ENGINEERING-ECONOMIC MODELING: ENERGY SYSTEMS

Part I

(Special Issue from ENERGY - THE INTERNATIONAL JOURNAL)

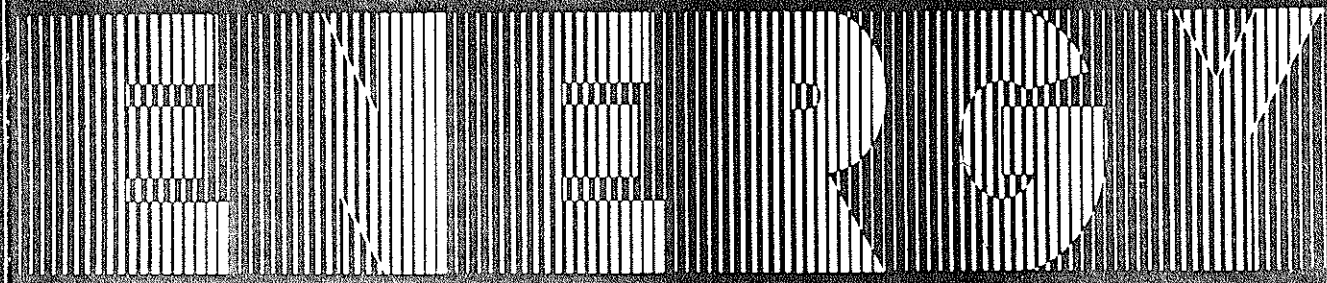
Guest Editors: John P. Weyant and Thomas A. Kuczmowski

EMF SR4

1990

VOLUME 15 NUMBER 3/4 MARCH/APRIL 1990 ISSN 0360-5442

TECHNOLOGIES · RESOURCES · RESERVES · DEMAND



IMPACT · CONSERVATION · MANAGEMENT · POLICY

The International Journal

SPECIAL ISSUE

ENGINEERING–ECONOMIC MODELING: ENERGY SYSTEMS

Part I

Guest Editors:

John P. Weyant and Thomas A. Kuczmowski

PERGAMON PRESS

Oxford · New York · Beijing · Frankfurt
São Paulo · Sydney · Tokyo · Toronto

ENERGY—THE INTERNATIONAL JOURNAL

(Technologies, Resources, Reserves, Demand, Impact, Conservation, Management, Policy)

Editor Professor S. S. Penner, *Energy Center, University of California (San Diego), La Jolla, CA 92093, U.S.A.*

Associate Editors

Term ending 31 December 1990

Professor Dr-Ing G. Adomeit, *Lehrstuhl für Allgemeine Mechanik, Rhein.-Westf. Technische Hochschule Aachen, 5100 Aachen, F.R.G.*

Professor Corrado Casci, *Centro Nazionale di Ricerca, Politechnico di Milano, Milano, Italy*

Professor Jaime Cervantes de Gortari, *Faculty of Engineering, National University of Mexico, Mexico, DF Mexico 04510*

Dr Larry Icerman, *2999 Calle Cerrada, Santa Fe, NM 87505, U.S.A.*

Dr Frank Kreith, *1485 Sierra Drive, Boulder, CO 80302, U.S.A.*

Professor Abrahim Lavi, *Carnegie-Mellon University, Schenley Park, Pittsburgh, PA 15213, U.S.A.*

Dr T. A. Reddy, *Center for Energy and Environmental Studies, The Engineering Quadrangle, P.O. Box GN 5263, Princeton University, Princeton, NJ 08544-5263, U.S.A.*

Professor S. Staniszewski, *Director, Warsaw Technical University, Institute of Heat Engineering, ul. Nowowiejska 25, 00-665 Warszawa, Poland*

Professor Georgios Tsatsaronis, *Department of Mechanical Engineering, Tennessee Technological University, Box 5032, Cookeville, TN 38505, U.S.A.*

Software Section Editor

Professor Anthony V. Sebald, *Department of Applied Mechanics and Engineering Sciences, University of California (San Diego), La Jolla, CA 92093, U.S.A.*

Term ending 31 December 1991

Professor A. J. Appleby, *Director, Center for Electrochemical Systems and Hydrogen Research, Texas A&M University, 238 Wisenbaker ERC, College Station, TX 77843-3577, U.S.A.*

Dr Garth S. Harris, *Arthur Young, P.O. Box 2146, Auckland, New Zealand*

Professor Manfred Heindler, *Institute for Theoretical Physics, Technical University of Graz, Peterstrasse 15, A-8010 Graz, Austria*

Professor Robert Herendeen, *Department of Forestry, 110 Mumford Hall, 1301 W. Gregory Drive, University of Illinois, Urbana, IL 61801, U.S.A.*

Professor Sidney Loeb, *Chemical Engineering Department, Ben-Gurion University of the Negev, Beer Sheva 84105, Israel*

Dr David B. Reister, *Engineering Physics and Mathematics Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831, U.S.A.*

Professor Arthur H. Rosenfeld, *Department of Physics and Energy Analysis Program, Lawrence Berkeley Laboratory, University of California, Berkeley, CA 94720, U.S.A.*

Dr T. A. Siddiqi, *Research Associate, East–West Environment and Policy Institute, East–West Center, 1777 East–West Road, Honolulu, HI 96848, U.S.A.*

Professor Kenneth E. Train, *Department of Economics, University of California, Berkeley, CA 94705, U.S.A.*

Dr Noel D. Uri, *U.S. Department of Agriculture, 1301 New York Ave., NW, Washington, DC 20005, U.S.A.*

Dr Göran Wall, *Physical Resource Theory, Chalmers University of Technology and University of Göteborg, S-41296 Göteborg, Sweden*

Dr T. Y. Yan, *Mobil Research & Development Corp., P.O. Box 1025, Princeton, NJ 08540, U.S.A.*

Term ending 31 December 1992

Professor B. W. Ang, *Department of Industrial and Systems Engineering, National University of Singapore, Kent Ridge, Singapore 0511*

Dr M. L. Barrère, *Directeur des Recherches, ONERA, 92320 Châtillon-sous-Bagneux, France*

Dr M. R. Brambley, *Batelle Pacific N.W. Laboratories, P.O. Box 999, 2400 Stevens, Richland, WA 99352, U.S.A.*

Professor A. B. Cambel, *Director of Energy Programs, George Washington University, Washington, DC 20052, U.S.A.*

Professor Bruce Hannon, *Energy Research Group, University of Illinois, Urbana, IL 61801, U.S.A.*

Dr Eric Hirst, *Energy Division, Oak Ridge National Laboratory, Oak Ridge, TN 37830, U.S.A.*

Professor R. Kümmel, *Physikalisches Institut der Universität Würzburg, D-8700 Würzburg, F.R.G.*

Professor R. Ramanathan, *Department of Economics, University of California, San Diego, La Jolla, CA 92093, U.S.A.*

Dr Kirk R. Smith, *Resource Systems Institute, East–West Center, 1777 East–West Road, Honolulu, HI 96848, U.S.A.*

Professor Robert H. Socolow, *Center for Energy & Environment Studies, School of Engineering/Applied Science, Princeton University, Princeton, NJ 08544, U.S.A.*

Professor Ho-Ming Yeh, *Department of Chemical Engineering, National Taiwan University, Taipei, Taiwan, Republic of China*

Dr Wang Yin-ren, *Beijing Research Institute of Coal Chemistry, CCMRI Beijing, People's Republic of China*

Publishing Office: Journals Production Unit, Pergamon Press plc, Hennock Road, Marsh Barton, Exeter EX2 8NE, Devon, U.K. (*Tel.* Exeter (0392) 51558; *Telex* 42749).

Subscription and Advertising Offices: *North America*—Pergamon Press Inc., Maxwell House, Fairview Park, Elmsford, NY 10523, U.S.A. *Rest of the World*—Pergamon Press plc, Headington Hill Hall, Oxford OX3 0BW, U.K. (*Tel.* Oxford (0865) 64881).

Published Monthly

Subscription Rates (1990): Annual institutional subscription rate (1990) DM 1025.00; 2-year institutional rate (1990/91) DM 1947.50: personal subscription rate for those whose library subscribes at the regular rate (1990) DM 189.00. Prices are subject to amendment without notice.

Back Issues: Back issues of all previously published volumes, in both hard copy and in microform are available direct from Pergamon Press offices.

Copyright © 1990 Pergamon Press plc

It is a condition of publication that manuscripts submitted to this journal have not been published and will not be simultaneously submitted or published elsewhere. By submitting a manuscript, the authors agree that the copyright for their article is transferred to the publisher if and when the article is accepted for publication. However, assignment of copyright is not required from authors who work for organizations which do not permit such assignment. The copyright covers the exclusive rights to reproduce and distribute the article, including reprints, photographic reproductions, microform or any other reproductions of similar nature and translations. No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means, electronic, electrostatic, magnetic tape, mechanical, photocopying, recording or otherwise, without permission in writing from the copyright holder.

Photocopying information for users in the U.S.A. The Item-fee Code for this publication indicates that authorization to photocopy items for internal or personal use is granted by the copyright holder for libraries and other users registered with the Copyright Clearance Center (CCC) Transactional Reporting Service provided the stated fee for copying beyond that permitted by Section 107 or 108 of the United States Copyright Law is paid. The appropriate remittance of $3.00 per copy per article is paid directly to the Copyright Clearance Center Inc., 27 Congress Street, Salem, MA 01970, U.S.A.

Permission for other use. The copyright owner's consent does not extend to copying for general distribution, for promotion, for creating new works, or for resale. Specific written permission must be obtained from the publisher for such copying.

The Item-fee Code for this publication is: 0360-5442/90 $3.00 + 0.00

⊗™ The paper used in this publication meets the minimum requirements of American National Standard for Information Sciences—Permanence of Paper for Printed Library Materials, ANSI Z39.48-1984.

ENGINEERING–ECONOMIC MODELING: ENERGY SYSTEMS

Part I

Guest Editors:

John P. Weyant and **Thomas A. Kuczmowski**
Energy Modeling Forum, Stanford University, California

PERGAMON PRESS
OXFORD · NEW YORK · BEIJING · FRANKFURT
SÃO PAULO · SYDNEY · TOKYO · TORONTO

Copyright © 1990 Pergamon Press plc

All Rights Reserved. No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means: electronic, electrostatic, magnetic tape, mechanical, photocopying, recording or otherwise, without permission in writing from the publisher.

