufacturing costs.

- Facilitate interaction between manufacturing and engineering at the phase when manufacturing input to the design will have the maximum impact on producibility.

- Provide design engineers with a comprehensive source of data or a tool to enable trade-offs to be conducted between manufacturing costs, including test, inspection, and evluation (TI&E), and the performance of the system.

I	II	III	IV	V	VI
PROCURED ITEM COSTS	MATERIAL REMOVAL COSTS	DETAIL FABRICATION COSTS	MATERIAL TREATMENT COSTS	PERMANENT JOINING* COSTS	ASSEMBLY* COSTS
• Forgings Hand Conventional Blocker Precision • Castings Sand Permanent Mold Investment Die Casting • Extrusions • Materials • Fastener Systems • Emerging Proc. Isothermal Forging Powdered Metal Pultrusion HIP Healing CAM	• Machining Turning Milling Drilling • Chem Milling • EDM • ECM • Emerging Proc. Laser Fluid-Jet CAM EB Cutting	• Metallic Forming Cutting • Non-Metallics Forming Cutting Molding Laminating • Emerging Proc. Superplastic Forming Flow Forming Hydrostatic Forming Thermoplastic Forming CAM	• Heat Treatment • Surface Treatment • Emerging Proc. Laser Treating Nonenvironmental Polluting Treatments	• Welding • Adhesive Bonding • Brazing • Emerging Proc. Diffusion Bonding Weld Bonding Laser Welding Ultrasonic Welding Plasma Arc *Subassembly	• Metallic Assy. Mechanical Fastening • Non-Metallic Assy. Mechanical Fastening • Emerging Proc. Bimetallic Rivets Microwave Curing *Major and Final Assembly

Categories = e.g., Material Removal
Sections = e.g., Machining
Subsections = e.g., Turning and Milling

Figure 28. MC/DG Volume Contents: Manufacturing Technologies for Airframes.

Preliminary Design Phase

During this phase, the final design concept or configuration is selected. If in preparation for a design competition, the "homework" has been accomplished and the customer's requirements have been defined with a reasonable degree of accuracy, this phase constitutes primarily a review of the conceptual design more closely meeting the requirements of the "Request for Proposal" (RFP) at the most competitive cost. During this critical phase, the overall configuration is "frozen." The following are examples of the major decisions which have to be made:

- Performance parameters.
- Envelope the basic configuration and size the design.
- Define weight targets.
- Material selection for major components.
- Major manufacturing subassemblies or break-backs.
- Degree of commonality with existing designs.
- Major "make-or-buy" decisions.
- Major testing requirements, e.g., aerodynamic, acoustic, structural fatigue and materials.
- Optimum master schedule.
- Manufacturing plan/tooling policy.
- Manpower and skill requirements.
- Funding requirements.
- Marketing plan.

It is incumbent on the design engineer to indicate, by providing adequate and factual test data, that the projected performance claims made in the initial proposal to the customer will be met. Of equal importance, is the responsibility of manufacturing to satisfy the customer that the product can be built on schedule, at the contractual price, and without overruns. In today's climate of double-digit inflation and material availability uncertainties, this poses a major problem to both the manufacturer and the customer. The methods of determining cost and their accuracy must be capable of substantiation with factual backup data for both in-house evaluation and customer review.

Current Methods of Estimating Manufacturing Costs

The bases for most cost estimations of products are "historical data," drawn from a data bank. These costs are in categories such as the following:

- Material cost trends and projections.
- Man-hours per pound of structure.
- Standard hour data.
- Learning curve slopes.
- Productivity trends.
- Average tooling costs by type of tool.
- Average cost per part.
- Processing costs.
- Production control costs.
- Test, inspection, and evaluation (TI&E) costs.
- Overhead pools.
- Administrative costs.

MC/DG As a Costing Methodology Aid

Historical data, regardless of the constraints, will continue to be used to estimate manufacturing costs. However, the designer can utilize MC/DG formats such as shown in Figures 29 and 30 to conduct trade-offs between structural performance and manufacturing cost. Such formats showing qualitative data, put the designer on the lowest cost track early in the development phase where the leverage exists to reduce cost.

The MC/DG is not a cost-estimating manual, but rather a guide toward lower cost, enabling designers and manufacturing engineers to avoid cost-drivers. As such, it becomes a tool to evaluate the compatibility of the proposed design with a market, and customer's accepted "bench-mark" or "baseline" of low cost design. A low manufacturing cost for a design utilizing high cost material or with extensive lead-time, may not meet the criteria for affordable performance. The alternatives for different materials, design concepts and manufacturing technologies, can be evaluated with the MC/DG. The design approach is illustrated in Figure 31 for a fuselage shear-panel study. The methodology to define base parts (simplest geometry) and designer-influenced cost elements (DICE) are schematically shown in Figure 32.

Production Design Phase

Although the principal design decisions pertaining to the material to be used on the primary structural components are made in the previous design phase, many important decisions are also necessary during the production design phase.

Detail design of each of the frequently thousands of parts, requires a comprehensive knowledge of the effect of designer-influenced cost elements (DICE) on discrete parts. DICE might add 25 percent to manufacturing man-hours. Selection of material for the detail parts must be compatible with the structural configuration, loading, corrosion, and many other requirements, but the designer frequently has considerable latitude in selecting the material form; e.g., bar, plate, sheet, forging, and casting for metals, or tape or broadgoods for fibrous composites.

MATERIAL	DESIGNER INFLUENCED COST ELEMENTS / BASE PART MANUFACTURING METHOD	STANDARD JOGGLE	FLANGED HOLES	BEADS	HEAT TREATMENT	SPECIAL FINISH	SPECIAL TOLERANCE	LINEAL TRIM	END TRIM	CUTOUTS W/O FLANGES
ALUMINUM	BRAKE FORM	L	L	X	H	L	H	L	L	L
	BRAKE/BUFFALO ROLL	L	L	X	H	L	H	A	L	A
	BRAKE STRETCH	L	L	X	H	L	N	A	A	A
	DIE FORM	N	N	N	N	L	N	L	L	L
	DROP HAMMER	N	N	N	L	L	H	L	X	A
	FARNHAM ROLL	X	L	X	L	L	H	L	X	A
	ROUTED FLAT SHEET	X	L	X	L	L	H	L	X	L
	RUBBER PRESS	N	N	H	N	L	A	L	L	L
	STRETCH FORM	X	L	A	N	L	N	A	X	A
	YODER ROLL	L	L	X	H	L	H	A	A	A
	YODER STRETCH	L	L	H	N	L	N	A	L	A
TITANIUM	BRAKE FORM R.T.	A	L	X	X	L	H	N	H	L
	R.T. BRAKE/HOT STRETCH*	A	L	X	X	L	L	H	H	H
	CREEP FORM*	X	L	X	X	L	L	H	H	H
	FARNHAM ROLL	X	L	X	X	L	H	H	H	H
	HOT PRESS*	N	L	N	X	L	L	N	N	L
	PREFORM/HOT SIZE*	N	L	N	X	L	L	N	N	L
STEEL	BRAKE AND BUFFALO ROLL	A	L	X	N	L	H	H	A	L
	BRAKE FORM R.T.	A	L	X	N	L	H	L	L	L
	BRAKE/R.T. STRETCH	A	L	X	N	L	A	H	L	A
	FARNHAM ROLL	X	L	X	N	L	H	H	L	A
	RUBBER PRESS	N	N	N	N	L	A	L	L	L
	STRETCH FORM	X	L	X	N	L	A	H	A	L

LEGEND

RATING	
X	NOT APPLICABLE
N	NO ADDITIONAL COST INCL. IN BASE PART COST
L	LOW ADDITIONAL COST
A	AVERAGE ADDITIONAL COST
H	HIGH ADDITIONAL COST

Percentage Cost Ranges For Above

L	Up to 10%
A	10–30%
H	Above 30%

*Denotes one or more elevated temperature processing steps.

Figure 29. Typical Format Showing Relative Cost Impact of Various Manufacturing Technologies for Aluminum, Steel and Titanium

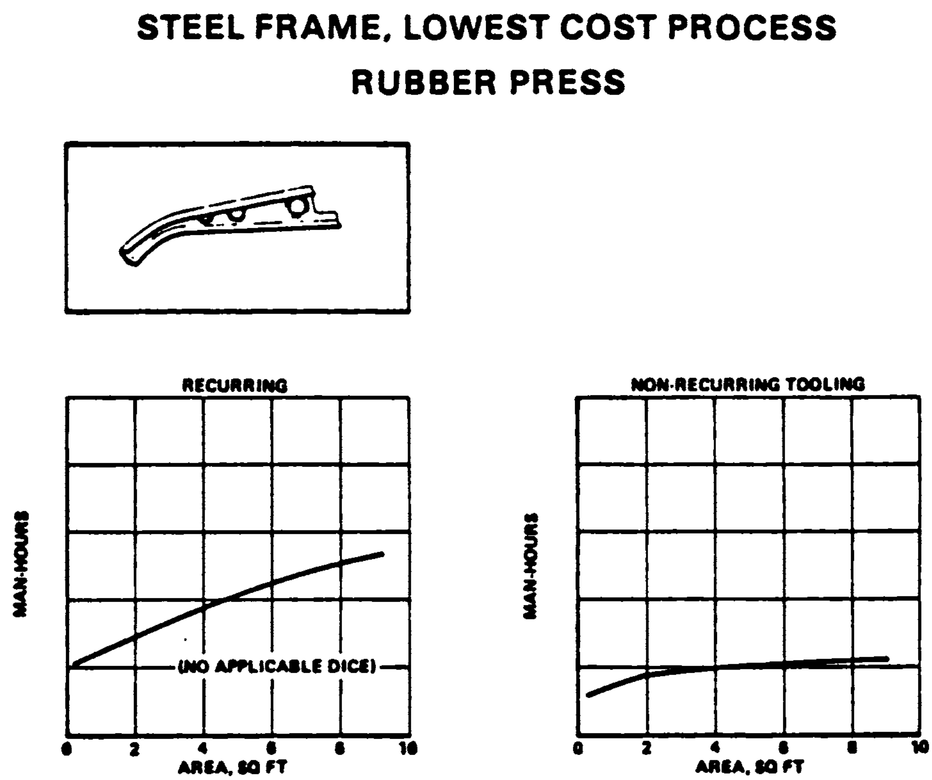


Figure 30. Typical Designer-Oriented Cost-Estimating Format Showing Man-Hours (Values Omitted for Proprietary Reasons)

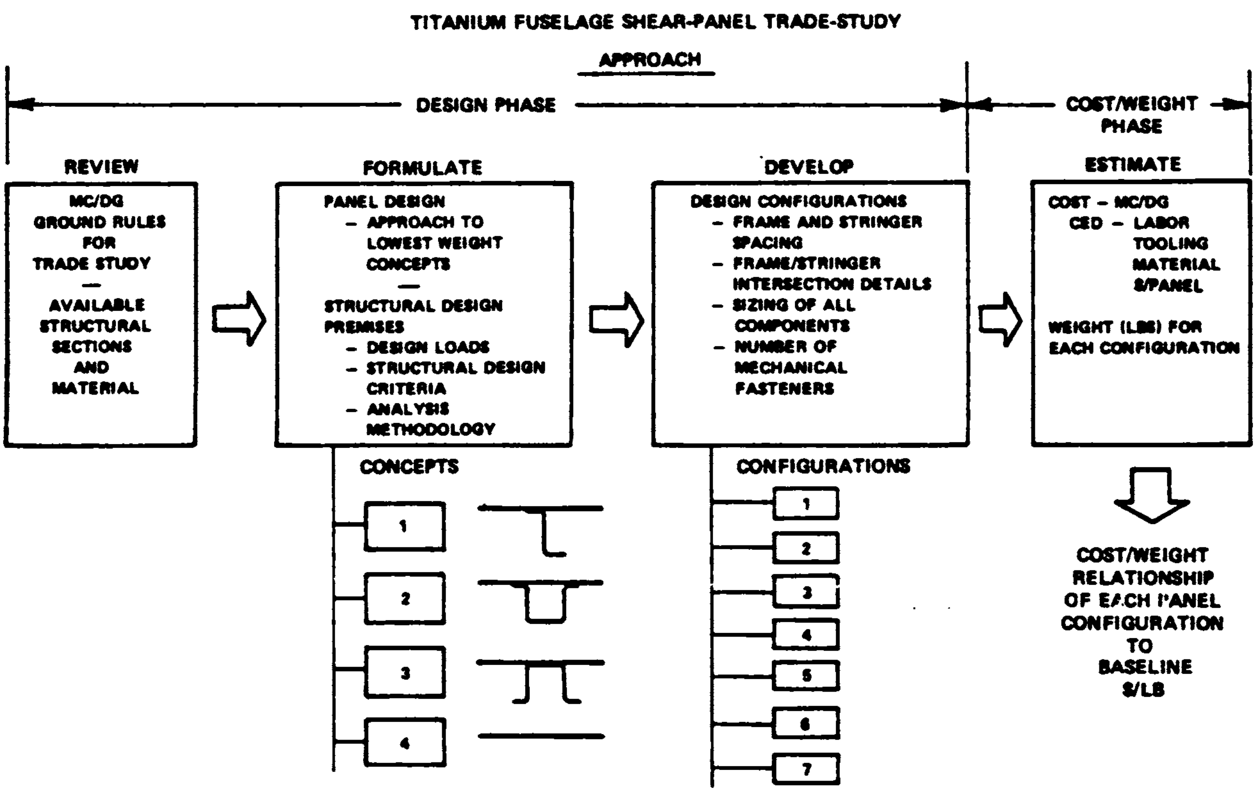


Figure 31. Typical Example of Design Approach Utilizing MC/DG

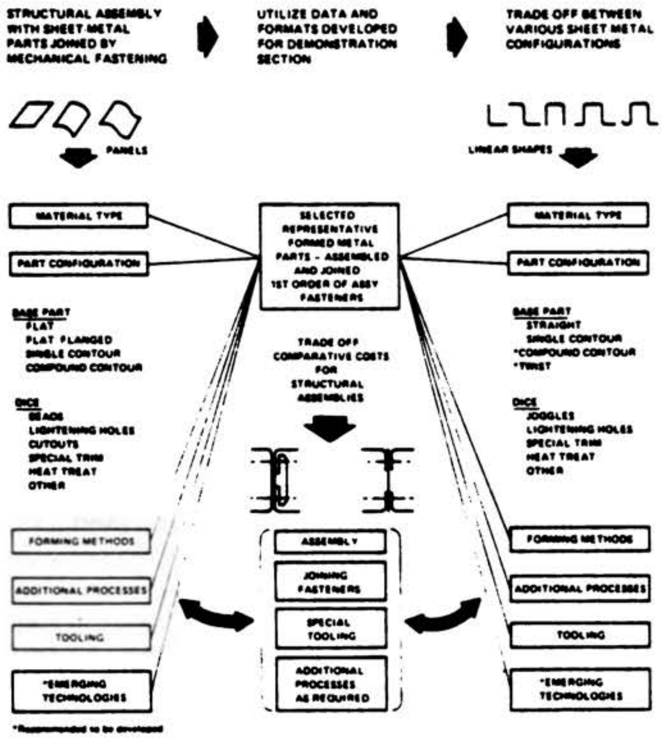


Figure 32. Basepart and Designer Influenced Cost Element (DICE) Approach in Utilizing MC/DG

It is at this critical phase, that the design is "locked-in" and changes made after design release must meet the "form, fit, and function" requirements. This results in added costs for retesting, retooling, scrap or salvage of completed parts, and interruption of the learning curve, resulting in schedule delays.

During this phase, schedules are tight and decisions are made rapidly. The designer-oriented formats in the MC/DG have been especially developed to minimize the possibility of schedule slippage. The MC/DG will be available in hard copy (3-ring binder and pocket version), and also as a computerized data base and interactive computerized system (Manufacturing Cost/Design System-MC/DS).

Production Phase

Provided that an acceptable design has been released and every effort has been made to utilize the lowest cost materials and manufacturing technologies, the product should proceed into production at an acceptable cost level. Unfortunately, this is an ideal situation, but continued application of design-to-cost (DTC) and manufacturing-to-cost (MTC) (Figure 26), as exemplified by the MC/DG, will help to achieve the desired goal of affordable performance.

Most production programs experience a "start-up" and "shake-down" phase to correct engineering and tooling discrepancies and achieve shop learning. This is normal, but the cost and schedule impact can be minimized if corrective action is taken in a timely and organized manner.

An indirect benefit of the MC/DG is the mutual understanding and interaction that results between design engineering and manufacturing.

The MC/DG continues to be a valuable tool throughout the production program. For example, when fully developed, it:

- Provides data necessary to evaluate the cost impact of proposed or necessary changes brought about, for example by changes in system missions.
- Provides justification and a method to evaluate cost impact of necessary changes resulting from the requirement to substitute alternative materials due to shortages, lead-times or increased performance requirements (cost could either be lowered or increased).
- Provides cost analysis data necessary to justify the feasibility of introducing new materials or emerging technologies into ongoing programs.
- Provides a "bench-mark" or "baseline" to document gains in productivity.
- Promotes better industry and customer relations by providing a common baseline or starting point for cost vs. performance studies.
- Provides methods to determine "break-even" points for introduction of, for example, forgings or precision castings, vs. "hog-outs."

COST OF SUBSTITUTION FOR SCARCE RESOURCES IN AN INDUSTRIAL ECONOMY

Robert B. Gordon
Yale University

Many of the materials used in modern industry and agriculture are obtained from mineral resources that we perceive to be finite. We expect, therefore, that they will be exhausted eventually. Two general solutions to this problem have been offered. One is the assertion that the requisite materials can be obtained from the ordinary rocks of the earth's crust as sufficient energy from nuclear power sources becomes available for required concentration and extraction procedures (see, for example, Brown, 1954). The other is that increased technical sophistication will make it possible to find substitutes for materials which become scarce or unavailable. The technical problems involved in making such substitutions are discussed here with a view of working out a methodology that can be used to estimate substitution costs. If there is to be salvation through substitution, not only must the replacement of scarce materials by abundant ones be technically feasible, but the costs of making the substitutions must not become an excessive economic burden.

Importance of Subetitution Costs

Experience in confronting materials shortages throughout the period of rapid industrial growth of the last two centuries supports the idea that technological advances will always make possible substitution of alternative materials for scarce resources. Examples of industrial development curtailed or restricted by shortages of materials resources are confined to restricted geographical areas or short time intervals. It is helpful to begin the discussion by examining some examples of successful substitution for scarce resources.

Charcoal vs. Coal in British Iron Industry

The response of the 18th century British iron industry to dimimishing charcoal supplies and limited water power sources is frequently cited as an example of how technological solutions to problems caused by finite resources develop (Deane, 1965). At the beginning of the 18th century charcoal was the only fuel that could be used to produce pig iron in a blast furnace. The air blast for the furnace was supplied by pumps driven by water power. Pig iron was produced at places where ore, wood for charcoal burning, and water power were all available in the same locality, since transport of bulk commodities over long distances was impractical at that time (Deane, 1965). Water power is a finite (but renewable) resource whose magnitude is determined by the stream flow and fall at the site where power is generated. Wood for charcoal burning is renewable at the forest placement rate, but existing forest stands were being exploited much more rapidly than the replacement rate in 18th century Britain--i.e., the sustainable wood yield was much less than the rate of use. Local shortages of wood for charcoal burning were recognized as early as 1558 in the well-established iron-making districts of Great Britain (Aitchison, 1960) and growth of the industry would have been severely curtailed after 1750 had an alternative fuel not been found.

The technical development required to permit substitution of coal for wood in iron smelting was worked out by Abraham Darby in 1709 (Hyde, 1977). The costs involved in making the substitution are discussed later in this paper; they became favorable for use of the substitute fuel after 1750. The rate of transition from charcoal to coal firing of blast furnaces in Britain after this date can be compared with the law proposed by Fisher and Pry (1975)

$$\frac{f}{1-f} = \exp 2 \quad (t-t_o) \qquad (1)$$

to describe substitutions based on technical change. Here f is the fraction of substitute material used, t is time, t_o the time at which substitution is half complete and is a time constant. Hyde (1977) has tabulated data from which f can be computed during the transition period. The amount of substitution is plotted in the form of Equation (1) in Figure 33; the data are a reasonably good fit to the equation. The "take-over time," the time interval between f = 0.1 and f = 0.9, is found to be 38 years and t_o is 1774. Comparison with the table of take-over times given by Fisher and Pry shows that the transition from the use of charcoal to coke in iron making in the 19th century was fairly rapid even by 20th century standards.

The problem of adequate pumping power for the air blast in iron smelting was also solved by adoption of an existing technological development, in this case one not originating in the iron industry. The steam engine, originally developed for mine pumping, was used to run the air pumps of blast furnaces where adequate water power was not available. Water power continued to be used at older blast furnace installations through the 18th century, but after about 1750 new furnaces used steam-powered blowing machinery. The use of steam power made it possible to locate furnaces near sources of fuel and ore without regard to availability of water power. It also made furnace operation independent of variations in stream flow and eliminated the need to shut down during the summer (low flow) period (Hyde, 1977). The adoption of steam-powered blowing permitted growth of the industry and resulted in cost savings in iron production.

Both substitutions made to overcome limits on growth due to scarce resources in the 18th century British iron industry utilized technology that was available when required and proceeded rapidly once economic conditions were favorable. They took place during a time when the iron industry was expanding rapidly in size and utilized a little-exploited mineral resource, the rich coal deposits of Great Britain.

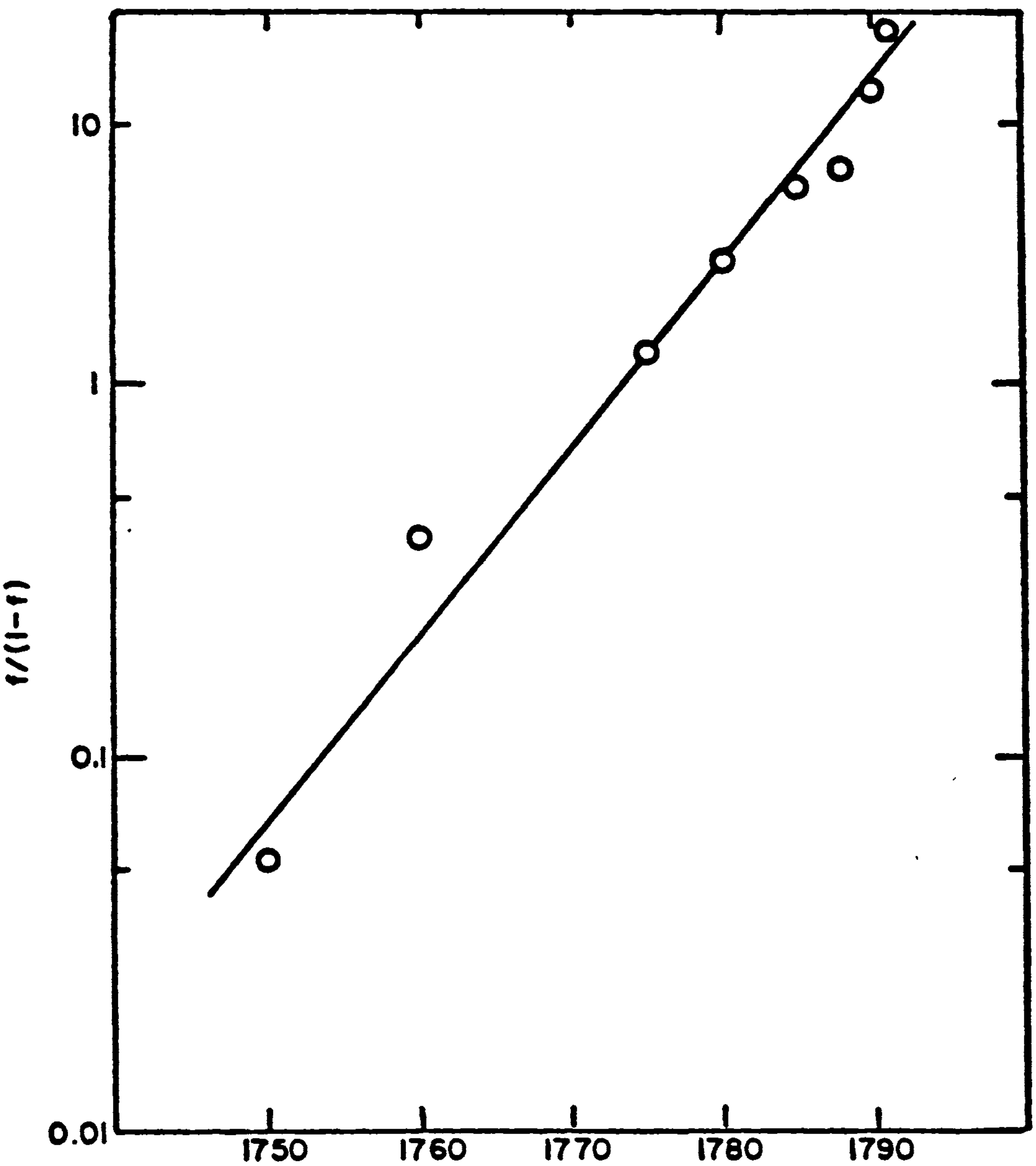


Figure 33. Substitution of Coal for Wood in Iron Smelting in Britain

Molybdenum vs. Tungsten in High-Speed Steel

A more recent example of successful substitution for a scarce resource in the iron industry is the use of molybdenum in place of tungsten in high-speed steel, the steel used to make cutting tools that can operate at the high temperatures developed in metal-cutting machines that achieve high production rates. The early forms of high-speed steel contained up to 18 percent tungsten and were in general use up to the time of World War II. The introduction of 8.5 percent molybdenum makes it possible to reduce the tungsten content of high speed steel to 1.5 percent with no loss of the properties required in cutting tools. A method of protecting the surface of molybdenum high-speed steel during heat treatment so that it would not be degraded by decarburization was developed in 1930 at the Watertown Arsenal (Roberts, Hamaker, and Johnson, 1962). Thus, when a shortage of tungsten developed during the war, it was possible to adopt technological advances already made to effect replacement of a scarce material by a more abundant one. The transition to the use of molybdenum high-speed steel was completed after the war because it is a lower cost product that offers service equivalent to that of the more expensive alloy using tungsten.

Substitution in the Future

Fisher and Pry (1975) list 14 examples of take-over times for the replacement of a scarce by a more abundant material, and there are certainly many more examples that have not yet been examined quantitatively. Near universal success in the past in finding substitutes for resources that have become scarce and at the same time achieving a reduction in costs is the basis of the optimistic view of the possibilities of substitution for materials that may become scarce in the future. This past experience may not be an altogether reliable guide in the years ahead because:

- It is likely that there are some materials used in particular applications in contemporary technology for which there are no possible substitutes.

- Substitutions have been made in the past on a pair-wise basis, as in the replacement of charcoal by coke in the blast furnace or tungsten by molybdenum in high-speed steel. It was not, for example, necessary to consider how the replacement of charcoal by coke influenced the availability of coal for mine pumping. As it becomes necessary to make more substitutions in the future it becomes less likely they can be considered as isolated pairs that can be examined independently of other uses for the materials in question.