Published as Volume 15, Number 3/4 of the journal *Energy* and supplied to subscribers as part of their 1990 subscription. Also available to non-subscribers.

Printed in Great Britain by BPCC Wheatons Ltd, Exeter

CONTENTS

ENGINEERING–ECONOMIC MODELING: ENERGY SYSTEMS

Part I

Preface v

SECTION 1: MODELING FOR POLICY DEVELOPMENT

Ryo Imoto	145	Government's point of view on modeling
Arnold B. Baker and **Anthony J. Finizza**	149	Corporate point of view on modeling
Walter C. Labys	155	Evaluation of models
Yoichi Kaya	163	Credibility of models
Paul J. Werbos	171	Rational approaches to identifying policy objectives
John P. Weyant	187	Application of models in the process of legislation
John P. Weyant	203	Policy modeling: an overview

SECTION 2: DEMAND FORECASTING METHODOLOGIES

Andrew J. Lipinski	207	Introduction
Paul J. Werbos	213	Econometric techniques: theory versus practice
Walter C. Labys and **Hiroshi Asano**	237	Process models
David B. Reister	249	The hybrid approach to demand modeling
Bertrand Chateau and **Bruno Lapillonne**	261	Accounting and end-use models
Andrew J. Lipinski	279	Probabilistic forecasting
Ahmad Faruqui, **Thomas Kuczmowski**, and **Peter Lilienthal**	285	Demand forecasting methodologies: an overview for electric utilities

INDEXED IN Curr. Cont. ASCA, Biosis Data., Cam. Sci. Abstr., Chem. Abstr. Serv., Curr. Cont. Eng. Tech. & Applied Sci., CABS, Eng. Ind., Environ. Per. Bibl., INSPEC Data., Curr. Cont. Sci. Cit. Ind., Curr. Cont. SCISEARCH Data.

ISSN 0360-5442
ENEYDS 15(3/4) 145–394 (1990)

SECTION 3: MODELING ENERGY–ECONOMY INTERACTIONS

William W. Hogan and **Shailendra C. Parikh**	297	Modeling interactions
P. Capros, P. Karadeloglou, G. Mentzas, and **J.-E. Samouilidis**	301	Short and medium-term modeling and problems of models linkage
Xia Shi and **Yingzhong Lu**	325	Multisectoral planning models
Walter C. Labys and **John P. Weyant**	341	Integrated models
Gale Boyd, John Fox, and **Donald Hanson**	345	Set of models
Nazli Choucri and **Christopher Heye**	363	Simulation models
Andrew J. Lipinski	379	Cross-impact models
Janusz Cofala, Thomas Kuczmowski, and **John Weyant**	387	Energy–economy modeling: a survey
Software Survey Section	I	

PERGAMON PRESS
Oxford · New York · Beijing · Frankfurt · São Paulo · Sydney · Tokyo · Toronto

Energy Vol. 15, No. 3/4, p. v, 1990
Pergamon Press plc. Printed in Great Britain

PREFACE

This special issue surveys different modeling techniques for energy forecasting, planning, and decisionmaking. It is one of two related special issues of Energy on this subject. The survey is not problem - but methodology - oriented, i.e., it does not explicitly address problems of energy planning or policy, but identifies methods that have proven useful in addressing these issues. Some fundamental problems encountered in applying quantitative methods to policy analysis, including defining objectives and insuring model credibility, are addressed in the first part of this volume. This is followed by a review of forecasting methods, including econometric, process engineering, and hybrid approaches. The final part of this volume deals with the theory and practice of modeling energy-economy interactions, with the linkage of energy sector and economy-wide models at various levels being the main tool for most practical applications.

Besides our obvious debt to the authors of individual papers in this volume, who also served as reviewers of contributions other than their own, we thank Stanford Penner for inviting us to contribute this special issue, and Andrew Lipinski, recently retired from the Institute for the Future, for organizing the contributions of forecasting methods, and his advice on many other aspects of the volume. We also thank Edith Leni of the Energy Modeling Forum at Stanford University for her tireless work in editing and producing the final copy for this volume. Finally, without the continual efforts and timely advice of Peter Shepherd and John Bailey of Pergamon Press this project would never have come to fruition.

John P. Weyant
Thomas A. Kuczmowski
Palo Alto, California
December 1989

SECTION 1

MODELING FOR POLICY DEVELOPMENT

Energy Vol. 15, No. 3/4, pp. 145–147, 1990
Printed in Great Britain

0360-5442/90 $3.00 + 0.00
Pergamon Press plc

1.1. GOVERNMENT'S POINT OF VIEW ON MODELING

RYO IMOTO

Nuclear Power Division
Ministry for International Trade and Industry
JAPAN

Abstract - Government generally needs many kinds of forecasts and outlooks as fundamental input data in making political decisions. Engineering-economic models are a useful and powerful method for obtaining forecasts of population, economic growth, fuel prices, assessments of limitations on resources, etc., which are fundamental data for policymaking. In recent years, methodologies of synthesis evaluation applied models including and based on preference of political decisions of policymakers have also been studied. This paper shows how economic models can be applied for middle- to long-term political decisionmaking using methodologies of energy technology evaluation as an example, which is a typical political subject, and it refers to a problem awaiting a solution from the government's point of view.

1. THE OBJECTIVE OF TECHNOLOGY EVALUATION

The objective of government energy technology evaluation is to identify the most suitable political decision that promises to satisfy multiple political objectives such as reducing the ratio of oil imports, minimizing costs to energy consumers, ensuring sufficient energy supplies for maintaining industrial prosperity, maximizing safety and reliability, and avoiding environmental pollution, etc.

In general, there are three different types of relationship between technology and society. The first type of relationship is where technology induces social change (i.e., changing private life styles and/or markets, promoting new industries). The second type of relationship is when the needs of society and values of the times inspire the development of new technologies. The third type of relationship is an intermediate case, where cause and effect between technology development and changes in lifestyles move in parallel.

In an era where there are many breakthrough scientific discoveries and rapid technological progress, society tends to be changed by technology more than vice versa. But nowadays many technologies are almost mature, and developments of technologies may entail large risks, so societal needs and values are tending to drive technology development. As far as energy technology is concerned, it is clear that society and/or major coalitions of individuals in it have strong power to choose desirable technologies.

So it is essential to analyze and synthesize many items regarding a new technology's development such as R&D costs, features of technologies, commercializing time, amount of energy supplied, difficulty of development, environmental load, etc. to evaluate it while considering uncertainty regarding future trends in society and the economy.

2. METHODOLOGIES OF ENERGY TECHNOLOGY EVALUATION

In energy technology evaluation, it is important to specify a number of fundamental inputs and conditions beforehand such as the technologies to be evaluated, the evaluation area (all over the world, a whole country or a district), the amount of risk associated with each energy supply technology, time to introduction (long or short) and the evaluation criteria. Different inputs and conditions can lead to different results and may even require different evaluation methods. In general, the more detailed the analysis, the more valuable the outputs. But simple analysis can also be useful in cross-checking the results from--and supplement--the outputs of the more detailed analysis. The figure below shows a sample of the framework of synthesis methodologies of energy technology evaluation.

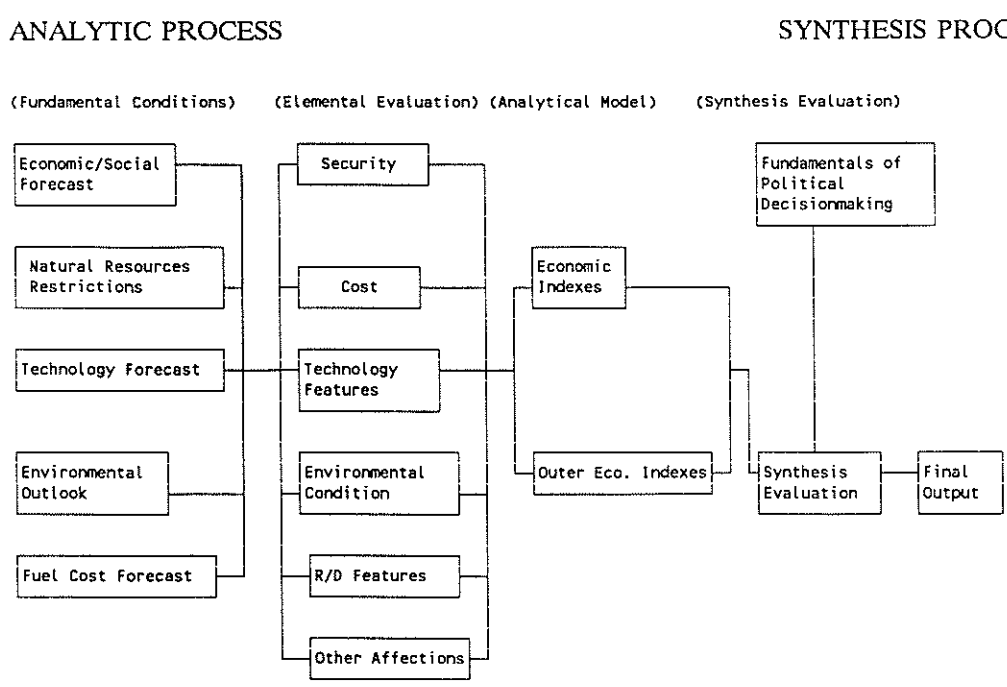


Figure 1. Framework for Synthesis Methodologies of Energy Technology Evaluation

Analytical Process

In this process some fundamental conditions, i.e., economic/social forecasts, restrictions on natural resources use, etc. are developed. To make the analysis concrete some evaluating factors of new technologies need to be drawn out by means of statistics investigations and scenario analysis. Then using these fundamental conditions, elemental evaluations can be made. Elemental evaluations are evaluations of each new technology with respect to various factors, i.e., security implications, costs, technical features, environmental features, R&D features and other criteria.

Energy efficiency and reliability of the technology are the most important technical features to be evaluated. R&D planning is a problem of resources distribution, so it requires the best allocation of research workers, research facilities, time, money, energy, and intelligence for the purpose of maximizing R&D output results. The quantitative economic data derived from these elemental evaluations will be put into economic models and the expected benefits and costs evaluated. In parallel, external factors (i.e., security, probability of development success, environmental criteria, etc.) are made into quantitative indexes.