- Substitution of alternative materials in complex technological applications often results in unanticipated difficulties and side effects that may greatly raise the cost of making the substitution. Such problems may arise even in simple applications, but they are more likely to occur, and to be more costly, as the complexity of technological systems increases.

Substitution may be impossible in the case of material exploited because of a unique property. Only a few examples are found in modern technology, however. One is the use of silver halides in photographic materials. The sensitivity of the silver halides to light and their response to chemical "developers" depends on a unique set of physical characteristics--the presence of the right combination of defect structure, band gaps, and atomic mobility in AgCl and AgBr crystals (Mott and Gureny, 1940). No substitute materials have been found in many years of exploration for alternatives, and solid state physics offers no promise that they will be found in the future. Diminished availability of silver may be a constraint on the use of photography.

A second example of a unique material application is the use of helium to attain temperatures below about 20°K. Theoretical understanding of the properties of helium at low temperature shows that this is a unique material and holds out no prospect that a substitute refrigerant can be found for the very low temperature range. Recognition of this was responsible for the initiation of helium recovery and storage programs in the U.S. by the Federal Government a number of years ago, even though the programs were uneconomic under existing market conditions. The subsequent demise of these programs was in response to removal of government protection from marketplace economics (Cook, 1979).

A third example is the use of the metal niobium in superconducting alloys. Those with the highest transition temperatures contain large proportions of niobium. Solid state theory does not yet predict the transition temperatures of complex alloys, so a unique role for niobium in this application cannot be proved, but only modest increases of transition temperature have been attained by recent research efforts. Any large-scale use of niobium-containing superconducting alloys would create an immediate supply problem. The helium and niobium problems are coupled. If an alloy with a superconducting transition temperature in the liquid hydrogen range could be found, the future demand for helium as a refrigerant would be much reduced.

It is expected that many of the substitution problems that will arise in complex technological systems will turn out to be coupled. At the time molybdenum was substituted for tungsten in high-speed steel there were few competing uses for molybdenum. That is not true today. Most of the alloying elements used in steel today also have other uses; if the pattern of use of one of the elements in steel-making changes, these other uses will also be affected. The problem of substitution in coupled systems has not yet been addressed, and a discussion is not attempted here. It is reserved for a future paper.

If substitution of a more-abundant for a less-abundant material is technically feasible, it is necessary to know the cost of making the substitution in order to find out if it is advantageous to make the change. Increased technological complexity makes evaluation of substitution costs a more difficult problem, but to make predictions about the economic consequences of reduced availability of materials, accurate assessment of substitution costs is required. The problems that arise in making such assessments are examined next.

Calculation of Costs of Substitution

The cost of substituting one material for another in the manufacture of a given product is made up of the following components:

a. The cost of the substitute materials used, less the cost of the original material not used.

b. The cost of redisigning the product and of the incremental plant and equipment needed to make the product out of the substitute material.

c. The change in the cost of manufacturing the product when the substitute material is used. Cost increments due to changes in cutting, forming, and joining methods are included, as is the cost of additional materials required to make the product when the substitute materials is used in place of the original.

d. The costs that arise from changes in product performance that are a consequence of the use of the substitute material. These include:

   1. Costs that arise from the weight change of the product.

   2. Costs due to changes in the service life of the product.

   3. A cost allowance for unforeseen contingencies that arise from failure of the product to perform as expected.

All of these may be expressed as cost per unit weight of the substitute material used. To express the total transition cost for design and equipment as a unit cost, divided the total cost of changing over to a product made of the substitute material by the output during the life of the new equipment and multiply by a capital recovery factor (which may be taken to be 15 percent at the present time).

The unit cost of substitution, $c_{s,o,i}$, for substitute materials s used in place of original material o in product i is

$$c_{s,o,i} = p_s - p_o \frac{Q_{o,i}}{Q_{s,i}} + f \frac{C_{t,i,s}}{\Sigma Q_{s,i}} + c_{m,i,s} - c_{m,i,o} + \Delta c_{i,s-o} \qquad (2)$$

where

p_s = price of the substitute material.

p_o = price of the original material.

$Q_{o,i}$ = quantity of the original material required to make product i.

$Q_{s,i}$ = quantity of substitute required to make product i.

$c_{t,i,s}$ = transition cost for product i and substitute material s.

$c_{m,i,o}$ = manufacturing cost of product i made from original material o.

$c_{m,i,s}$ = manufacturing cost of product i made from substitute material s.

$\Delta c_{i,s-o}$ = costs due to change in product performance when substitute material is used.

f = capital recovery factor.

and the sum Σ is taken over the time the new equipment produces product i with the substitute material.

One goal of substitution cost analysis is to calculate a switch price, the price at which substitute materials take over the function currently performed by the original material. The switch price, $p^*_{o,i,a}$, for product i and substitute material s is the value of p_o that makes $c_{s,o,i} = o$.

The immediate objective is to work out a systematic way of evaluating the terms in the substitution equation. The evaluation of the ratio $Q_{o,i}/Q_{s,i}$ is based on the idea that there is some property, or combination of properties, that determines the performance of a product made out of a given material in each of its applications. For example, in the electrical wiring of a house it is the resistivity of the metal used to make the wire that determines, for a given current-carrying capacity, the size of the wire that must be used. If a heat exchanger (such as an automobile radiator) is made of a substitute metal having a lower thermal conductivity than that of the material originally used, then the size of the radiator will have to be increased if the same amount of heat transfer is to be effected. A first approximation to the design changes required can be made from comparison of the properties of the substitute and original materials. These determine the amount of substitute material, Q_s, that will be required.

In making these estimates for different uses of a given material it is helpful to classify uses according to which material property, or combination of properties, is most important in each different application. These can be called "engineering-use categories" to distinguish them from the "demand categories" by which commodity markets are described. To illustrate, take the metal copper as an example. A set of engineering-use categories that incorporate nearly all the uses to which this metal is put will include:

- Heat exchangers (thermal conductivity is the dominant property).

- Motors, generators, and transformers (electrical conductivity, ease of forming joints by soldering).

- Pipe (ductility, corrosion resistance, ease of making solder joints).
- Power transmission wire (electrical conductivity, corrosion resistance).
- Communication wire (electrical conductivity, corrosion resistance).
- Machinery (corrosion resistance, strength, machinability, bearing properties).
- Chemicals (toxicity).

Within each of these categories, $Q_{o,i}/Q_{s,i}$ should be about the same. Other terms in the substitution cost that are determined by physical properties may also be nearly the same within one engineering-use category. Before attempting to make generalizations about these terms, it will be useful to examine several examples of determinations of substitution costs.

Historical Example: The Substitution of Mineral Coal for Charcoal in Iron Making

The principal products of the iron industry in the 18th century were pig iron and bar iron. Pig iron was produced from ore in a blast furnace and could be cast into useful articles (such as pots). It is a moderately strong material but, because of its high carbon and silicon content, is also brittle and cannot be mechanically worked, as by rolling or hammering. Malleable iron, known as bar iron, was made in the 18th century by reducing the carbon and silicon content of pig iron through a series of processes carried out at an establishment known as a "forge." Early in the 19th century about half of the pig iron production was used for making bar iron; 100 years later this had increased to 70 percent. Before 1709, charcoal was the only fuel that could be used in the iron-making blast furnace, while after that date and until about the end of the century both charcoal-produced and coal-produced pig iron were available to the forge operator. Either type of starting material could be used to make bar iron, but a process modification was required when the starting material was changed. The charcoal-fired blast furnace operated at a lower temperature than the coal-fired furnace and yielded pig iron with a lower silicon content; less fuel was required to convert this pig to bar iron at the forge.

For this example consider the substitution of coal-produced for charcoal-produced pig iron as the starting material for making bar iron. In the cost of substitution equation, let

p_s = cost of coke-produced pig iron (excluding capital costs).

p_o = cost of charcoal-produced pig iron (excluding capital costs).

$Q_{o,i}$ = tons of charcoal-produced pig iron required to produce 1 ton of bar iron.

$Q_{s,i}$ = tons of coke-produced pig iron required to produce 1 ton of bar iron.

c_t = capital cost of coke-produced pig iron less the capital cost of charcoal-produced pig.

Δc_m = fuel and other costs in making 1 ton of bar iron at the forge from coke-produced pig less the fuel and other costs of making 1 ton of bar iron from charcoal-produced pig.

$\Delta c_{i,s-o}$ = costs that arise from differences in the properties of bar iron made from coke-produced and charcoal-produced pig.

Conversion of charcoal-produced and coke-produced pig in the forge yield bar iron with equivalent properties. Hence, $\Delta c_{i,s-o} = 0$. The remaining costs can be evaluated from data published by Hyde (1977). The manufacturing costs of making bar iron from the two types of pig are given in Hyde's table for the time interval 1730-1739. Conversion of coke-produced pig requires more starting material and more fuel than does conversion of charcoal-produced pig. Other costs are the same. Hence Δc_m = £1.24/ton, the cost of the additional charcoal fuel used in the forge. Coal came into use in forges later than in blast furnaces and in a variety of different processes during the late 18th century. Estimates of c_m are not available for the processes and would probably be quite difficult to make. We will assume that, for our purposes, a sufficiently accurate estimate of the manufacturing cost difference during the transition period can be made by correcting the 1730-39 estimate for the rise of price of charcoal used in forges in successive decades. (These prices are given by Hyde; capital costs for coke-fired blast furnaces are also listed). For the charcoal-fired furnace we use his estimate of £0.66/ton and assume this to be constant through the 18th century. The ratio $Q_o/Q_s = 0.92$, according to the data from Hyde. The prices p_s and p_o are also from Hyde. These data have been used in the following computation of substitution cost (Table 9):

Table 9. Cost of Substitution in the Manufacture of Bar Iron

Time Interval	P_2	P_o	Q_o/Q_s	C_t	c_m	c_s
1730-1750	5.50	5.03	0.92	0.54	1.24	2.85
1750-1760	3.36	5.37	0.92	0.14	1.46	0.02
1760-1770	2.44	6.29	0.92	-0.16	1.65	-1.86
1770-1780	2.71	7.02	0.92	0.34	1.80	-1.61

Note: all costs are expressed as £/ton of iron

The steady rise in p_o and in c_m during the time interval shown in the table is due principally to the rise in the price of charcoal, which started to increase rapidly after 1750. This increase resulted from rising demand and increasingly restricted supplies. During the 1750's, p_s was substantially lower than p_o, but c_s was still positive. Despite the lower cost of coal-produced pig iron, it was still not the preferred material for making bar iron because of the higher capital manufacturing costs its use entailed. Data showing the relative amounts of two types of pig iron used for making bar iron at forges during the transition period are not available but, since most pig iron produced during the transition period was converted to bar iron, it is likely that the transition curve shown in Figure 33 applies and the mid point of the transition interval was 1774. The data clearly show that, although a substitute material was available because of a technological advance made in 1709, and a shortage of wood for charcoal making had been an official and public concern for many years, the substitute material was not adopted on more than a local basis until c_s became negative.

We turn now to examples of substitution costs in the more complex technology of modern industry and examine two illustrations of the costs involved in replacing copper by aluminum. Copper is a geochemically scarce element, while aluminum is abundant (Skinner, 1976) and has properties that make it the most likely substitute for copper in many applications.

The Cost of Substitution for Copper in Automobile Radiators

Suppose that aluminum is to be substituted for copper in the manufacture of automobile radiators. The function of a radiator is to transfer heat from the cooling fluid circulated through the engine to the ambient air. The heat transfer effected through the radiator depends on the temperature difference between the coolant fluid and the air, on the thermal resistance of the radiator. This resistance is the sum of resistances due to boundary layers at the fluid-metal and air-metal interfaces and the resistance of the metal wall of the radiator, which depends on the thickness of the metal and its thermal conductivity. The thermal conductivity of aluminum is 54 percent less than that of copper. An aluminum radiator made to the same pattern as a copper one would have a greater thermal resistance; if it is to perform the same service as the copper radiator, it must be made larger. How much larger depends on the relative magnitudes of the different contributions to the thermal resistance. If resistance due to heat flow through the metal is much larger than the other sources of thermal resistance, the increase in size can be estimated by assuming that:

- The operating temperatures and wall thickness in the two radiators are the same.
- Changes in the fluid-flow characteristics in the two radiators can be neglected.
- It is heat flow through the metal walls of the radiator that determines the amount of heat transferred.