Synthesis Process

Finally, when a synthesis evaluation is attempted, final policy recommendations are derived by considering both economic and non-economic indexes within the political value system of decisionmakers. In these uncertain times, responsible government decisionmakers (probably in a way similar to decisionmaking in private companies) should choose the best policy from the available alternatives. Above all, political decisionmaking by the government should influence society widely and deeply. So the most probable policy is one that provides political fulfillment.

Difficulty of Energy Technology Evaluation

There are two reasons that make it difficult to evaluate energy technologies. One is technical and the other is social. Energy is, of course, a fundamental material of life and the economy. Energy technologies must continuously supply energy at as low a cost as possible. So they must be large-scale, with high efficiency and be operated for a long time with high reliability. Furthermore, they should have excellent safety and environmental features to protect people who live near energy conversion facilities. Therefore, energy technologies need to fulfill contradictory requirements--low cost, but not so low cost as to be unsafe or overly polluting.

Besides these technical concerns, the latest trends in the energy situation (i.e., reductions in the demand for energy, low prices for oil, interfuel competition, increases in safety and environmental demands) require new directions in the development of future energy technologies. It will be necessary to continually examine and analyze various requirements requested by people who demand energy and by people living near energy facilities. So the government needs to make political decisions in energy technology R&D that cover all of these considerations.

Necessity of Synthesis Evaluation by Economic Models and its Problem

At present, economic models help a great deal in political decisionmaking. Especially in analytic processes they have important roles and are largely reliable. And it is evident that political energy decisionmaking is supported by synthesis evaluations. But there are some problems doing synthesis evaluation/decisions using economic models. Such models should include many factors and be capable of making tradeoffs between them. In many cases, there are non-economic indexes (i.e., safety, security, environmental conditions, etc.), which are generally difficult to quantify, besides economic indexes. Non-economic indexes seem to be becoming more important as social requirements become more diversified. Synthesis evaluations should include economic and non-economic considerations.

So, in general, it is very difficult to develop final synthesis evaluations with sufficient reliability strictly from economic models. But the study of economic models is rapidly advancing. In fact, in recent years economic models for decisionmaking, handling multi-dimensional factors including non-economic indexes, have been developed and refined. Government officials who take part in political decisionmaking need economic models effective for synthesis evaluations as soon as possible because decisionmakers now desire more objective and reliable policies in this age of uncertainty.

REFERENCES

EPRI (1982). Electric Generation Expansion Analysis System, EPRI-2561(1), Palo Alto.

The Institute of Applied Energy (1981). "Development of Methodologies for Energy System Evaluation," The Research Report, Tokyo.

Uchiyama, Y., (1989). "General Remarks--Methodologies of Energy Technology Evaluation," Journal of Energy and Resources, Vol. 10 No. 3, Tokyo.

Energy Vol. 15, No. 3/4, pp. 149–153, 1990
Printed in Great Britain

0360-5442/90 $3.00 + 0.00
Pergamon Press plc

1.2. CORPORATE POINT OF VIEW ON MODELING

ARNOLD B. BAKER and ANTHONY J. FINIZZA

Atlantic Richfield Company (ARCO)
Los Angeles, California
U.S.A.

Abstract - This paper describes the use and limitations of economic models in the corporate environment. It reviews alternative model types and their applications for business environment analysis, investment alternatives, and strategic decisions. In addition, it suggests potential areas for model improvement that would benefit the corporate decision making process.

Economic models can help to quantify thought processes and extend the ability to think through complex, multivariable problems. In the corporate environment, they can also help to standardize and simplify decision making processes, and help to understand the uncertainties in the business environment in which we operate. In the simplest terms, models are devices to uncover knowledge and help gauge uncertainty.

In many instances, however, mathematical models often take on a life of their own, yielding answers that are received as "the true word." This we believe to be counter-productive, if not highly misleading. Models are only tools that mirror and extend thought processes. They must be used cautiously, with their limitations clearly recognized. They are not a substitute for thinking, but only a stimulus and aid in the process. In fact, many of the most effective models used are combinations of judgmental and quantitative models, which more closely replicate the "mental" model of the corporation.

Corporations do have a wide variety of economic modeling needs. While specific needs vary with particular lines of business and management styles, these needs can roughly be broken down into three areas: Evaluation of specific investment decisions; analyzing and modeling the future business and market environment; and delineating fundamental business strategies in alternative environments.

Each of these needs involves somewhat different types of models. We will briefly review each of these corporate areas,indicating the types of models and purposes for which they are used, and then provide our view of the role of models and their limitations in the corporate environment.

Micro-models for Investment Decisions

These are the most economically and technically pure models, and are usually developed or implemented, at least initially,by those closest to the project development. The objective is to clearly lay out the financial aspects of a business investment for a variety of operational and market assumptions.This includes testing whether or not the project will meet the minimum corporate financial criteria and how robust the project will be to alternative environments. The time profile involved is the life of the project, which may range from ten to thirty years or

longer. These models may include process models, optimization models, accounting models, engineering models, and in the case of oil and gas, competitive bidding models. Different models are used for different classes of projects, but in general, the corporation will adopt a specific corporate-wide model (or at least, a financial methodology) for a specific class of projects to assure analytical consistency.

These models focus on the detailed micro level of the investment--e.g. where to locate; how will the plant operate; where to drill etc. In effect, these are financial "bench scale" working models of the project in question. To deal with financial and market uncertainty, alternative assumptions are made for input and product prices and sales quantities and for the cost of capital.

While, as noted above, this analysis is generally carried out by those technically closest to the project, the market and capital cost assumptions generally are provided by a central corporate group to assure corporate-wide analytical consistency.

The output of these models is then reviewed and compared with other projects or with corporate standards, and depending on the specific capital cost involved and corporate decision authority, are either approved at the local level or moved up the corporate hierarchy for further review and approval.

Macro-models of the Business and Market Environment

These are an entirely different set of models. These tend to be more econometric and are generally utilized by staff economists and market analysts, with the output reviewed by senior management. The objective here is to analyze and provide clear signals on alternative business environments and markets that the corporation is likely to face. In some instances forecasts are needed by month or quarter for up to five years; in others, annual forecasts are needed for 20 to 40 years into the future.

While project evaluation models tend to be developed in house, models in this group tend to be purchased from outside vendors, and include elaborate macroeconomic models, integrated world energy models, and sectoral U.S. and fuel market models. They are usually based on historic price and income elasticities and/or input-output relationships. These types of models were initially developed and began to be widely utilized in the late 1960's and early 1970's when world economic and energy markets appeared to be in somewhat steady state growth. They are still popular today and largely form the basis of economic forecasting. Their strength lies in their adding up properties of a wide number of economic and policy variables. Generally the economic relationships are pre-wired, with the casual user able to select policy levers (i.e. tax rates, monetary policy, etc), and the more well trained user able to re-estimate individual equations and modify them as desired.

Energy versions of these models and even energy versions that are directly linked to these macroeconomic models are also available and widely used. In general, they are not quite as effective as the macro models themselves, since energy data is typically poorer quality than economic data, there are much longer lags in reporting energy data, and the linkages between energy and economic variables is somewhat less clear than between economic variables themselves.

As noted, some of these models may be sector specific, such as models of U.S. oil production, natural gas production, on road gasoline demand, electric utility fuel demand, or residential fuel demand. They may be used directly as forecasting tools, or simply as tools to answer "what if" kinds of questions and market sensitivities. For example, because U.S. oil production is a key variable in world oil markets and oil prices, we might wish to know how U.S. oil production would respond to a range of different oil prices, and whether sharply lower

oil prices would likely lead to sharply reduced oil production over a short time period. By understanding that relationship, we would be in a better position to forecast long term oil prices.

A common feature of the industry-specific Macro-models is their use as a scenario generator rather than a base case generator. Often times the model is benchmarked to a set of industry conditions that were produced analytically outside the model context. Then, a set of price and income elasticities and other parameters are calculated to closely replicate what the analyst has determined to be the base case. The model is then used to calculate the effects of alternative policies and other factors on the base case. These models are quite helpful in determining the sensitivity of factors to the base case. It is important for the sensitivities not to stray too greatly from the base case or the results are suspect.

As a practical matter, political "models" can come into play here as well. These may be formal probabilistic models of country political risk, or they may be more judgmental and broader in nature. The need here is to capture a sense of political forces that tie into and influence primary driving economic variables. Pure country political risk models may come more into play as a part of the process of project evaluation (i.e., trying to decide whether to build a plant in a particular country and needing to quantify the risk of expropriation of assets, the likelihood of changed fiscal regimes, etc).

These models help to generate the key external business environment planning and project evaluation variables. Major planning variables (which vary by corporate line of business), such as GNP growth, inflation, interest rates, and oil prices are frequently reviewed by corporate senior management who makes the final decision on what their level and trend will be.

The Nexus: Business Strategy Analysis

There are two somewhat different types of models here. Both are utilized by divisional or corporate planning personnel, with the output generally reviewed by senior management.

The first type is a general planning model that mimics the corporate division or corporation as a whole. Frequently, each division will have its own divisional model, with corporate headquarters having a model which aggregates the output of the divisional models. These models tend to be accounting type of models and provide financial summary data in alternative business operating environments. These are frequently used to generate annual long range plans of two to five years or more.

Many of these are available from third party vendors, and to varying degrees can be tailored to the structure of the particular corporation using it.

The output of these models is used by senior management to test aspects of long range plans, and help to support fundamental strategic decisions. For example, what will we look like if we buy or sell that division, if oil prices are halved or doubled, if we reduce manpower costs by 10 percent, etc.

Because these models are accounting type by nature, they tend not to include interdivisional and second round financial interaction of such external effects. Rising oil prices may, for example, may depress margins for refined product prices and raise the feedstock costs of liquid based petrochemicals. These factors would not usually be considered internally in the models, but would need to be considered externally to arrive at the integrated corporate impact.

A somewhat different type of model is a more direct strategic planning model. It is a mixture of somewhat more limited accounting functions with some interactive feedback effects, and may contain additional financial measures as the present value of a corporation's assets. These models try to model a corporation in a functionally reduced form, and are specifically

used to test alternative strategies. For example, what happens to the present value of a company if it bets on one economic growth scenario and a totally different one results? Are there elements of strategies that will work effectively in both cases? Which elements offer the greatest risk and reward?

These models are more generally designed within the corporation. In many instances, formal mathematical models are not used, and the mental models of senior executives serve as the basis for key strategic decisions.

Model Limitations for Corporate Needs

Perhaps the most applicable group of models for corporate needs are those used in project evaluation and investment decisions. They are micro oriented, involve very specific assumptions about the project, and because they tend to be engineering or process oriented, fulfill their function well - examining the costs of specific projects under alternative market, price, and cost assumptions.

Long range planning and strategic planning models probably are next most useful as tools, as long as one recognizes their limitations. An accounting model will only add up what you put into it, and corporate interdivisional interactions are very important in understanding the bottom line output. Strategic planning models tend to be theoretical, and care must be taken to translate their output into real world dollars and cents.

Business and market environment mathematical models tend to be the least useful in the current business environment. The macro aspects of these models tend to miss key trends and discontinuities such as oil price shocks and technological innovation. They too often tend to be black boxes in which key relationships get buried between input and output.

Also, because these models tend to be so large, personnel need to be dedicated to understand and run them, with high start up costs, and high costs of changing over to competitive models. And perhaps most importantly, because of these models size and complexity, people tend to believe the model output without thinking about it. Many corporate model users place a high premium on small, understandable macro models where these problems can be avoided.

In the current business environment, there is no single macro model we can rely on, or for that matter, would want to. The key is to understand the major economic relationships in a company's lines of business; understand the basic market sensitivities and how they impact profitability; and understand the fundamental trends of these variables over time.

This approach is eclectic. It forces hard thinking about what is happening in the business environment, and thinking at the micro level. It is labor intensive and time consuming, but once the key factors on which a particular company needs to focus have been identified, periodically reviewing them is less labor intensive. In addition, this approach forces users to be close to the underlying data, and provides them a feel for major trends. It allows users to borrow from the latest market research and latest third party model output, and integrate that with their own analysis. And it leaves users with an understanding of forces and a few key numbers, rather than a lot of numbers that cannot be explained.

Ultimately in a large corporation, model results do not matter. What matters is an understanding of the problem and the best alternative solutions. These usually need to be quickly determined, and yet must have analytical credibility. And they must be presented to decision-makers in a way that is focused and easily understood through the decision-maker's own mental models of the world, frequently gained through years of hands on experience.

In sum, given our uncertain business environment, the simpler the model the better. The more it is used for testing hypotheses and understanding relationships rather than pure forecasting, the more meaningful the results are likely to be.

This we believe to be true given the capability of today's models. The next generation of models using artificial intelligence and enhanced personal computers, may offer something much better still. By modeling the decision/thought process, they may be able to judgementally integrate all of the disparate pieces of information users must now do themselves, and it may make the learning process considerably easier by quantitatively documenting it.

Energy Vol. 15, No. 3/4, pp. 155–162, 1990
Printed in Great Britain

0360-5442/90 $3.00 + 0.00
Pergamon Press plc

1.3. EVALUATION OF MODELS+

WALTER C. LABYS

Department of Resource Economics, West Virginia University
Morgantown, West Virginia
U.S.A.

Abstract - This section addresses the important issue of how business managers and government planners can gain confidence in using engineering-economic models as part of the decision making and planning process. Purposes and methods of model evaluation are discussed from the viewpoint of the modeling methodologies employed and the nature of the model validation procedures applied to them.

1. INTRODUCTION

It has been shown thus far that business managers and government planners are employing engineering-economic models to a greater extent to overcome economic uncertainties. In fact there has been a proliferation of these kinds of models, some of which have been academic and experimental, but many of which have been very large scale and have demanded substantial efforts on the part of model planners and model builders to apply them. Thus, it has proved to be more than a minor issue to determine not only the inherent validity of these models but also their creditability, or the trust that can be placed in them. Model evaluation has thus become an essential step in improving the effectiveness of engineering-economic planning.

The need for evaluation and better credibility is indirectly illustrated by the recent study of Greenberger and his associates (1983) who have surveyed the views of energy experts and others regarding the quality, the attention received, and the influence of fourteen major energy studies of the past decade. The survey indicates that while attention received and influence were positively correlated, quality tended to be negatively correlated with both attention received and influence. The study offers no explanation for this disconcerting negative correlation, but throughout emphasizes the importance of timeliness and effective communication of modeling results.

The elements of model evaluation have been much discussed and debated. Two summary studies include Gass (1981), and Labys and Wood (1985) who attempt not only to abstract from the modeling process to develop validation as part of that process but also to explore the implications of the process. Validation here refers to the correspondence of the model to the underlying processes being modeled. This form of validation includes three elements: (1) the

+This chapter represents an update of two earlier papers, W.C. Labys, "Measuring the Validity and Performance of Energy Models," Energy Economics 4 (1982) pp. 159-168; and W.C. Labys and David O. Wood, "Energy Modeling," in W.A. Vogely (ed.). Economics of the Mineral Industries, American Institute of Mining Engineers, New York, 1985. pp. 303-336.

structural features of a model, (2) the inclusion of relevant variables such as policy instruments and concepts of importance to be analyzed, and (3) the model's capability.

Structural evaluation, of course, is the essence of scientific analysis. It is based upon (1) the conceptual specifications of the model, (2) the specification and application of the measurement process by which the model data are generated or obtained, (3) the specification and analysis of the scientific hypotheses derived from theory underlying the model and to be tested via analysis of the model data, and (4) the selection of the final model best supported by the scientific laws, principles, maintained hypotheses, and tested hypotheses which have emerged from the research process. Validation of the model structure starts with the replication of measurements and hypotheses testing, but also includes analysis and/or counter-analysis involving the variables and concepts integral to the policy issues for which the model was intended.

The second validation issue, content validity, is usually singled out from structural validity for the latter purpose. Both policy evaluation and policy analysis require that engineering-economic models reflect the appropriate policy concepts and instruments. A policy evaluation model will be simpler than a policy analysis model in this regard in that only the policy actually implemented and being evaluated must be included. Engineering-economic models which emphasize policy are more complicated in that the policy instruments and concepts suitable for the alternative policies of potential interest and importance to the various constituencies concerned with the issue(s) of interest must be included. Further the model must be explicit concerning the resolution of "facts" and/or value judgments which are in dispute among the various constituencies. Crissey (1975) has made the evaluation of such "facts" or contention points a central feature of his approach to policy model analysis.

The third element of model validation is predictive capability, determining if the scientific information and results included in the model are sufficient to discriminate among the policies being considered. If the range of scientific uncertainty spans the range of policy dispute, then the model's usefulness in policy research is very limited. Model-based studies may sometimes only report point predictions and not information on prediction confidence levels or sensitivity analysis of predictions to changes in input data and/or structural coefficients, consistent with known or conjectured uncertainties in the underlying measurement processes and scientific results. The lack of this additional information may suggest an unjustified precision of analysis. Analysis of predictive power is thus an important aspect of policy model analysis, quite independent of the structural validity of the components of the model.

Closely related to the various dimensions of model validity is the validity of the data associated with an engineering-economic model. Data validation must include not only evaluation of the measurement process by which the data component of model structure is developed, but also the processes by which the data required for model applications are obtained. While data and measurement process evaluation are closely related to model evaluation, particularly to that of model structural and predictive capability, it is probably useful to single out this aspect of validation, since it typically receives too little attention in policy modeling and research.

Crissey (1975) has a similar perspective on the elements of policy model validation. He emphasizes that the credibility and utility of a policy model will depend upon its treatment of the factual, behavioral, evaluative, and structural issues in dispute. Disputed issues should be represented in the model in a manner facilitating analysis of alternative resolutions. Such issues comprise the model's contention points. According to Crissey, a contention point is said to be critical if changes in its resolution significantly affect the model conclusions, and it is a contingency point if changing the resolution of this contention point in combination with others results in a significant change in the model result (Crissey, 1975; pp. 83-88). This concept of model contention points provides a useful focus for structural, content, and predictive validation.

In contrast to validation, model verification refers to the evaluation of the actual model implementation. At issue is the correspondence of the implemented model--usually a computer

program--to what the modeler intended. Verification is thus more mechanical and definitive than model and data validation. Gass (1979) has suggested that policy model verification is the responsibility of the modeler, and that evaluation should be limited to review of the verification process.

A final aspect of engineering-economic model evaluation concerns usability. This dimension of evaluation refers to both the sufficiency of documentation to support model understanding and applications, and the efficiency of the overall system. Technical documentation and materials sufficient to inform potential users of the nature of the model's structure, content, and predictive characteristics, as well as to support interpretation of model-based results, are essential for any policy model. The need for documentation to support independent application of the model, including user guides, system guides, and test problems will depend upon the model application environment. Of course, even if the intent is for the modeler to conduct all applications, there still should be evidence that application procedures have been developed, and that a reasonable applications practice is in effect.

The Validation Problem

Here a clarification must be made between the various procedures as well as the vocabulary which has been employed to describe the validation process. Model evaluation and assessment appears to describe the overall process of examining model performance. As an example, the MIT Model Assessment Group (1979) have identified four different approaches to evaluation:

(1) Review of the literature
(2) Overview assessment
(3) Independent audit
(4) In-depth assessment

The major distinction between the approaches concerns the materials used in evaluation. A summary of the relationships between evaluation approaches is given in Figure 1. A review of this literature for a model, or set of similar models, focuses upon model formulation, measurement and estimation issues relating to model structure, applicability for analysis of specific policy issues, and so on. Such a review may be both descriptive and evaluative. An example is the review by Taylor (1975) of electricity demand models comparing model structure with an "ideal" structure. In its various forms, literature review and analysis are the traditional means of model analysis. Issues of approach, logic, measurement and interpretation are formulated and analyzed. Issues of actual implementation are less susceptible to analysis with this approach.

An overview assessment uses the underlying technical model documentation, especially the computer code, for a more precise analysis of the model's structure and implementation. An overview evaluation can identify a policy model's critical points, but it will only occasionally be able to pass judgment on the adequacy of the model's treatment of them. The overview report is a useful intermediate stage in the assessment process, but assessment of the model's validity and applicability generally requires the acquisition and analysis of experimental data.

An independent audit evaluates a model's behavior by analyzing data derived from experiments that are designed by the assessors but run by the model builders. An important element of the procedure is that the assessment group is "looking over the model builder's shoulder" while the experimental runs are being made. This is essential to the accurate interpretation of the results produced by the experiment. An audit report should use the experimental data together with the analytical material developed in previous stages of the evaluation process to determine the model's validity in as many key areas (critical points) as possible. Audit procedures have the advantages of being relatively quick and inexpensive. With complex models, however, there will generally be some critical points that cannot be fully evaluated through an audit.