Under these assumptions an aluminum radiator yielding equivalent cooling would have a radiating surface 1.9 times greater than that of the equivalent radiator made of copper. The additional area could be realized by increasing the cross-sectional area of the radiator, or its length, or both in appropriate proportion. We will now estimate the incremental costs involved in making and using this radiator.

a. Materials Cost. The average 1978 price for copper was $1.45/kg and of aluminum, $1.17/kg. The density of copper is 8960 kg/m^3 and that of aluminum, 2700kg/m^3. In a straight volume substitution, 1 kg of aluminum replaces 3.32 kg of copper. Under the assumptions listed above, a radiator made of copper need be only 54 percent as big as one made of aluminum to effect the same heat transfer. Hence, in this application 1 kg of aluminum (cost $1.17) replaces only 1.78 kg of copper (costing $2.56). The materials cost increment for this substitution in radiators, therefore, is $1.41/kg.

b. Design and Tool Cost. Data for this estimate were supplied by the G & O Manufacturing Company of New Haven. They estimate that it would cost $2.5 million to redesign their product and retool their operation to produce an aluminum radiator. In the 10 years after the transition they would produce 2.0 x 10^7 kg of aluminum radiators. The design and retooling cost is calculated as the investment required, times a capital recovery factor of 0.15, divided by the mass of the radiators produced annually. This is $0.19/kg.

c. Manufacturing Costs. Since both copper and aluminum are easily cut and formed, we assume no increment in the cost of these operations when the substitution is made. Copper radiator parts are joined by soldering and aluminum parts by vacuum brazing. The cost of the brazing equipment is figured into item b above, and we assume that there is no incremental cost in joining operations by the two methods once the requisite equipment is in place and operating. No additional materials are required to make a radiator of aluminum rather than copper. Thus, the manufacturing cost increment is nil.

d. Service Costs. An aluminum radiator is 44 percent ligher than an equivalent-service copper radiator; its use in a vehicle may result in improved fuel economy. We assume that when a lighter radiator is used, compensating design changes are made that keep the vehicle performance unchanged. The Office of Fuel Economy, National Highway and Traffic Safety Administration, asserts that, under these conditions, the relation between distance traveled per gallon of fuel, f, and vehicle effective weight, w, is

$$f = k\,w^{-0.8}$$

where k is a constant. The effective weight of the vehicle is the curb weight plus 300 lb. of passenger weight. The value of k was determined from the average of 1979 U.S. passenger car weight and fuel consumption, viz., 3400 lb. and 19 mpg. If there is no other change in vehicle weight--i.e., no change in the total weight of coolant or in the vehicle structure to accommodate the aluminum radiator--then f is increased by 0.029 mpg. If fuel costs $1/gal and the vehicle is driven 100,000 miles, the fuel cost is reduced by $1.72/kg of aluminum used in the radiator. If the vehicle life is five years, the saving is reduced by 1.1^{-5} = 0.62 and is $1.07/kg.

Alternative assumptions are possible. For example, if there is no compensating design change to retain performance when the substitution is made, the exponent in the above equation for f becomes -0.4. The saving in fuel cost due to the use of the aluminum radiator is then only \$0.50/kg. If it is necessary to change the size of the vehicle to accommodate the aluminum radiator (which is larger than the equivalent copper radiator), there may be an increase rather than a decrease in vehicle weight. Suppose, for example, that radiator thickness is increased from 1 in. to 1.86 in. and that the vehicle length is increased 0.86 in. to allow space for the new radiator. If the principal structure of the vehicle were 120 in. long and weighed 3400 lb., the increment in length would add approximately 24 lb. to its weight and the net weight change when the aluminum radiator was used would be a 17 lb. increase. There would then be an increase in total fuel cost over the life of the vehicle of \$2.24/kg.

Practical experience that shows how long aluminum radiators will last in service is not now available. It is anticipated that the potential for corrosion problems in aluminum radiators is higher than it is in copper ones, but in the absence of data we will assume that the service life of aluminum radiators can be made to be as great as that of copper ones and charge no cost increment for a change in service life. However, an allowance has to be made for the possibility that this expectation will not be fulfilled and for the fact that the repair of aluminum radiators will cost more than the cost of radiator repair work done now (primarily because the aluminum cannot be soldered). To allow for this we include a contingency cost of 25 percent of the cost of making the aluminum radiator--i.e., \$0.33/kg.

The substitution cost of aluminum for copper in automobile radiators under the two sets of assumptions discussed is shown in Table 10.

Table 10. Cost of Substitution of Aluminum for Copper in Automobile Radiators

P_s	P_o	Q^o/Q_s	C_t	Δc_m	Δc_{s-o}	c_s	NOTES
1.17	1.45	1.78	0.19	0	-1.07+0.33	-1.96	1
1.17	1.45	1.78	0.19	0	2.24+0.33	1.35	2

NOTES: All in \$/kg
1. No size change in vehicle.
2. Vehicle size increased to accomodate larger radiator.

The most important term is that due to the change in vehicle operating expense due to the change in radiator weight. The assumptions made are not necessarily the most extreme estimates of this cost increment that could be made. If metal resistance is not the dominant source of thermal resistance in the radiator, Q_o/Q_s will be larger and the size increase of the radiator will be smaller. This will make c_s more favorable for substitution. The results show that a reliable estimate of c_s in this case requires a more detailed analysis of technical detail than has been presented here, as well as assumptions about the life and operating characteristics of the vehicles in which the substitutions are made.

In the next example these difficulties do not come up.

Substitution of Aluminum for Copper in Home Wiring

The condition of equivalent service from the two materials in this application is that the electrical resistance and length of the wire remain the same when the substitution is made. Let

R = electrical resistance = $\rho\ell/A$.

ℓ = length of wire used.

A = cross-sectional area of the wire.

ρ = electrical resistivity.

W = weight of wire per unit length = Ad.

d = density of the metal used.

The subscript "A" is for aluminum and "C" for copper. From the conditions that $R_A = R_C$ and $\ell_A = \ell_C$ we find

$$\frac{W_A}{W_C} = \frac{A_A d_A}{A_C d_C} = 0.53$$

when

$$\rho_C = 1.65 \times 10^{-8}\Omega_m \qquad d_C = 8960\ \text{kg/m}^3$$
$$\rho_C = 2.88 \times 10^{-8}\Omega_m \qquad d_A = 2700\ \text{kg/m}^3$$

In a typical house, 300 m of #14 AWG copper wire is used in the electric power distribution system. This weighs 5.6 kg and can be replaced by 3.0 kg of aluminum wire. The cost saving on material in one home is then $4.61 at average 1978 prices. If direct substitution of aluminum for copper wire is made with no other changes in the electrical system, there is no redesign cost. The same, or similar, wire-drawing equipment can be used for aluminum as for copper, and the difference in manufacturing cost will be small. The weight difference between copper and aluminum results in no change in performance in house-wiring. The other cost that needs to be considered, however, is the contingency cost for potential failure of the aluminum wire to function as well as expected in the house-wiring application.

Aluminum was first used for house wiring in the 1960's. Between 1.5 and 2.0×10^6 homes had all-aluminum wiring by 1972 (Newman, 1975). By 1970 it was recognized that there was a serious safety problem in many aluminum-wired homes due to overheating in wall receptacles, switches, and junction and panel boxes. In some cases, temperatures high enough to ignite an adjacent wooden structure were attained, and several serious fires have been traced to overheated junction boxes.

Four factors contribute to excessive heat generation at junctions between aluminum and other metals carrying electric current (Mittleman, 1969; NBS, 1974). These are:

- High electrical resistance due to failure to break the oxide film on the aluminum wire when a mechanical joint is made.
- Creep of the aluminum wire in mechanical joints, which results in a loss of contact pressure.
- Expansion of the aluminum wire due to heating, causing accelerated creep.
- Electrolytic corrosion at dissimilar-metal contacts.

High temperature at a junction with aluminum wire may result even when current well below rated circuit capacity is flowing. Hence, fuses and circuit breakers do not provide protection against overheating due to high junction resistance. Junction failures are a direct consequence of physical properties of aluminum that are different from those of copper, viz:

- Rapid formation of a hard, continuous oxide on a freshly exposed metal surface.
- Lower melting temperature, T_m, so that at room temperature the ratio T/T_m is higher for aluminum than copper. (Creep rate scales as T/T_m and so is relatively greater in aluminum wire.)
- Lower yield strength of the pure metal.

Satisfactory junctions with aluminum wire can be made, but have to be designed to make allowance for the difference in properties of aluminum and copper.

One method of estimating the contingency cost for failures in aluminum house wiring is to use data for the number of receptacle repairs and replacements determined by inspection of a sample of homes built in 1969-70 and surveyed in 1977-78 by the Franklin Institute (Anon. 1979). They report:

	Al (WBS)*	Cu (WBS)*
Number of repairs per home-year	0.060	0.018

*WBS = wiring binding screw; i.e., Al or Cu wire secured under a binding screw in the receptacle.

Suppose that in a typical home there are 30 electrical receptacles. Then in an aluminum-wired home there will be 10 receptacles per kilogram of aluminum wire used in the electrical system, and in a copper-wired home there will be 5.4 receptacles per kilogram of copper wire. Assume that the life of the home electrical system is 50 years. The number of receptacles that will require repair or replacement during the service life of the electrical wiring is then:

For Al wire: 50 x 0.060 = 3/house.
For Cu wire: 50 x 0.018 = 0.9/house.

In the aluminum-wired house, 10 percent of the receptacles will require attention, while the copper-wired house only 3 percent will. A local electrical contractor estimated that the average cost of repair or replacement would be about $10. Hence, the contingency cost difference required to allow for the different junction failure rates experienced with the two types of wiring is:

For Al wire:	10 x 10% x $10 =	$10/kg
For Cu wire:	5.4 x 3% x $10 =	1.6/kg
		$8.4/kg

This is a contribution to the cost of substitution that must be included when no cost allowance is made for redesign of the home wiring system and the aluminum wire is installed in the same way as copper wire so that there is no allowance for an installation cost increment. The substitution cost is shown in Table 11. An argument that Δc_{s-o} should be even larger can be made on the grounds that the cost of having wiring out of service during the time that repairs were made was not included. The cost of loss of property and life in fires originating at overheated junctions has not been included either.

Table 11. Aluminum for Copper Substitution Cost in Home Electrical Wiring

p_s	p_o	Q_o/Q_s	c_t	Δc_m	Δc_{s-o}	c_s
1.17	1.45	1.88	0	0	8.4	6.94

Note: All expressed in $/kg.

At the time aluminum wire was first offered as a substitute for copper wire there was an immediate cost advantage in its use. The results presented in Table 11 show that, had the full cost of using aluminum without design changes being made in household electrical receptacles been known, the decision to use aluminum wire would not have been taken. In a recalculation of c_s under today's conditions, allowance would have to be made for the cost of design and manufacture of receptacles compatible with aluminum wire and the costs of testing and certification of its use. Estimates of these costs have not yet been made.

Discussion and Conclusions

Contributions to the cost of substitution from factors other than materials costs are important in all three of the examples examined. In the substitution of coke-produced for charcoal-produced iron, the increment in manufacturing cost was the most important of these extra charges. Bar iron is a relatively simple metallurgical material, and it was not made to very precise specifications in the 18th century. Exact specification of properties was not required by iron users of the time, and the mechanical systems they were building were relatively insensitive to variations in material properties. The lack of an entry for contingency cost in Table 9 is made with the benefit of hindsight, but it seems clear that such costs were much less likely to be large in the technology of the 18th century than in the technology of today. Costs that enter Δc_{s-o} are expected to be increasingly great as the technical sophistication of society increases. Technical systems are increasingly interdependent, so that an unanticipated failure in one system is more likely to raise costs in others. Despite the availability of a great deal more technical expertise today, it would be hard to make a convincing case that the consequences of making changes in technical systems can be better foreseen today than in the past. There is no difficulty in compiling a list of technical systems that have failed to perform as expected and have incurred costs well above those designed for. The problems that arose from the use of aluminum in house-wiring provide an excellent example.

A second generalization that may be drawn from the examples examined here is that, even when substitution of one material for another is technically feasible, the substitution will not be made until there is an economic advantage that can be seen in the framework of cost information perceived by those responsible for making the decision on which material to use. In the British iron industry example, the decision between coke-produced and charcoal-produced iron was made by individual ironmasters who were concerned with both fixed and variable costs; the transformation to the widespread use of coke-produced iron did not begin until the total costs recognized by ironmasters were favorable. That their understanding of the costs was reliable is shown by the fact that the transformation was completed and became the accepted practice of the industry. In the case of the substitution of aluminum for copper in automobile radiators, c_s may be positive or negative, depending on the assumptions made. (The choice of assumptions among those arguing this problem may depend on noneconomic factors.) The industry seems to be in a state of indecision about this particular substitution today. The decisions made to use aluminum house wire in the 1960's were taken by persons primarily concerned with $p_s - p_o\ (Q_o/s)$. This appreciation of the costs of substitution was seriously incomplete, and the use of aluminum house wire at the time proved to be a false start. (The use of aluminum electrical wire in homes is now prohibited by some building codes.) The actual cost of making the substitution, when it is eventually done, will be much greater than originally assumed.

The principal sources of uncertainty in estimating cost of substitution are evaluation of all of the costs due to change in the performance of the product made with substitute material and the proper allowance for contingency costs. The uncertainty in determining both is likely to increase as more substitutions are forced by material shortages and as substitutions are made in more complicated systems.

REFERENCES

Aitchison, L. 1960. A History of Metal, London, MacDonald and Evans, p. 443.

Anon., National Controlled Study of Relative Risk of Overheating of Aluminum Compaired with Copper Wired Electrical Receptacles in Homes and Laboratory, Franklin Research Center, Report F-C4812-01, 1979.

Brown, H. 1954. The Challenge of Man's Future, Viking Press, New York.

Cook, E. 1979. The Helium Question, Science 206, pp. 1141-1147.

Deane, Phyllis. 1965. The First Industrial Revolution, Cambridge University Press, pp. 104-105.

Fisher, J. C. and R. H. Pry. 1975. A Simple Substitution Model for Technological Change, Mineral Resources and the Environment, U.S. National Academy of Sciences.

Hyde, C. K. 1977. Technological Change and the British Iron Industry 1700-1870. Princeton University Press, Table 4-6.

Mittleman, Joseph. 1969. Connecting Aluminum Wire Reliability. Electronics, p. 94.

Mott, N. F., and R. W. Gurney. 1940. Electronic Processes in Ionic Crystals, Oxford, Chapter VII.

NBS Quarterly Report on Aluminum Wiring; Combined Report from the Measurement Engineering Division and Center for Building Technology; p. E-17, 1974.

Newman, Rae. 1975. Hazard Analysis of Aluminum Wiring, U.S. Consumer Product Safety Commission NIIC-0600-75-H006.

Roberts, G. A., J. C. Hamaker, and A. R. Johnson. 1962. Tool Steels, American Society for Metals, p. 8.

Skinner, B. J. 1976. A Second Iron Age Ahead?, Am. Sci. 64, pp. 258-269.

TECHNOLOGICAL SUBSTITUTION:
REVISITING THE METHODOLOGY

Harold A. Linstone
Futures Research Institute
Portland State University
Portland, Oregon

The purpose of this paper is to review briefly the forecasting methodology in the context of technological substitution. The discussion is divided into four parts: data-based, model-based, Delphi, and multiple-perspective concepts. Sufficient references are included to permit the interested reader to delve more deeply into this very active field.

Data-based Methods

Traditionally we have looked to scientists and technologists for forecasts of technology and substitution of the new for the old. By far the most common method they have used is trend extrapolation--according to mathematician Eric Bell one of the two great evils we inherited from the Greeks.

Figures 34 and 35 present two trend extrapolations showing a technological characteristic varying with respect to time. The first is taken from an early book on technological forecasting (Ayers 1968), the second from the most recent issue of Science (Kear 1980). The points to be emphasized in such time series are these:

a. They are data-based--i.e., they rely on empirical input from the recent past.

b. They assume continuation of the system structure--elements and interactions--as well as environment from the past into the future.

c. Macro-system forecasts tend to be more reliable then micro-system forecasts.

An implication of a and b is that trend extrapolations are conservative and tend to lack imagination--they perceive the future like a driver moving the car forward with his eyes on the rear-view mirror. The scientist is comfortable with data, but there exist no future data, so he is constrained by the available data to an extension of the past. This is usually reasonable for near-term or short-range forecasts, but can prove hazardous for long-range predictions.

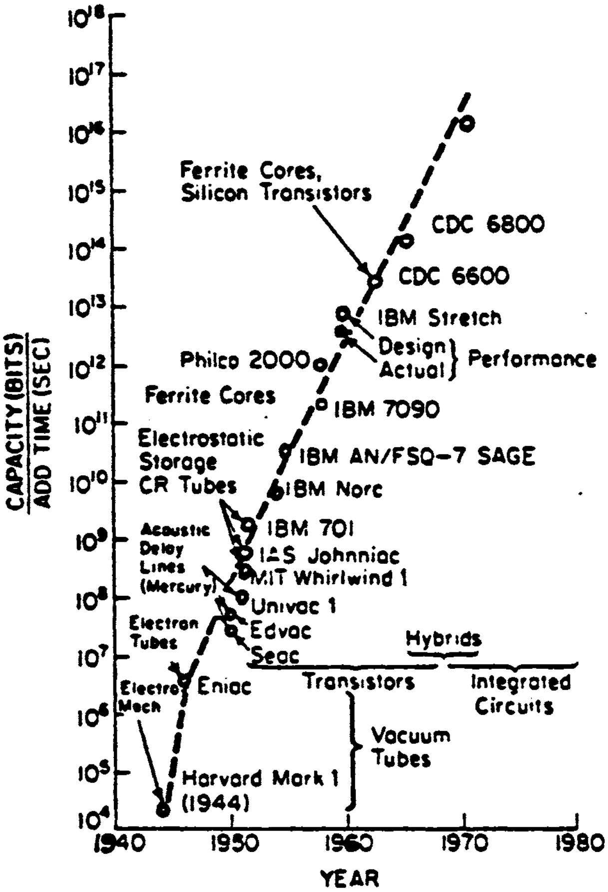


Figure 34. Computer Performance

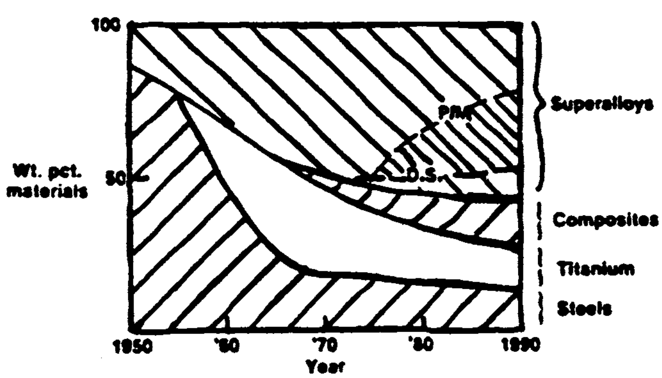


Figure 35. Weight Percent of Materials Used in Most Advanced Aircraft Gas Turbine Engines

Forecasts are highly sensitive to the, often opaque, core assumptions. Ascher (1978) has demonstrated that a forecast using a very crude methodology and sound core assumptions is usually more accurate than one with a very sophisticated methodology and inappropriate core assumptions. The lesson is obvious: the assumptions need at least as much attention as the data search and collection.

An implication of b and c is that envelope curves are more likely to indicate forthcoming substitutions than study of any one component. In Figure 36, for example, neither the cyclotron nor the betatron curve offers a good basis for anticipating synchrotrons. An envelope curve comes much closer to suggesting that the cyclotron and betatron will face substitution by a new system which is not an extension of either.

The macroview is more concerned with the function or task to be done, while the microview deals with a specific means of accomplishing that task. Thus in 1929 aeronautical engineer Nevil Shute Norway firmly predicted that by 1980 commercial aircraft would be limited to a cruising speed of 110 to 130 mph and a range of 600 miles (Ayers 1969). He assumed that we would rely on propeller aircraft for the next 50 years. But a forecast of air transportation systems need not--in fact, should not--have made such a restrictive assumption.

The forecaster concerned with technological substitution should always carefully examine the tacit assumptions underlying the forecast.

Models

The difference between unlimited exponential and limited logistic curve growth (Figure 37) can be viewed from the perspective of substitution. The former neither motivates substitution nor represents its pattern. If a material is in unlimited supply, exponential growth is theoretically possible (although practically impossible on a finite earth). If it is in limited supply or has saturated the market, we are usually concerned about substitution. We either face an inability to fill the demand in the face of resource shortages and rising prices (e.g., wood fuel), or we seek to create a new market by making a readily available product or process obsolete (e.g., black and white television).

The logistic (or Pearl) curve presents us with a surprisingly useful model of the technological substitution process. The underlying assumption is that the rate of adoption of a new product or process is proportional to the fraction of the old one still in use. Mathematically we express this relationship as follows: if f is the fraction of the market captured by the innovation, t the time, and b a constant, then

$$\frac{1}{f}\left(\frac{df}{dt}\right) = b(1-f)$$

If t_o is the time at which the fractional substitution reaches its midpoint--i.e., f = 1/2--intergration yields the familiar logistic curve

$$f = \frac{1}{1+\exp b(t-t_o)}$$

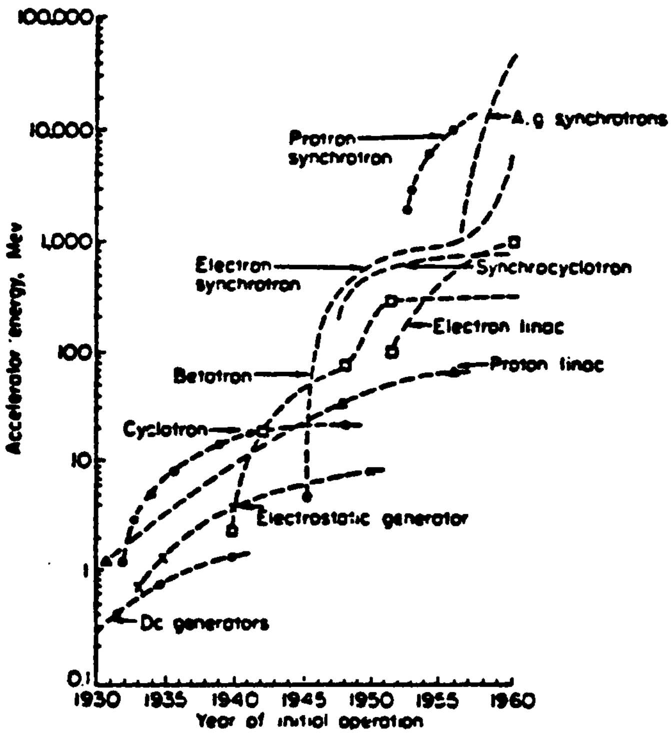


Figure 36. The Rate of Increase of Operating Energy in Particle Accelerators

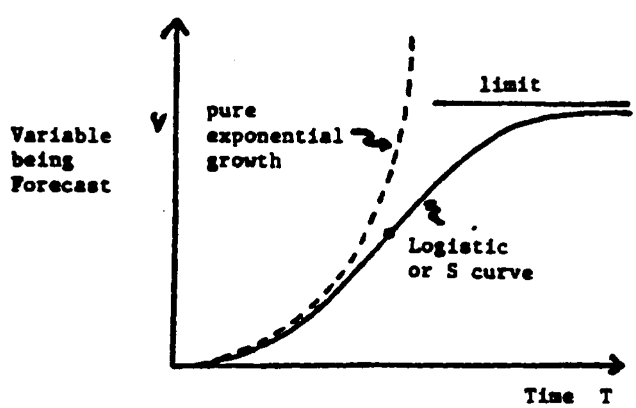


Figure 37. Growth Curves

Fisher and Pry (1971) have analyzed 17 cases of substitution (Table 12) and found striking support for this model once the data are normalized (Figure 38).

Blackman (1972) and Floyd (1968) have developed variations which can, together with Fisher and Pry, be incorporated in the more general equation

$$(1-\sigma)\left[\ln \frac{f}{F-f}\right] + \sigma \left[\ln \frac{f}{F-f} + \frac{F}{F-f}\right] = C_1 + C_2\, t.$$

where F = upper limit of the market share f,

C_1, C_2 = constraints

σ = dimensionless factor, $o \leq \sigma \leq 1$.

The case $\sigma = 1$ corresponds to Floyd's model, $\sigma = 0$ to Blackman's model, and $\sigma = 0$, $F = 1$ to the Fisher-Pry model. Sharif and Kabir (1976a) find that $F = 1$ is usually a reasonable assumption (i.e., 100 percent substitution) and permits simplification of the preceeding equation to the form:

$$\ln \frac{f}{1-f} + \sigma \frac{f}{1-f} = C_1 + C_2 t.$$

Here the second term on the left side of the equation is a "delay factor" and a "delay coefficient." Sharif and Kabir provide guidelines for selecting the proper delay coefficient.

Another interesting variation stemming from Fisher and Pry has been put forward by Marchetti (1977) in his work on energy substitution. He is concerned with more than two energy sources competing at one time for the market--e.g., a mix of coal, oil, gas, and nuclear energy. This situation is sometimes known as "multilevel substitution." As he recalls,

> I had to extend the treatment slightly with extra stipulation that <u>one</u> of the fractions is defined as the difference to 1 of the sum of the others. This fraction follows the [Fisher-Pry] equation...most of the time but not always. It finally shows saturation...The fraction dealt with in this way corresponds to the oldest of the growing ones. The rule can be expressed in the form: first in - first out.

He is then able to fit more than a century of data nearly perfectly using only two constants for each energy source, the dates for 1 percent and 50 percent market fraction (Figure 39). Neither war, wild price oscillation, energy price, depression, nor available reserves seem to disturb the basic trend of substitution. The U.S. and world energy substitution trends are shown in Figure 40 beginning in 1850 and becoming a forecast in 1970. The transparency of the model permits examination of various future options: introduction of solar or fusion energy, faster nuclear penetration, a nuclear moratorium, etc. The method illuminates several key points for Marchetti: (a) the robust internal dynamics of the energy substitution process (e.g., the coal share started to decline around World War I despite infinite reserves); (b) the slow, inexorable pace of the process; (c) criticality of the initial push, the take-off, for the long-term pattern of an energy source; and (d) the value of a macroscopic view of the substitution process (suggested earlier).

Table 12. Takeover Times (ΔT) and Substitution Midpoints, T_o, for a Number of Substitution Cases.

Substitution	Units	Δt Years	t_o Year
Synthetic/Natural Rubber	Pounds	58	1956
Synthetic/Natural Fibers	Pounds	58	1969
Plastic/Natural Leather	Equiv. Hides	57	1957
Margerine/Natural Butter	Pounds	56	1957
Electric-Arc/Open-hearth Specifically Steels	Tons	47	1947
Water-based/Oil-based House Paint	Gallons	43	1967
Open-hearth/Bessemer Steel	Tons	42	1907
Sulfate/Tree-tapped Turpentine	Pounds	42	1959
TiO_2/PbO-ZnO Paint Pigments	Pounds	26	1949
Plastic/Hardwood Residence Floors	Square Feet	25	1966
Plastic/Other Pleasure-Boat Hulls	Hulls	20	1966
Organic/Inorganic Insectides	Pounds	19	1946
Synthetic/Natural Tire Fibers	Pounds	17.5	1948
Plastic/Metals Cars	Pounds	16	1981
BOF/Open-Hearth Steels	Tons	10.5	1968
Detergent/Natural Soap (U.S.)	Pounds	8.75	1951
Detergent/Natural Soap (Japan)	Pounds	8.25	1962

Note: Δt is the time from 10 to 90 percent takeover.