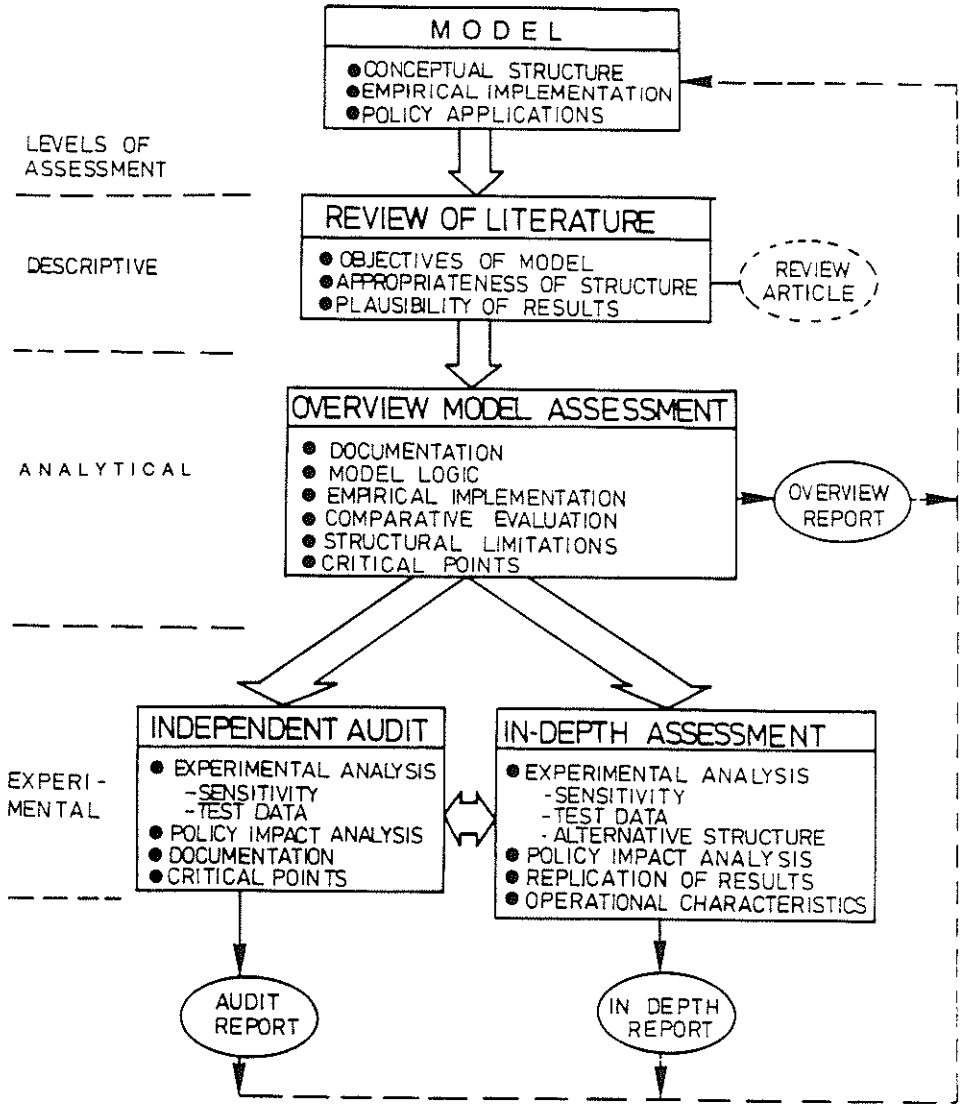


Figure 1. Approaches in Model Assessment

An in-depth assessment develops experimental data through direct, hands-on operation of the model. Direct operation makes it feasible to carry out more complex tests, particularly when the tests require modifications in model structure rather than simple changes in model parameters and/or data. Because of the significant costs of in-depth evaluation, it is probably most efficient to conduct exploratory analysis through an independent audit before embarking on more detailed evaluation. After an in-depth evaluation has been completed, audits might subsequently be used to update the evaluation as new versions of the model are developed. However, and to the contrary, an argument against validation has been made by House (1979) who decries the implication that economic models can foretell the future, let alone predict turning points accurately.

Even a model that is so designed that it 'apes' the past completely has nothing in its making that gives one any confidence that it will be able to forecast the future.

According to House, attempts to validate models on the basis of experience deriving from the natural and physical sciences are not likely to be successful. In the natural sciences, validation can more reasonably be provided by structuring the modelling methodology to ensure

that it will produce results that conform to the measured data. But in the social sciences, the discovered 'laws' are not natural nor exact. Thus, replication of the past is likely to say little about the future because future occurrences will deviate from the historical circumstances.

Concerning the availability of stringent validation tests, Greenberger et al. (1976) stress that there is no uniform procedure for model validation. The authors emphasize that all models are simplifications of the reference system or economic structure. They are never entirely valid in the sense of being fully supported by objective truth. The most appropriate expressions of a model's performance are that it be 'useful', 'illuminating', 'convincing', or 'inspire confidence' rather than it be 'valid'.

Confidence can be increased by having an engineering-economic model reproduce past behavior of the reference system, by exploring its sensitivity or response to perturbations, by critically examining the premises and theories on which it is based, and finally by putting it to use. However, this point of view should not be one of performing validation but rather of supplanting it. Such as approach to increasing confidence is an admission that some engineering modelling methodologies often do not rest on formally identified data or statistically significant coefficients.

The econometrician, in contrast, rigorously defines the conditions for validation utilizing statistical tests. Fromm (1973) emphasizes that models can only improve policy decision making if the premises on which the models are constructed are made explicit and the systems are subjected to a battery of validation tests.

All too frequently what purports to be a model based on theory and evidence is nothing more than a scholar's preconception and prejudices cloaked in the guise of scientific rigor and embellished with extensive computer simulations. Without empirical verification there is little assurance that the results of such exercises are reliable, or that they should be applied for normative policy purposes.

Apart from the econometrician's reasons for validation, we have neglected the role of error-cost analysis in the economic choice of models. That is, modelling methodologies vary in their costs as well as in their scope and accuracy. As stated by Chambers et al. (1971), the policy maker must fix the level of inaccuracy he can tolerate - in other words, decide how his decision will vary, depending on the range and accuracy of the forecast. This allows him to trade-off cost against the value of accuracy in choosing the modelling methodology.

Figure 2 shows how cost and accuracy are likely to vary and graphs this against the corresponding cost of a model's errors, given some general assumptions. The figure intends to show that ideally the policy maker must hope to weigh the cost of a more accurate (expensive) model against that of a less accurate (inexpensive) model or a different modelling methodology.

The most sophisticated model that can be economically justified is one that falls in the region where the sum of the two costs is minimal. There may be some debate as to the extent that this approach might apply to the vast array of energy models already in existence. There are also a host of other questions such as whether a more complex or disaggregated model structure is necessarily more accurate than a simpler, highly aggregated structure. However, without some notion of a model's validity or accuracy we cannot begin to think of applying it.

Surveyed briefly here have been the elements of model validation and verification, and approaches to independent evaluation. It should be emphasized, however, that the modeler has the primary responsibility for designing, conducting, and reporting the results of validation and verification efforts. When the modeling research is reported in refereed journals and monographs, the modeler will be guided by general scientific practice. However, modeling research intended to be used in planning analyses, and often not submitted for refereed publication, places an additional burden on the modeler. In addition to providing documentation

consistent with "good scientific practice," the modeler must also satisfy the information needs of a (usually) less technically oriented audience interested primarily in model based results, not in the modeling research itself.

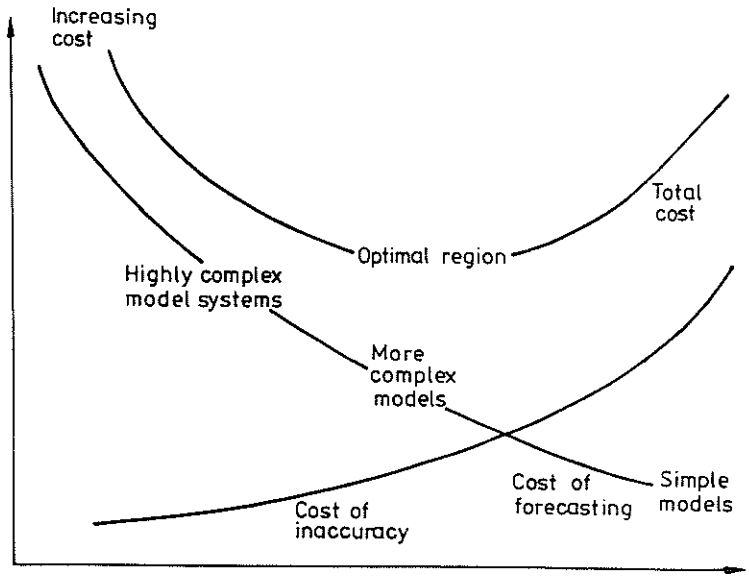


Figure 2. Trade-Off Between Cost and Accuracy

The essential requirement is to provide these diverse model audiences - other modelers, policy analysts, and those affected or influenced by model based studies - with the information to judge for themselves the adequacy and legitimacy of the models, and their relevance in any particular applications. No amount of independent evaluation can substitute for a well conceived and executed plan to satisfy these legitimate information requirements. Hence the prospective modeler intending to contribute to the analysis and understanding of the current and prospective energy policy issues should be aware that good models and analysis are not sufficient to ensure the credibility and utility of model based policy studies. The findings of Greenberger and his associates, mentioned at the beginning of this section, that for the major energy studies of the past decade (and many relied heavily on models), perceived quality was negatively correlated with attention received and influence, is a sobering reminder and challenge to energy modelers to consider carefully, and to satisfy, the legitimate information needs of their audiences.

Validation in the Modelling Process

Engineering-economic modelling can be generally described as a process consisting of four interrelated activities: model specification, parameter estimation, simulation and forecasting, and model validation. In addition to being sequential, the interrelation can be thought of as involving feedback. During the validation process, additional shortcomings of a model may be detected and reformulation and estimation may again be necessary. It would be extremely informative for users of engineering-economic models, if some standardized validation approaches could be employed and appropriately documented. Following the presentation of validation criteria in this paper, the following three-step procedure might be considered.

(1) Model specification and content:
- Identification of relevant policy instruments useful for the model's intended application.
- Selection of appropriate hypotheses to test.