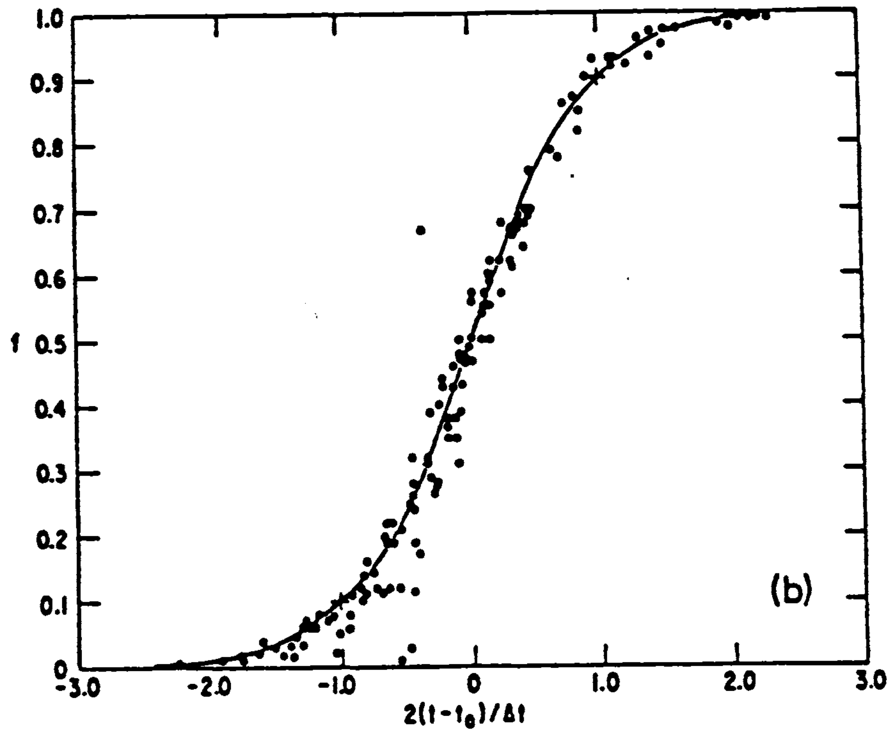


Figure 38. Fit of Substitution Model Functional to Substitution Data for all 17 Cases vs. Normalized Units of Time

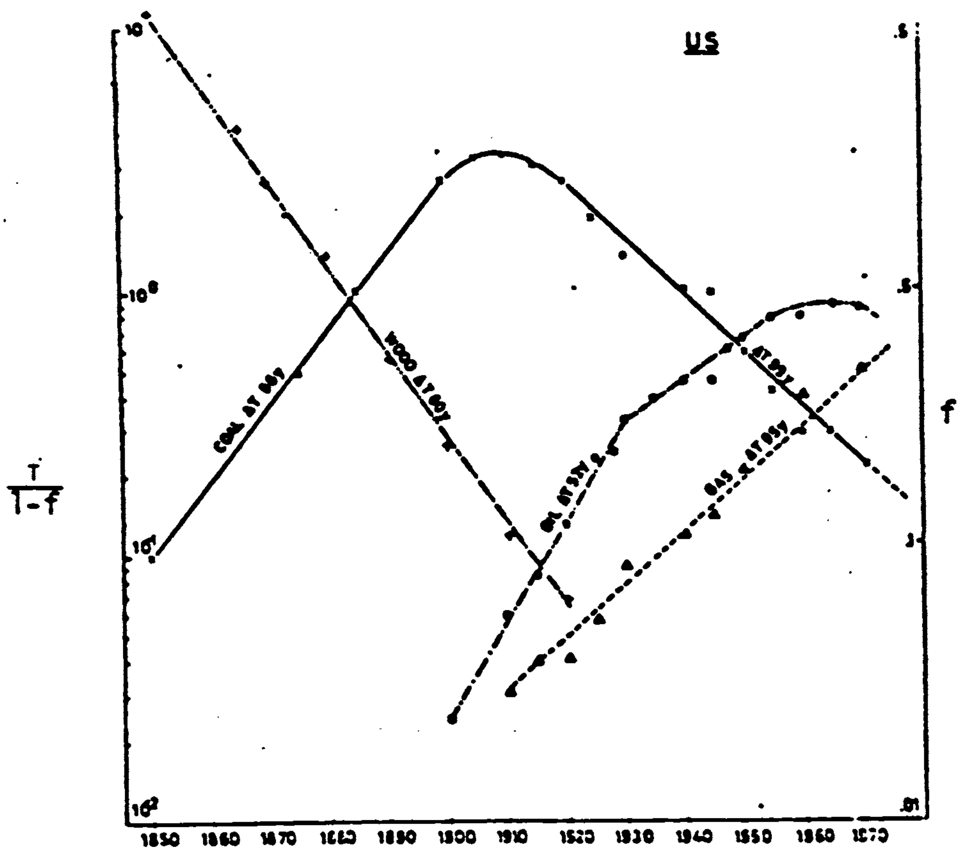


Figure 39. Fitting of the Statistical Data on Primary Energy Consumption in the U.S. Straight Lines are Represented By Equations of Type 2. Rates of Penetration are Indicated by the Time To Go From 1 Percent to 50 Percent of the Market (ΔT years). The Knee In the Oil Curve and the Saturation Regions can be Calculated by the Rule "First In-First Out"

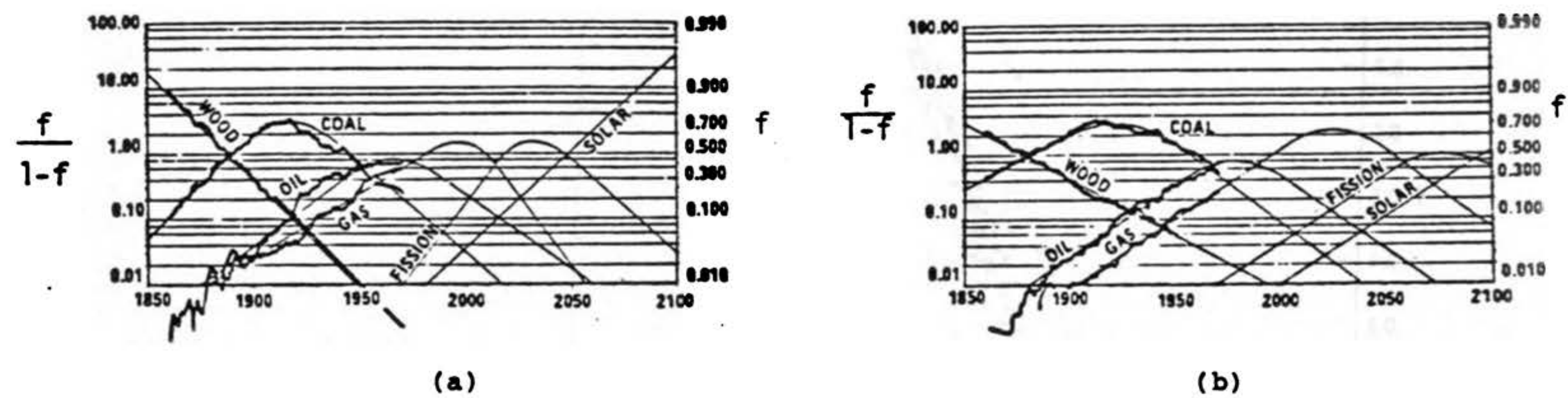


Figure 40. Energy Market Penetration History and Projection

Like Ascher, Marchetti appreciates the value of relatively simple models. However, a balanced survey requires consideration, even if briefly, of more complex models. A favorite approach of modelers with an engineering background is system dynamics. First applied outside electrical engineering to industrial organizations by Forrester, this type of model gained great popularity with the Club of Rome's "Limits to Growth." Sharif and Kabir (1976b) have addressed the multilevel substitution process using system dynamics as shown in Figure 41. They apply it to commercial aircraft propulsion systems, a mix of piston, turboprop, and turbo engines.

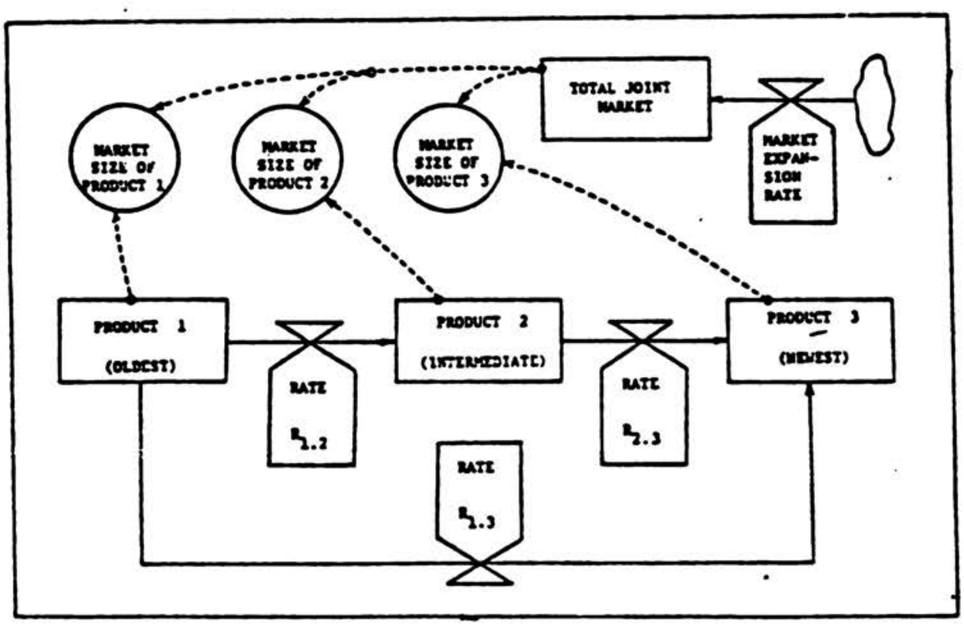


Figure 41. Simplified System Dynamics Diagram Showing the Basic Structure of the Multilevel Substitution Model

Economists have also developed a variety of substitution models, beginning with Zvi Griliches (1976) and Edwin Mansfield (1961). The latter for example, develops a deterministic model which computes the rate at which a new product innovation displaces an existing product in a given market as a function of (1) the proportion of the firms already using the new product, (2) the profitability of the new product relative to the old product, and (3) the investment needed to adopt the new product.

Blackman, Seligman, and Sogliero (1976) extend this work by developing an innovation index which indicates the relative tendency of various industrial sectors to innovate. The index is derived from various input variables which reflect the extent to which resources are allocated to achieve innovation and output variables which measure the extent to which new product and process innovation is achieved.

Stern, Ayres, and Shapanka (1976) take into account the fact that alternate technologies do not compete merely on the basis of price, but rather on the basis of the "utility" they provide. Further, the markets for which they compete may not be the same. These authors develop a model which includes numerous indicators of utility and market. They illustrate their concept with a case study of the substitution of plastic for glass in bottles. Table 13 indicates the range of indicators used.

Table 13. Utility Analysis Attributes for Plastic/Glass Bottle Substitution

<u>Attributes of Concern to Bottlers</u>

Chemical inertness
Heat resistance
Nonpermeability
Pressure resistance
Flexibility of color and shape
Adaptability to existing bottling equipment

<u>Attributes of Concern to Consumers</u>

Transparency
Tradition
Reusability
Convenient dispensing

<u>Attributes of Concern to Bottlers and Consumers</u>

Unbreakability
Light weight

<u>Attributes of Social Concern</u>

Disposability
Recyclability

Not surprisingly, input-output models are also used. Their advantage is the use of real, measurable physical or monetary quantities. The substitution alters the values, and the impacts can be readily traced. Ayres and Shapanka (1976) use four steps: identification of trends that can be expressed as substitutions, translation of this information into a form necessary for an input-output model, use of logistic curves to obtain trend extrapolation, and modification of existing input-output coefficients in conformance with trend extrapolations.

We may also obtain insights from models not directly concerned with technological substitution. One example will suffice. Economists such as Kuznets and Kondratieff have developed models to explain business cycles. Recently Mensch (1979) has found interesting relationships between technological innovations and the Kondratieff 50-year cycle model. This "long wave" consists of the sequence prosperity-recession-depression-revival. Table 14 presents the scheme according to Van Duijn (1977).

Table 14. Kondratieff Cycles

Prosperity	Recession	Depression	Recovery
1783-1803	1815-1826	1826-1837	1837-1847
1847-1866	1866-1875	1875-1884	1884-1893
1893-1913	1921-1929	1929-1938	1938-1949
1949-1967	1967-1975		

Mensch points out that basic invention occurs at a fairly steady pace while basic innovation experiences strong surges. For example, computers, radar, television, the atomic bomb, jet engines, and automatic automobile transmissions were basic innovations which clustered in a relatively brief time span--when the U.S. emerged from the depression and commenced its economic recovery. At such times there is a willingness to take risks and initiate major new capital investment. As the recovery continues and prosperity commences, the emphasis shifts to product improvement rather than basic innovation. Prosperity reaches its peak, and excess capacity leads to layoffs and recession. Figure 42 illustrates the phenomenon. At the very least this model suggests the need to consider different levels of substitution--e.g., total system and subsystem.