(2) Model Structure and estimation:
- Coefficient tests including structural change
- Goodness of fit tests
- Model selection
- Pseudo forecasts
- Sensitivity analysis

(3) Model Simulation and Forecasting:
- Single point criteria
- Interval criteria
- Error-cost analysis
- Turning point analysis
- Comparative errors
- Error decomposition
- Sensitivity analysis

What are some of the problems we face in applying such a battery of criteria. First of all, validation is difficult to imagine for large-scale energy models with many technically determined coefficients such as those including programming methodologies. In addition, most of these models cannot be run as a total machine model. They sometimes require the intervention of the submodels to 'massage' the data to get a full run of the system. Finally, many of these planning models forecast many years into the future (2000 or beyond).

Earlier the economic criteria in the selection of alternative model structures had been mentioned. The criterion places pressure on energy modelers to undertake validation even under the most stringent of conditions. For example, Fromm et al. (1974) point to the possible development of Federal standards for model validation and some of these standards are reflected in the Gass (1981) summary of model validation perspectives. In that context validation was seen as a means of avoiding or discarding bad models and enhancing the credibility of good ones for policy applications. To a large extent, the Office of Statistical Standards of the U.S. EIA has made progress in the energy model audit procedure of the Quality Assurance Division, e.g., see Cohen (1986).

If we are going to pay more than lip service to validation, what should it involve? In the case of large-scale, long-term models, we should examine the internal consistency of a model as well as its consistency with theory. Also of interest is a comparison of the results of a particular model with the results of other models or of other surveys that purport to measure the same factors. This comparison need not only relate to the level of the variables; it can also embody a sensitivity analysis of the response of the alternative models to similar perturbations. This validation process implies not only validation in the nature of comparative results but also taking a serious look at the data and documentation to see if variations can be explained reasonably.

Many of these factors have been embodied in the overview model assessment of Wood (1981). The assessment steps specified include: review of the literature, overview model assessment, independent audit, and in-depth assessment. These steps are viewed as interactive and complementary. They involve not only comparing different model runs based on similar data and assumptions but also assessor-modeler interactions.

For medium and short-term models of a moderate scale, increased use can be made of the parametric and non-parametric validation techniques described above. Where temporal energy activities are being described, techniques such as goodness-of-fit, single point criteria, turning point analysis and comparative errors can be employed. Where possible, validation should take place separately in the ex-post or sample period from the ex-ante or prediction period. Such a suggestion obviously applies better to econometric models than input-output or programming models. For engineering-economic models, greater development and application of sensitivity analysis would seem necessary.

For the future we await more demanding criteria which could be applied more meaningfully to model validation. For the present we hope that the criteria mentioned here can be applied fruitfully and effectively.

REFERENCES

Chambers, J.C., Mullick, S.K. and Smith, D.D. (1971). "How to Choose the Right Forecasting Technique," Harvard Business Review, XX: 45-74.

Chen, D.T. (1981). "Accuracy in Commodity Price Forecasts with the Wharton Monthly Model: An Error Decomposition Analysis," Presented at the Conference on Commodity Price Analysis and Forecasting, Ames, Iowa.

Cohen, S. (1986). "Model Verification," Presented to the American Statistical Association Committee on Energy Statistics, Washington, D.C.

Crissey, B.L. (1975). "A Rational Framework for the Use of Computer Simulation Models in a Policy Context," Ph.D. Dissertation, Johns Hopkins University, Baltimore.

Fromm, G. (1973). "Policy Decisions and Econometric Models," Presented at NSF Seminar on Modelling, Washington, D.C.

Gass, S.I. (ed.) (1979). Validation and Assessment Issues of Energy Models, Proceedings of an NBS Workshop, U.S. Department of Commerce, Washington, D.C.

Gass, S.I. (ed.) (1981). Validation and Assessment of Energy Models, (Proceedings of an NBS Symposium. U.S. Department of Commerce, Washington, D.C.

Greenberger, M. (1983). Caught Unawares, Cambridge, MA: Ballinger Publishing Co.

Greenberger, M., Crenson, M.A., and Crissey, B.L. (1976). Models in the Policy Process, New York: Russell Sage Foundation.

House, P. (1979). "What's Really New About Modelling?" Journal of Policy Modeling 1: 159-178.

Labys, W.C. (1982). "Measuring the Validity and Performance of Energy Models," Energy Economics 4: 159-168.

Labys, W.C. and Wood, D.O. (1985). "Energy Modeling," in W.A. Vogely (ed.) Economics of the Mineral Industries, American Institute of Mining Engineers Inc., New York, pp. 303-336.

MIT Model Assessment Group (1979). Independent Assessment of Energy Models, Prepared for the Electric Power Research Institute, Report No. EA-1071, Palo Alto, CA.

Taylor, L.D. (1975). "The Demand for Electricity: A Survey," Bell Journal of Economics and Management 6: 74-110.

Wood, D.O. (1981). "Energy Model Analysis and Evaluation: Current Practice," in S.I. Gass (ed.) Validation and Assessment of Energy Models, National Bureau of Statistics, Washington, DC: Government Printing Office.

Energy Vol. 15, No. 3/4, pp. 163–170, 1990
Printed in Great Britain

0360-5442/90 $3.00 + 0.00
Pergamon Press plc

1.4. CREDIBILITY OF MODELS

YOICHI KAYA

Department of Electrical Engineering, University of Tokyo
Tokyo
JAPAN

Abstract - The paper presents a study on model credibility, which was completed in Japan. More than 10 economywide models were selected and GNP results compared. The economists found that the model with the smallest prediction error was a very simple model: a linear extrapolation of the trend in the previous year. In other words the hypothesis that the GNP growth rate during a year is the same as the rate during the previous year proved to be an accurate forecast. Undoubtedly not only modelers but those who are not familiar with modeling methodology would be surprised.

Let us begin with a short review of economic forecasting in Japan. Econometric models are today widely used as tools for forecasting, and a number of research institutions, think-tanks and city banks regularly publish their own forecasts using the models that they have developed themselves. One economist has tried to compare the results of these different forecasts with actual economic data in order to evaluate the usefulness of the models as forecasting tools. During the 1960's the Japanese economy experienced a high rate of smooth GNP growth and this growth rate was used as the variable for the purposes of comparison.

More than 10 models were selected and compared and the economist found that the model with the smallest prediction error was a very simple model - a linear extrapolation of the trend in the previous year. In other words, the hypothesis that the GNP growth rate during any year is the same as the rate during the previous year gives the best fit.

Can it be said then that a simple linear extrapolation of the past trend is the best model for economic forecasting? Undoubtedly not only modelers but also those who are not familiar with modeling methodology will say 'no'.

This illustrates that the credibility of models does not rely solely upon predictability. According to the author there are two types of measures for enhancing model credibility. There are those for conveying model information to users, and those for model validation. The former may be called 'external' measures, while the latter may be called 'internal' measures. The elements of these measures are shown in Figure 1, together with factors which may disturb the effectiveness of the measures. The following sections deal with the details of these elements.

Measures for Model Validation	Measures for Conveyance of Model Information to users	Disturbing Factors
statistical analysis	efforts for more transparency	mere escalation of model size
efforts for causal model more valid	better illustration of causal structure	non-modelable factors
analysis on simulation runs	calculation of key indicator	
sensitivity analysis	illustration of key outputs	

Figure 1. Measures for and Disturbing Factors to Enhancing Model Credibility

Purpose of Model Use

It is a fact that some models are used purely as forecasting tools. A typical example is the electric power demand forecasting model which is used for the prediction of demand for the next few hours or at the longest a day ahead. The forecasts are then utilized for the optimal dispatching of demand in different power stations within the network. Since the optimality of the operation costs depends heavily on the difference between predicted and actual demand, the predictability of the model is a key factor. A number of models have been developed for this purpose and most of them show a high goodness of fit with actual data.

However, we should note that the above models are built in exceptional circumstances. Enormous amounts of data are available so that various statistical methods may be applied to build models with high goodness of fit. Models thus built can be revised continuously so as to adapt themselves to structural changes of the system. In the case of ordinary socio-economic and energy modeling, data availability is much worse. The number of samples is usually 10-20 or even less, so that only the simplest least squares method is applicable to the estimation of model parameters. Consequently, the goodness of fit of some parameters may be low and the overall model may appear to be unreliable from a statistical and theoretical point of view.

We are in most cases not capable of using models as pure predictors directly connected to decision making, but are restricted to using them as tools that provide useful information on the simulated system. For planners and/or decision makers in companies and governments, this information is useful as it provides guidelines in the decision making process.

To be more specific let us consider two types of models - one optimization oriented and the other simulation oriented. The former type is widely used in energy modeling to express the optimum supply-demand relations in energy systems; usually the linear programming scheme is employed. Since the planning of a system is in a sense an optimization process although the criterion is different from planner to planner and in some cases not explicit, optimization models can be considered to be those that provide the ultimate desirable images of the planning target. Simulation type models on the other hand can be considered to be those that provide future scenarios of the target system built under the assumption that the past trend will continue and that the structure of the system will not change. The decision makers postulate certain prior conditions and then receive output from the model as to what will happen in the future given different policy decisions.

The extent to which the information and/or scenarios will be utilized in decision making depends upon how reliable decision makers feel the models are, in other words, upon the

credibility of models. The questions then arise - what kind of models do decision makers feel to be highly credible and what are the factors contributing to model credibility?

Measures for Conveying Information to Users

The above questions are inevitably connected to use of pure statistical models. There have been some attempts to apply the techniques of time series analysis such as AR and ARMA models to socio-economic modeling. These methods are designed in essence to express the present value of a variable as a linear or nonlinear function of its past values and then to evaluate the parameters of the functions from the time series data. In some cases the goodness of fit of the estimate is high and therefore the estimate seems to be statistically reliable.

I believe that most users are reluctant to call such models credible. If these models are to be used for only the very short term, predictions may provide very good estimates of variables with only small prediction errors. But in the medium and long term, as the actual data available to us is limited and there may be many changing factors in the system of which some are structural, these models are almost useless.

The reader will realize what this means of the review of GNP forecasting presented at the beginning of this chapter is recalled. Most readers may agree for example with the statement that GNP of the current year will grow by the same rate as in the previous year. However, in cases where little information on the national economy is available, very few readers may agree with the statement that GNP for the next 5 years will grow by the same rate as in the past several years. Only when they are informed of the trends of other economic variables and their contributions to GNP may they agree with the statement. This suggests that these models, which may be called descriptive models (Lewandowski, 1982), are of very limited use, whereas causal models involving causal relations among system elements are far more useful in most cases.