Delphi

Earlier we focused on data-based forecasting methods and on model-based techniques. These are the accepted modes of inquiry in science and technology. Delphi is a technique which, in addition to data- and model-based insights, permits the introduction of intuition. It is thus more subjective. Delphi may be described as an iterative questionnaire procedure which involves a group or panel while maintaining individual anonymity. It is sometimes termed a remote conferencing process. The key words are "iterative," implying here several rounds of questioning with feedback, and "anonymity," denoting nonattribution of responses to individuals. The advantages of the procedure

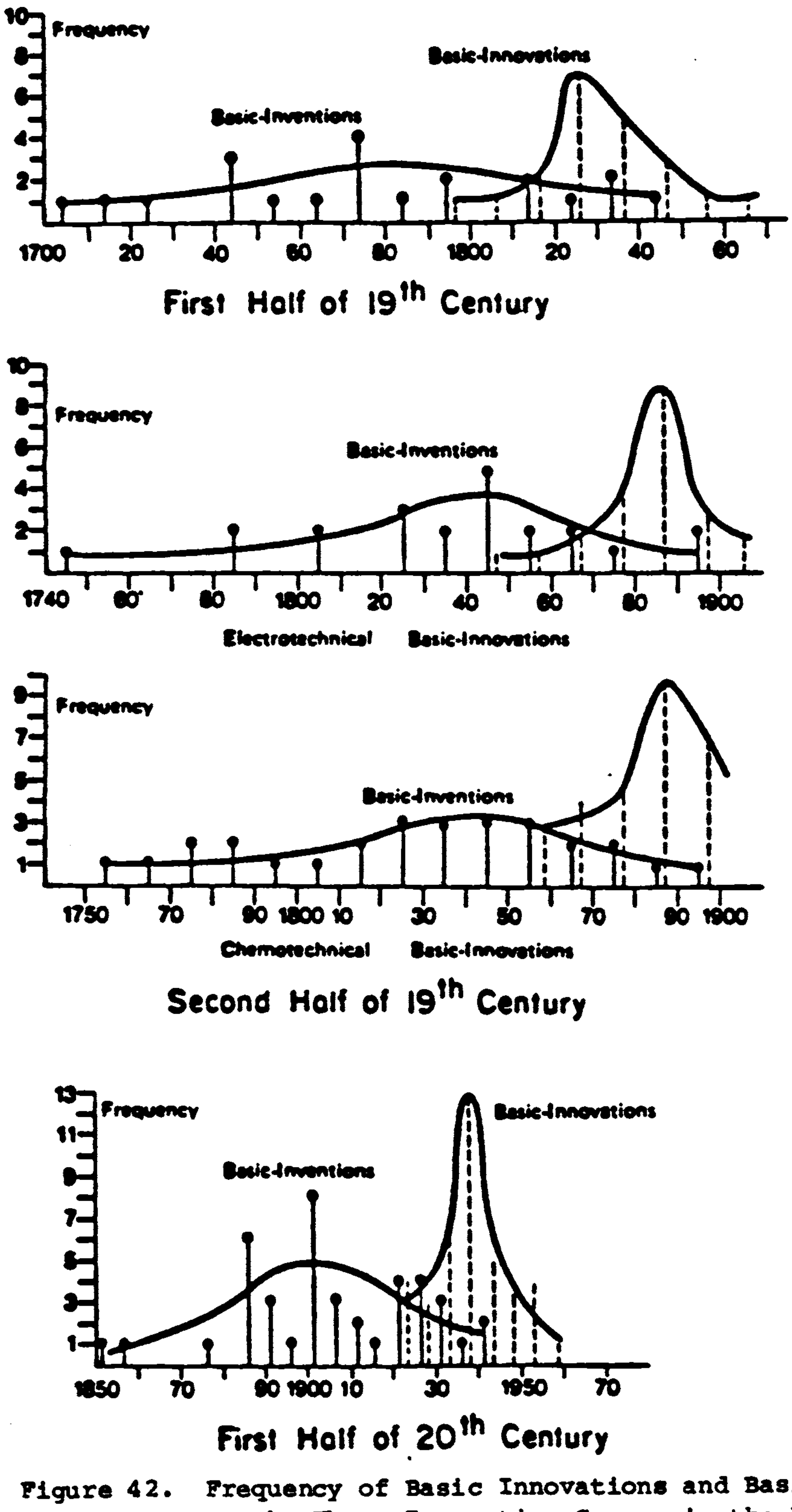


Figure 42. Frequency of Basic Innovations and Basic Innovations Corresponding to the Three Innovation Surges in the 19th and 20th Centuries (Source: Mensch)

are (1) structuring in a way often not possible with a conference, (2) avoidance of dominance by strong personalities or "big names" in the field of inquiry, and (3) ability to draw in widely dispersed individuals who cannot be brought together in a conference at one time. The first published Delphi study (done at the RAND Corporation) appeared in 1964; since then, many technological forecasting studies have been done in industry and think tanks using this technique (Linstone and Turoff 1975).

TRW Corporation's PROBE is a major in-house Delphi involving panels of scientists and engineers in 14 different fields, ranging from manufacturing processes to ocean technology. The materials panel, for example, includes 25 TRW technologists and developed forecasts in subareas such as composite materials, metals and alloys, photovoltaic and optical materials, coating, lubricants, and crystals. The format for all panels requires judgments concerning (1) desirability, (2) feasibility, (3) probability of occurrence (ever), and (4) timing of occurrence on the assumption that the event will occur. In the case of (1) and (2) an index from -1 to +1 is used.

Typical results from the materials panel after three rounds:

(a) Lubricants, seals, and adhesives

Conventional fasteners will be replaced by high-strength adhesives in 50 percent assembly applications.

Desirability	.09	Year of occurrence	10 percent chance 1975
Feasibility	-.14		50 percent 1978
Probability	.61		90 percent 1985

(b) Composite Materials

Molded synthetic materials--e.g., reinforced plastics--will replace steel forgings for the first time in steering-linkage components (sockets, pitmans, idlers).

Desirability	-.13	Year of occurrence	10 percent chance	1976
Feasibility	-.20		50 percent	1982
Probability	.46		90 percent	1990

A second example of a Delphi is the Plastics and Competing Materials study for 1985, undertaken by Enzer in 1970 (Linstone and Turoff 1975). The study focuses on material property changes likely to affect widespread material usage. The Delphi panel of experts is first presented with descriptions of major uses, properties, and proprietary qualities of 37 plastics and 16 nonplastics in widespread uses. The panelists are asked to identify likely changes in the properties which would affect the use of the materials by 1985 and to add new materials likely to be available at that time. The format and typical feedback of the initial round are illustrated in Figure 43. Information presented to the panel is in Roman type, feedback in italics. Items noted as being in "Package No. 2" are reassessed in Round 2 in greater detail. The format and typical results of this round are shown in Figure 44. In addition, overall U.S. plastics production is estimated by the panel.

EXISTING GENERAL PURPOSE & SPECIALTY PLASTICS

CODE: Property changes anticipated by the panel, likely to affect the widespread use by 1986, are noted in *italics* as follows:

Direction of expected change	% of the panel	40 or more	20 to 40	less than 20	None
	Improvement	+ + +	+ +	+	
	Degradation	− − −	− −	−	

1. MATERIAL & Typical uses	2. Proprietary qualities (relative to typical uses): (A) Assets (L) Liabilities	3. Property: Price	Processability	Tensile strength	Stiffness	Impact strength	Hardness	Useful temperature range	Chemical resistance	Weather resistance	Water resistance	Flammability	4. Comments
ACRYLICS Windows; fiber optics; building panels; lighting; tubing	(A) optical clarity; weather resistance (L) abrasion resist.	2 +	3	2	3	1 + +	2 +	1 + +	2	3	2	1 + + +	*(Included in package No. 2).*
CELLULOSICS Packaging; film; toys; telephones; instrument glass	(A) tough; clear (L) abrasion resist.	2	3	2	2	2	3	2 (−)	1	2	1	1 +	• *Not competitive with low cost plastics, e.g., vinyls.*
CAST EPOXY Printed circuits; potting compounds	(A) strong & flexible	2 + +	2	3	2	1 + + +	3	3	• 2 +	2	1 + +	1 + + +	*(Included in package No. 2).*
IONOMER Molded housewares; toys; extruded tubing; sheeting; packaging	(A) transparent, tough & flexible; chemical resistance (L) strength; temp. range	2 + +	3	1 +	1 +	3	1 + +	1 + +	2	2	2	1 + +	*(Included in package No. 2).*
MELAMINES & UREAS Dishes; wood laminates; appliance cabinets; electrical devices	(A) appearance (finishability); surface hardness (L) impact strength; temp. range	2	2	2	3	1 + +	3	2	2	2	2	2	
PHENOLICS Appliance cabinets & parts; bonding resins; electrical parts	(A) cost; strong, hard, rigid; abrasion resist. (L) chemical resist.	3	2 + +	2 +	3	1 + +	3	2 +	1 + + +	2	2	1 + + +	*(Included in package No. 2).*
LOW DENSITY POLYETHYLENE Dishes; bottles; pipes; tubing; film packaging	(A) flexible; cost (L) strength; weatherability; flammability	3	3	1 +	1	3	1	2	3	1 + +	3	1 + + +	• *Compounding can improve weather resistance.*

Degradations encircled indicate panel disagreements with the original rating not a forecast.

3 — Outstanding in property indicated; among the best performers available.

2 — Acceptable performance in this property; still suitable in most cases.

1 — Not acceptable if indicated property is important to intended use.

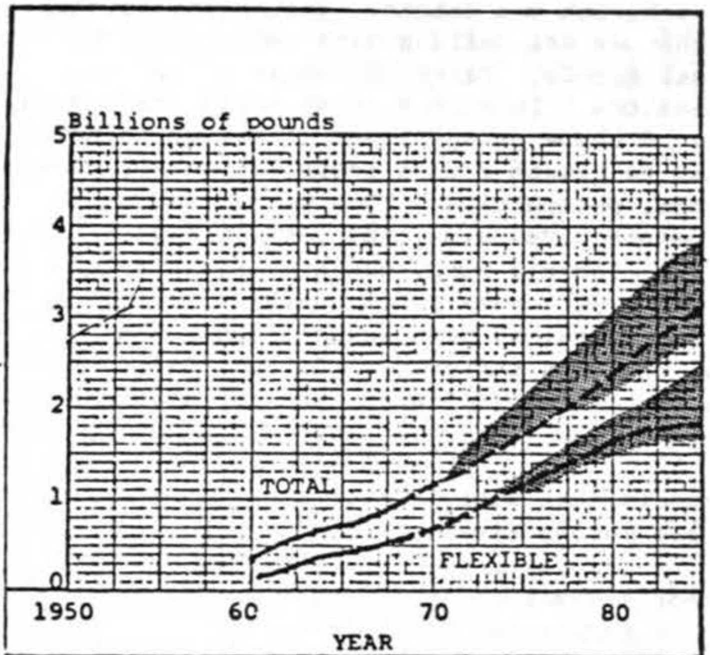
Figure 43. Typical Feedback of Results of Initial Estimates of Changes in Existing Materials

ENGINEERING PLASTICS

Materials for Which the Panel Anticipates A Significant Increase in Use by 1985			Comments by the Panel	Agree	No Consensus	Disagree	Applications		
Key Properties in Which Improvements Are Considered Likely by the Panel	Current Property Value	Estimate Property Value Likely by 1985					Market Breakdown (major uses)	Volume-Million lb/yr Current	Volume-Million lb/yr Estimate 1985
1. ABS.			Reduction in depolymerization will improve flammability.		x		TOTAL VOLUME	508	*1500*
Price, ($/lb)	0.28-0.44	*0.20-0.30*	Alloying and blending will improve temperature range and flammability.	x			Automotive	80	*250*
Maximum Service Temperature *(D 648) °F	180-245	*250*	Improved shaping of disperse particles as well as processing techniques will orient particles.	x			Major Appliances	60	*150*
							Pipe & Fittings	60	*200*
Flammability, in./min.* (D 265)	1.0-2.0	*0.65-2.0*	Reduction in compound prices; scale and competition lowering prices.	x			Bus. Machines, Phones	40	*90*
Impact Strength Notched Izod, ft. lb/in. *(D 256) ● 73°F	2.0-10.0	*2.5-11.0*	Availability of composite forms as sheet and ability to fabricate in inexpensive equipment will make this material competitive with glass-polyester.	x			Recreational Vehicles	45	*150*
● 40°F	0.8-3.5	*0.9-5.0*					Luggage	27	*65*
							Other:	196	*350*
Deflection temperature, °F *(D 648) ● 264 psi	214-244	*214-280*	Platability makes this material increasingly attractive for automobile parts, this will be especially important if low temperature strength and crack resistance can be improved.	x			*Interior panels & sheeting*		
● 66 psi	215-250	*215-280*							
			Properties will be very dependent upon filler.						
			ABS will include chemical and cross-linked materials.						
*ASTM Test Method									

Figure 44. Typical Results of Final Assessment of Important Changes in Existing Plastics

A graph of production (including distribution among major markets) from 1960 to 1970 is presented to the panel and an extrapolation to 1985 is requested. Typical results are shown in Figure 45. Solid lines represent statistics, dashed lines, the median forecast, and shaded areas, the interquartile range (i.e., 50 percent of responses within shaded area). Clearly Delphi can be used to probe technological substitution possibilities.



MARKETS	U.S. PRODUCTION (MILLIONS OF LB/YR.)	
	CURRENT	BY 1985
Appliances	58	130
Construction	206	1,050
Furniture	270	825
Packaging	170	475
Transportation	200	575
Other	90	440

Figure 45. Total U.S. Production-Foamed Plastics

Multiple Perspectives

First, we took a "technical" perspective in addressing technological substitution. Next, we introduced intuition, unavoidably subjective. We now open the breach further. In the early days of technological forecasting it was customary to emphasize a dichotomy--exploratory forecasting (what can we do technologically?) and normative forecasting (what should we do?) It was always assumed that what will be done is dependent on both, the "can" and the "should." Technological substitution seems to fit the pattern comfortably. It was further assumed that the rational approach inherent in the technical perspective--objectivity, reductionism, cost and benefit criteria, optimization, etc.--was the only valid one.

Yet there were nagging doubts: Why were key decisions apparently not based on such rational analysis? Why were reasonable alternatives never considered? Even more questions were raised as technology assessment became more formalized. Forecasts were no longer enough; there must be impact assessment and policy analysis. And there is inevitably feedback: the policy analysis can alter the forecast.

This has led us to the concept of multiple perspectives, of augmenting the technical, rational actor approach with two others: the organizational and the individual. We emphasize that we are dealing here with different paradigms, not different mathematical models. Table 15 compares the three perspectives along a number of dimensions. It should be stressed that we may look at a given technology from any of these perspectives--i.e., we do not mean to imply that a technology must be viewed from a technical perspective or an organization from an organizational perspective. Thus we may see a technology from an organizational perspective (e.g., the Detroit auto manufacturers' lack of enthusiasm for electric cars, the Army Staff's long resistance to the substitution of the superior M-16 for the M-14 rifle). Or we may see an organization from a technological perspective (e.g., the cybernetic theory of organizations, decision analysis). Herbert Simon or Cyert and March view organizations with a very different set of paradigms than do Machiavelli or Franklin Roosevelt.

So, too, we suspect that technological substitution also may--in fact often should--be viewed from multiple persepctives.

Figure 46 describes the concept schematically. We are currently testing the feasibility of its application to technology assessment for the National Science Foundation (Linstone et al 1980). One of the cases under study is the substitution of guayule for heavea as a natural rubber source. Examples of factors bearing on the substitution, categorized by the perspective providing the insight, are given in Table 16. Every application to date has tended to support the hypothesis that the organizational and individual perspectives are valuable in augmenting the technical perspective. And this effort unveils reasons for the frequent ineffectiveness of conventional systems analysis (e.g., computer simulations) and of technological forecasting and assessment in the context of private and public policy decisions.

Table 15. Three Perspectives

	TECHNICAL (T)	ORGANIZATIONAL (O)	INDIVIDUAL (P)
WELT ANSCHAUUNG	Science-technology	Organization	Psychology-behavior
CHARACTER-ISTICS	Cause effect	Cause-effect & challenge-response	Challenge-response
	Objective	Objective & subjective	Subjective
	Problem solving	Problem avoidance/delegation	Game-in-process for most
	Analysis	Analysis & synthesis	Intuition
	Prediction	Action/implementation	Fear of change and unknown
	Optimization	Satisfying	Creativity and vision by
	Use of averages, Probabilities	Standard operating procedures	Partial rationality
	Trade-offs	Porochial priorities	Inner world/self
	Complete rationalility	Factoring/Fractionating problems	Maslow hierarchy of needs
		Incremental change	Learning
		Recognition of partial unpredictability	Power/influence/dominance
	Left neocortex	Left and right neocortex	Left and right neocortex
PREFERRED INQUIRING SYSTEM	Locken-data	Heglian-dialectric	Intuition-noumena
	Leibnizian-model	Singerian-pragmatic	Merleau-Oonty-negotiated
	Kantian-multimodel		
TIME CONCEPT	Technological time	Social time	Biological time
	Zero discounting	Moderate discounting	High discounting

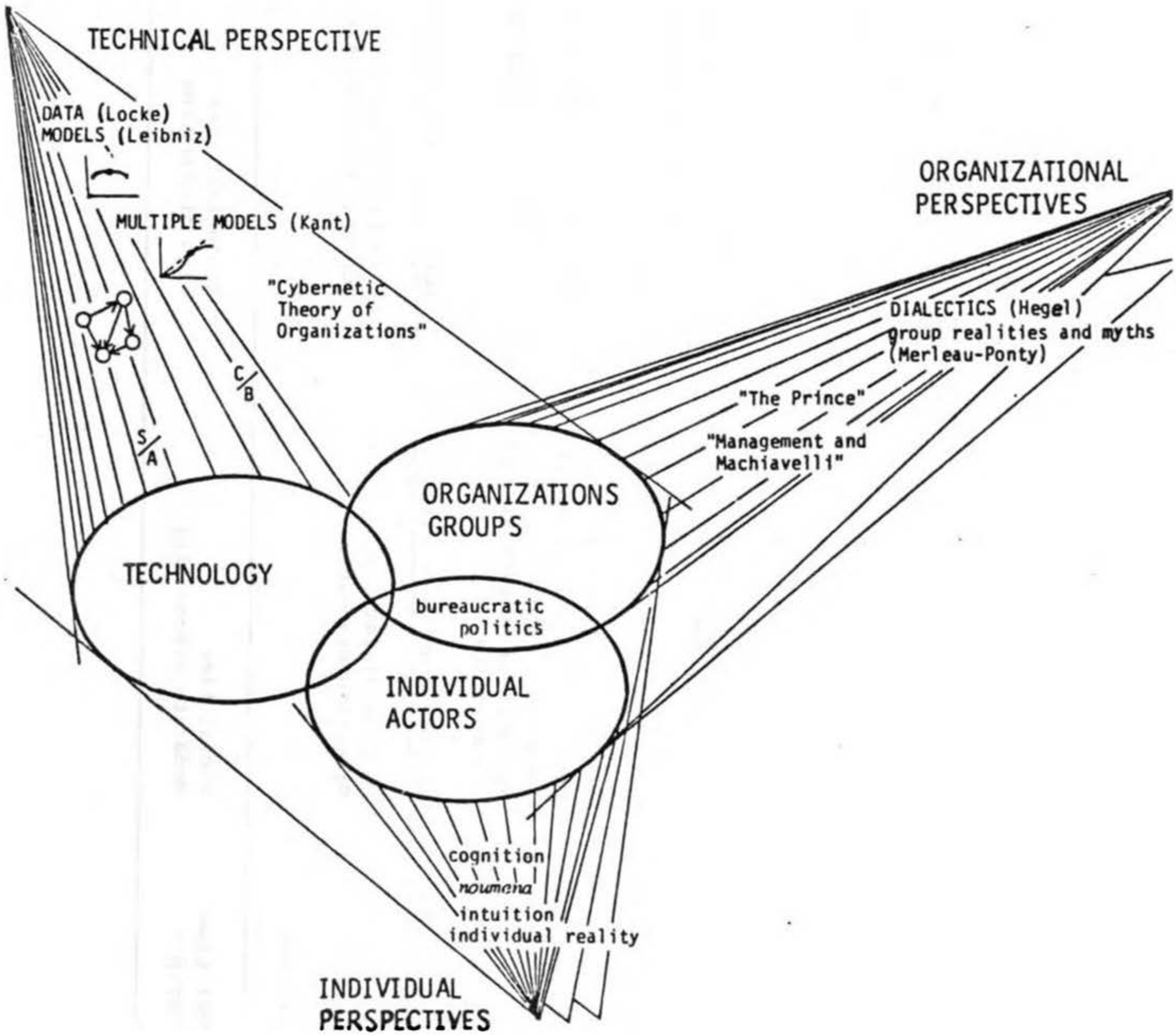


Figure 46. Multiple Perspectives-A Singerian Inquiring System

Table 16. Examples of the Use of Multiple Perspectives on the Guayule/Hevea Substitution

Technical Perspective	Organizational Perspective	Individual Perspective
Tests have shown guayule to be a satisfactory substitute for hevea in automobile and aircraft tires (the primary use of natural rubber).	Mexico has had a long history of interest in guayule (wild natural growth, a research institute in Saltillo, a pilot processing plant) but relations with U.S. can generate strain.	Ed Flynn ("Mr. Guayule Rubber News ") is a determined promoter.
In view of the 300% price increase of natural rubber since 1972, guayule is becoming competitive.	Research is not the key issue; rather, production start-up raises the question of assumption of financial risk between the tire and rubber companies and the government (Federal and California).	Effective leadership of, and cooperation between, the Federal Government (Alex Mercure, Chairman of Guayule Commission) and the California Department of Food and Agriculture (Isi Siddiqui) may spark implementation action.
Guayule development can meet 100% of the projected natural rubber shortfall by 1991.	The Department of Agriculture has not been aggressive; the Guayule Commission may become the lead group but has not done much to date; national security considerations may be decisive (Asian turbulence: Iran, Afghanistan, others?).	Rep. George Brown (Dem., California) has been a most effective advocate in the Congress and has been joined more recently by Sen. Domenici (Rep., N.M.). Texas is lacking a strong Congressional supporter.
Yield per acre is expected to double between 1985 and 2000.	Inbreeding appears to be a problem in tire and rubber industry management.	

REFERENCES

Ascher, W. 1978. Forecasting; An Appraisal for Policy--Makers and Planners, The Johns Hopkins University Press, Baltimore.

Ayres, R. U. 1968. Envelope Curve Forecasting, in J. Bright Editor, Technological Forecasting for Industry and Government, Prentice Hall.

Ayres, R. U. 1969. Technological Forecasting and Long Range Planning, McGraw-Hill Book Company.

Ayres, R. U. and A. Shapanka. 1976. Explicit Technological Substitution Forest in Long Range Input-Output Models, H. A. Linstone and D. Sahal, editor, Technological Substitution, Elsevier-North Holland, New York.

Blackman, A. E. 1972. A Mathematical Model for Trend Forecasts, Tech. Forecasting and Social Change, 3:441-452.

Blackman, A. W., E. J. Seligman, and G. C. Sogliero. 1976. An Innovation Index Based on Factor Analysis, Tech. Forecasting and Social Change, 4, pp. 301-316, 1973. Also, H. A. Linstone and D. Sahal, Editors. Technological Substitution, Elsevier-North Holland, New York.

Fisher, J. C., and R. H. Pry. 1971. A Simple Substitution Model of Technological Change. Tech. Forecasting and Social Change, 3:75-88.

Floyd, A. 1968. Trend Forecasting; A Methodology for a Figure of Merit, J. Bright Editor. Technological Forecasting for Industry and Government, Prentice-Hall.

Griliches, Z. 1976. Hybrid Corn; An Exploration in the Economics of Technical Change, Econometrica, 25:501-522.

Kear, B. H., and E. R. Thompson. 1980. Aircraft Gas Turbine Materials and Processes, in Science, Vol. 208:849.

Linstone, H. A. and M. Turoff, Editors. 1975. The Delphi Method: Techniques and Applications, Addison-Wesley Publishing Company, Reading, Massachusetts.

Linstone H. A., et. al. 1980. The Use of Alternative Decision Models for Technology Assessment, Progress Report No. 1 to NSF, Futures Research Institute, Portland State University.

Mansfield, E. 1961. Technical Change and the Rate of Imitation, Econometrica, 29:741-766.

Marchetti, C. 1977. Primary Energy Substitution Models: On the Interaction between Energy and Society, Tech. Forecasting and Social Change, 10:345-356.

Mensch, G. 1979. Stalemate in Technology: Innovations Overcome the Depression, Ballinger Publishing Company, Cambridge.

Sharif, M. N., and C. Kabir. 1976. A Generalized Model for Forecasting Technological Substitution, Tech. Forecasting and Social Change, 8:353-364. Also, H. A. Linstone and D. Sahal, Editors, Technological Substitution, Elsevier-North Holland, New York, 1976.

Sharif, M. N., and C. Kabir. 1976. System Dynamics Modeling for Forecasting Multilevel Technological Substitution, Tech. Forecasting and Social Change, 9:89-112. Also, H. A. Linstone and D. Sahal, Editors, Technological Substitution, Elsevier-North Holland, New York, 1976.

Stern, M. D., R. U. Ayres, and A. Shapanka. 1975. A Model for Forecasting the Substitution of the One Technology for Another, Tech. Forecasting and Social Change, 7:57-79. Also H. A. Linstone and D. Sahal, Technological Substitution, Elsevier-North Holland, New York, 1976.

Van Duijn, J. J., 1977. The Long Wave in Economic Life, De Economist, 125:544-576.

BIBLIOGRAPHIC DATA SHEET	1. Report No. NMAB-385	2.	3. Recipient's Accession No.
4. Title and Subtitle Analytical Techniques for Studying Substitution Among Materials			5. Report Date September 1982 6.
7. Author(s) Committee on Materials Substitution Methodology			8. Performing Organization Rept. No.
9. Performing Organization Name and Address National Materials Advisory Board National Academy of Sciences 2101 Constitution Avenue, N.W. Washington, D.C. 20418			10. Project/Task/Work Unit No. 11. Contract/Grant No. J0199138
12. Sponsoring Organization Name and Address Bureau of Mines U.S. Department of Interior Washington, D.C. 20241			13. Type of Report & Period Covered Final 14.

15. Supplementary Notes

16. Abstracts

The available analytical modeling techniques are assessed for their utility in treating and interpreting the mechanisms of materials substitution. The techniques are classified under eight categories, viz., extrapolation, case studies, optimization, econometrics, input/output, simulation, judgmental studies, and engineering studies. Invited papers on the techniques were presented and discussed at a workshop; these are presented as Part 2 of this report. The workshop and committee examined the strengths and weaknesses of the techniques from three points of view; requirements, potential uses, and criteria for selection, presented in tabular form herein. No specific recommendations are made. However, the committee concluded that no single technique can provide a comprehensive means of forecasting and analyzing substitution among materials; a good grasp of the strengths and weaknesses of the available techniques is necessary to understand the effects of various factors on materials substitution. It also concluded that the analyst is more important than the technique. The report outlines the potential of the available techniques for expanding our ability to treat and manage the substitution of one material for another.

17. Key Words and Document Analysis.

Materials Substitution
Methodology
Extrapolation
Case Studies
Optimization
Econometrics
Input/Output
Simulation
Judgmental Studies
Engineering Studies
Materials Selection
Materials Criteria
Potential Uses
Requirements
Techniques
Forecasting and Analyzing Substitution

17e. COSATI Field/Group

18. Availability Statement	19. Security Class (This Report)	21. No. of Pages
Release Unlimited	UNCLASSIFIED	183
	20. Security Class (This Page) UNCLASSIFIED	22. Price

FORM NTIS-35 (10-70) USCOMM-DC 40329-P