Given a causal model for the target system, users need to try to realize the characteristic of the model with the help of modelers. The deeper their understanding, the more credible the model will be.

Several methods have been used for this purpose. The first method is to explain key causal relationships in models descriptively or by illustrations. If users understand these relationships they may accept the results of model simulations not as those from crystal balls but as reasonable consequences of the causalty they see in the target system. A typical example of the use of illustrations is found in the famous report to the Club of Rome - "The Limits to Growth" (Meadows et al., 1972) which investigated the long term future of global issues (called 'problematique') including factors such as resources, food, environment, population and industrial development by use of System Dynamics methodology.

The model includes more than a hundred variables but one of the key causal relationships is expressed in Figure 2. It represents the relationship between population, industrial development and resources. If industrial development is accelerated it will give rise to an increase in service capital, which will then reduce both the death and birth rate. On the other hand, population growth will contribute to the reduction in industrial production per capita. The model can examine quantitatively which effects are larger. The report devoted a number of pages to the illustration of causal links amongst the system elements. The validity of the forecasts made in the report is even now still controversial but none will deny that these illustrations have greatly helped the readers of the report to understand the mechanism of the model.

The second method is to describe key characteristics of models in a simple way. In the case of energy systems planning, most decision makers are concerned with the elasticity of energy demand with regard to GNP and in some cases with price elasticities of energy. These values may be calculated from the model simulation runs and exhibited to decision makers, ideally with some comments from modelers on the differences of the values from those of other models and

also why they are different. Model users may then grasp the essential characteristics of the model without having to survey the details of the model structures.

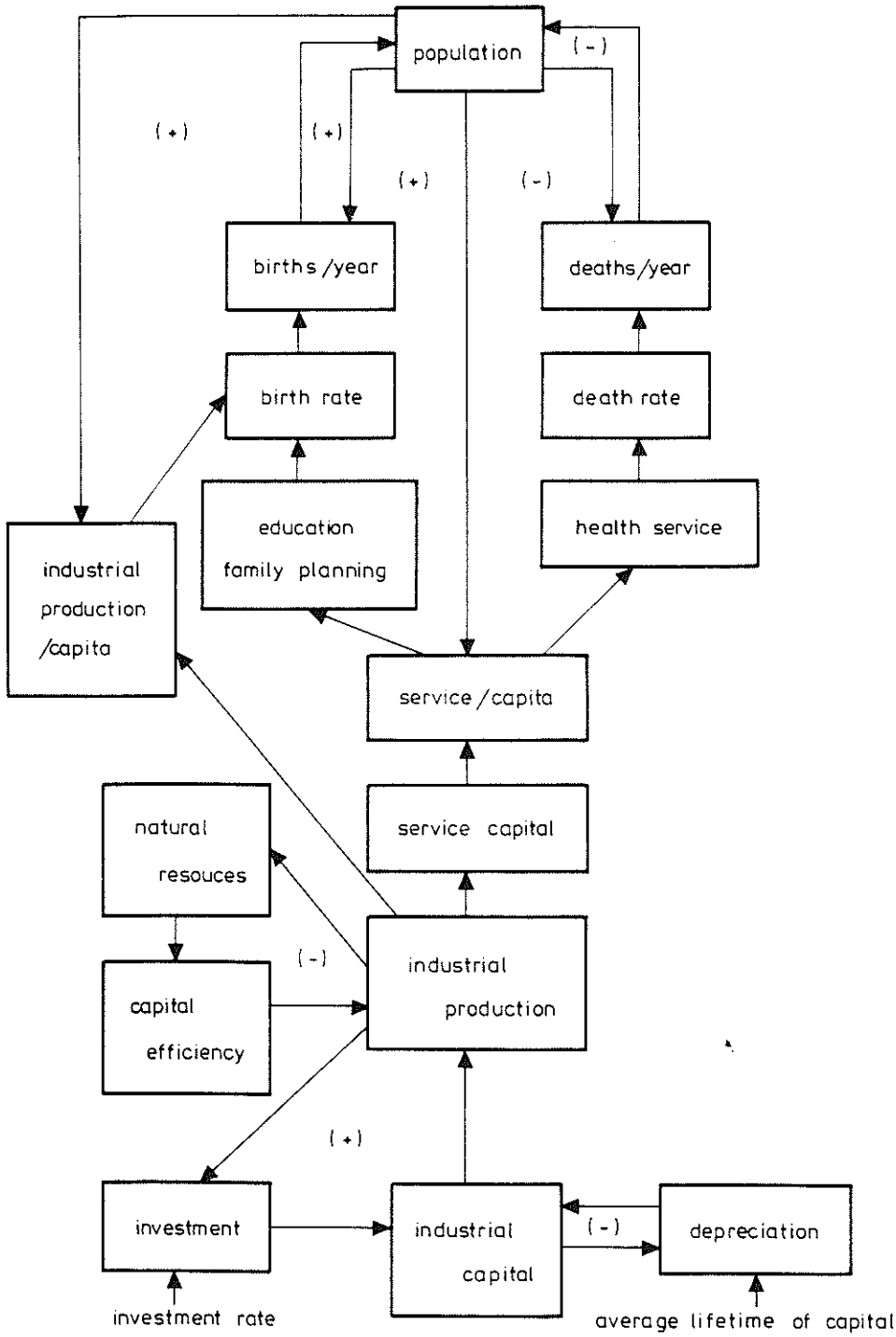


Figure 2. Feedback Loops Among Population, Capital, Services and Resources

The third method is to make parameter sensitivity analyses and show how variant the simulation results are. For example if a user is concerned with the economic growth of the national economy, modelers can show him a number of simulation runs that involve changing a

few key parameters such as oil prices. He will then see how uncertain growth is and what parameters are influential. An example of this kind is seen in the energy carbon dioxide model developed within the National Academy of Science (CDAC, 1983). In this model about 10 parameters are chosen and three values for each parameter are estimated corresponding to 5%, 50% and 95% points of its probability distribution (note that in this case probabilities are subjectively estimated by the modeler). Simulation runs are executed for each of 100 parameter sets randomly drawn from possible 3^{10} combinations of the above parameter values. Resultant atmospheric concentration of carbon dioxide is shown in Figure 3. It is seen that variability of its future values is high and the possibility of its very slow growth, which is undoubtedly desirable, is low although still present. This is certainly a useful piece of information when discussing anti-carbon dioxide measures.

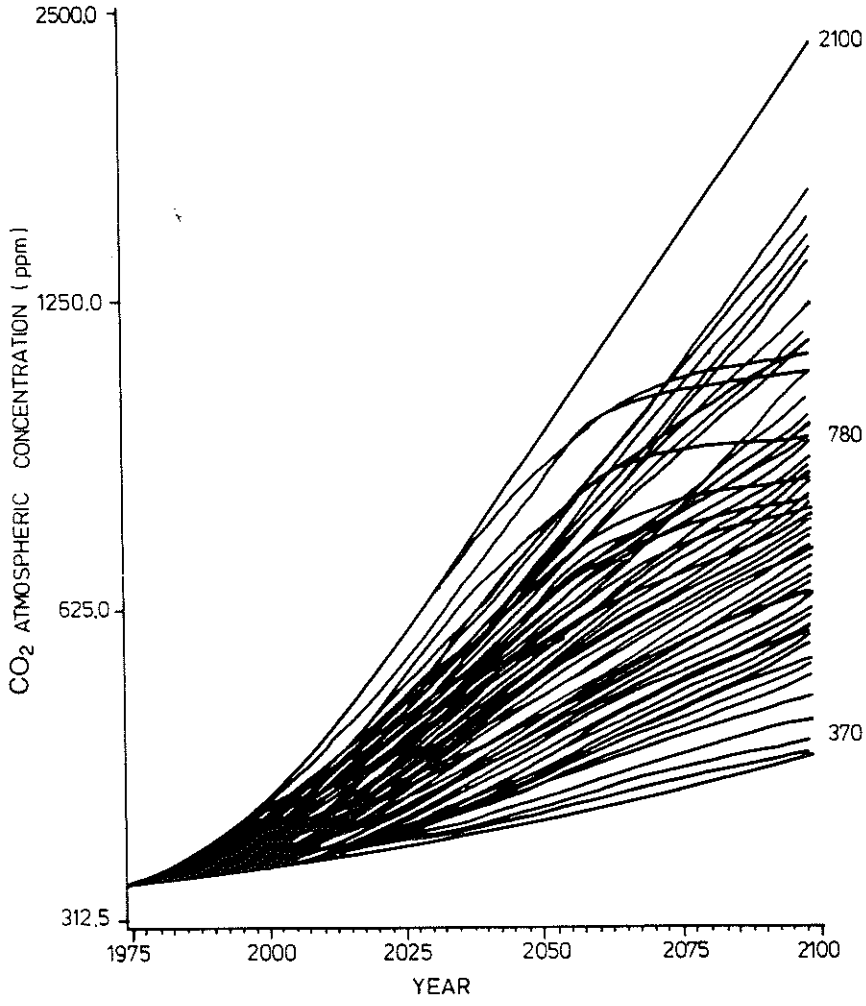


Figure 3. Uncertainty in $C0_2$ Concentration

Results: From the model developed by the US National Academy of Science for 100 randomly drawn runs.
Source: Changing Climate, National Academic Press, 1983.

The above measures are intended to provide model users with summarized information on models. In practice, users may raise a number of questions which are connected to the details of models. Modelers are responsible to answer these questions if requested but it is even more

important to leave models as transparent as possible so that the model users may examine details of models by themselves. In other words, as much information as possible on the model structures, equations and parameter values with statistical data (correlation coefficient, t-statistics etc) and data sources should be transferred to users. This is a fundamental requirement to all modelers. Quite often model users ask modelers to transfer only simulation results and supporting information is only wanted upon user request. It is an important requirement which contributes to the credibility of the model.

Validation of Models

The preceding section has dealt with measures for enhancing model credibility by delivering as much condensed information as possible to model users. However the basis of all measures lies in the modelers efforts to build models as valid as possible. In connection to the discussion in the preceding section, three types of modelers'efforts are notable. The first is the effort made for model validation through the careful examination of model equations. Ordinary statistical theory, both for the estimation and testing of parameters is certainly a basic tool for this purpose. High goodness of fit with the data is only one of the factors of model validation. It is not rare that any of several different types of equations show high goodness of fit, (R^2 and parameter variances), for the same endogenous variable, (the variable in the left hand of the equation), and the same data. Shown below is an example of this kind. The endogenous variable here is the automobile demand in Japan (cars per year), and the following two equations have been derived.

$$\begin{array}{lll} \text{ANPH} = & \text{-.26432E-01} & \text{-.77714E-01*APPH(-1)} \quad \text{+.16851*LICH} \\ & \text{(-2.1)} & \text{(-2.6)} \qquad\qquad\quad \text{(5.7)} \end{array} \tag{1}$$

$$\text{R*R} = \quad .9374 \qquad \text{D.W.} = 1.6$$

$$\begin{array}{lll} \text{ANP} = & \text{-1.8284} & \text{-.83249E-01*APP(-1)} \quad \text{+.34162E-01*GNE} \\ & \text{(-4.9)} & \text{(-3.5)} \qquad\qquad\quad \text{(7.9)} \end{array} \tag{2}$$

$$\text{R*R} = \quad .9738 \qquad \text{D.W.} = 1.9$$

where
ANP: new automobile demand per year
ANPH: ANP per household
APP: number of automobiles being used
APPH: APP per household
LICH: numbers of licensees per household
GNE: gross national expenditure
*(-1): the value of * in the previous year
R*R: R^2
D.W.: Durbin-Watson ratio

Numbers in the parentheses are values of t-statistics, and the data length is 14 years.

It can be seen that both equations show high goodness of fit and satisfy sign constraints of parameters (for example the coefficient of GNE in Meadows et al. (1972) should be positive). It seems appropriate in this case to adopt the second equation in the model, as it shows a higher goodness of fit than the first equation. However, the first equation was adopted because the modeler thought that the causality embedded in the first equation was more appropriate and particularly so with respect to future forecasts (i.e. the number of licensees in a family was seen as a more causally relevant variable than the level of income in explaining future automobile demand).This example shows how much effort modelers need to make when adopting equations which they think would be adequate from various points of view. This selection process of variables and equations is in a sense the heart of modeling particularly of socio-economic systems. Model users should never overlook the importance of this process.

The second method of model validation is examination of simulation runs. Simulation runs of a model can provide a lot of information on its dynamic characteristics. Even such models that produce outputs with high goodness of fit with the data can be dynamically unstable, and simulation runs are the only means to check model stability, particularly in case of large scale and/or nonlinear models.

The sensitivity analysis mentioned earlier is also a part of these sorts of validation efforts. By changing parameter values and providing corresponding model simulation runs, we see how sensitive the model behavior is to changes in model parameters. If the model is highly sensitive to a certain parameter, a more elaborate study on the value of that parameter is required. It takes time and effort but is indispensable in the process of building credible models.

Disturbance Factors

This section deals with factors which may disturb model transparency and therefore reduce its credibility.

The first factor involves a scale-up of the model size. Progress in computer technology has enabled the construction of large scale models with more than several hundreds variables. It seems better to build models as large as possible in order to describe details of socio-economic systems; some economic models are constructed with more than ten thousand equations and variables. For example, a multi-national model of 80 countries, of which each economy is expressed by 30 equations (as a country model it is rather a small size), has to include at least 80*30 = 2,400 equations. We should notice that mere large scaling may disturb the model validation process seriously. If a modeler wants to build a model with more than 2,400 equations, it is almost impossible for him to examine each equation carefully and revise it in an elaborate manner. There are still a few good large scale models in the world which were built very carefully, but they were built with the co-operation of a number of modelers (The famous LINK model originated by Prof. L. Klein is such an example). Therefore, modelers should always be cautious not to build unnecessarily large scale models, at least until a novel AI technology is developed for automatic checking and revising model equations.

The second problem is the existence of non-modelable factors. In building models we quite often encounter the situation where data of certain variables seems hardly available, although these variables are sometimes indispensable in the analysis of the target system. There are two cases: one, where data is tangible but hardly available, and another where variables are intangible.

Non-commercial energy in developing countries is a typical example of the former type. In some developing countries, non-commercial energy occupies more than 50% of total energy demand and modeling of non-commercial energy demand would seem to be essential in overall energy modeling in these regions. However, by definition non-commercial energy is the energy consumed outside of the marketplace and existing data generally consists of only rough estimates taken from small sample surveys. It is almost impossible to build models based on such data and there are few models to date which involve non-commercial energy sectors.

The latter case is typified in the treatment of the public acceptance of nuclear power plants. It is considered to be one of the key factors determining the future penetration of nuclear plants, but there have been few attempts to quantify degrees of public acceptance. This factor is therefore usually neglected in models (in some cases appropriate proxy variables may be used).

The problem of non-modelable factors is certainly not due to the idleness of modelers but a fundamental problem in modeling of socio-economic systems. The model users should realize

that a model is just a partial picture of the target system and therefore the model can provide useful information only within that picture.

There may by many other factors influencing model credibility but the most serious among them is, according to the author, the attitude of impatient users who are interested only in simulation results but not the in details of the models. As described here an understanding of model structures and model characteristics is the most effective way for deriving useful information from models and the author firmly believes that the credibility of models depends heavily upon the users understanding of this point.

REFERENCES

Lewandowski, A. (1982). "Issues in Model Validation," IIASA Paper R-R-82-37.

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

Report of the Carbon Dioxide Assessment Committee: Changing Climate (1983). National Academic Press.

Energy Vol. 15, No. 3/4, pp. 171–185, 1990
Printed in Great Britain

0360-5442/90 $3.00 + 0.00
Pergamon Press plc

1.5. RATIONAL APPROACHES TO IDENTIFYING POLICY OBJECTIVES

PAUL J. WERBOS[+]

National Science Foundation
Washington, D.C.
U.S.A.

Abstract - This paper focuses on the broad problem of specifying objectives or utility functions when making decisions about energy or related fields. Decisions of all sorts from the design of analysis through to global policymaking all depend on objectives, implicitly or explicitly. This paper begins by describing how objectives are commonly specified in bureaucracies, and how this affects decisions in practice. It evaluates these conflicting approaches, and suggests a way of trying to bring them more into harmony. It then applies this approach by developing substantive priorities which I would propose as a first cut for rational decision-making related to energy. The substantive analysis begins by revisiting the issue of discount rates, and suggesting that lower rates should be used than are commonly used today. It then discusses what we can learn from global energy models regarding the possibility of (and prerequisites to) long-term economic growth; it begins with a brief summary of the substantive lessons to be learned, and procedes to a discussion of how these models could be used to help establish priorities in more detail.

1. SCOPE OF THIS PAPER

This paper will discuss the problem of defining objectives. Choosing a set of objectives or a utility function requires that we make assumptions about values and utilities, which are very difficult in the real world. Therefore, this section will begin by discussing this problem in general terms, and evaluating the kinds of approaches which have been used in the past. Then the bulk of this paper will get down to specifics, by discussing substantive issues and variables which should be reflected in major decisions about energy.

The views expressed here are truly those of the author, rather than those of his employer. The section of government in which I work does not formulate objectives or make decisions about energy as such, and some of my views are probably different from the official views of the U.S. Department of Energy (DOE). Likewise, the examples below are not taken from the unit which I work in.

[+]This paper was submitted prior to the author's association with NSF therefore, it does not reflect on NSF or its views in any way.

2. A STATEMENT OF THE PROBLEM

Whenever humans take action, they face the task of deciding which actions best meet their objectives. Whenever they carry out this task, the results will depend very heavily on their choice of objectives. This section will focus on the broad problem of specifying objectives when making decisions about energy.

This section will talk about one narrow definition of the problem of defining objectives: the problem of specifying a utility function U when using explicit optimization methods. (Some methods would call this an objective function or whatever, but the terminology makes no difference.) Nevertheless, the discussion here has much broader implications.

Any analysis of energy decisions must relate them back to objectives, explicitly or implicitly, to be useful at all to decision-makers. Explicit optimization methods are often unpopular, because the ultimate decision-maker does not trust the modeller to account for all of his concerns. This mistrust can be overcome in part by considering a range of possible objectives and assumptions, by showing how they would lead to different strategies of action, and by explaining very carefully why the computer arrives at the recommendations it chooses. One can also use optimization methods implicitly, as a way of identifying strategies whose results are then discussed in more objective terms; one can perform sensitivity analysis to show how the bottom-line results are changed by slight adjustments in strategy. The validity of such an implicit analysis -- generated by a computer program or by a human brain -- still depends on the validity of the assumed objectives. Even simple cost-benefit analysis and cause-and-effect discussions, based on the use of simple arithmetic and graphics, may be useful or misleading, depending on the choice of target variables (and strategies) represented. Thus whenever an analysis of objectives is not performed, the analyst is leaving the validity of his results to chance. Again and again I have seen cases where complicated analyses were largely wasted (or abused) because the analysts "answered the wrong question." In the Defense area, one sometimes hears of efforts which move optimally, quickly, and efficiently over the nearest cliff. In summary, the problem of defining objectives is not confined to formal optimization methods, and it is critical to the validity of analysis.

3. HOW THE PROBLEM IS USUALLY DEALT WITH

There is a bewildering variety of methods in use for specifying utility functions. Some variety is necessary, because users in government, academia and business have different needs; likewise, there is a difference between people who use optimization in order to make recommendations and people who use it as a predictive tool (assuming that someone else will act to maximize something). Commonly used approaches include:

1. Ask your boss. Analysts often prefer this approach for two reasons: (a) it fits the idea that they only do objective, scientific work and let other people worry about values; (b) it fits the idea that they are strictly tools of the state (or the corporation), whose values are specified from above. When you take this approach, you are giving your boss the problem of choosing between the remaining approaches (unless he passes the problem in turn up to his boss, etc.).

2. Draw up a short list of alternatives, and use conventional bureaucratic procedures to make a choice. The list would consist of simple, obvious variables which are commonly discussed (such as oil imports in 1995 or residential gas prices next year) or used by other analysts in their optimization work. The choice is commonly based on what would be most acceptable, popular, simple, easy to implement, noncontroversial, familiar, and within the general limitations of scope and vocabulary that one's organization generally deals with. (This approach is rather like guessing what your boss would say if you asked him; it still involves passing the buck to some extent.)

3. Try to be Objective. For example, put together data on every variable of possible interest, and do a factor analysis. This method is quite popular in groups where neither the analysts nor the bosses have experience in thinking hard about social problems. About $80 billion in civilian funds were allocated on the basis of such approaches in a recent year; also, this approach has been used to construct a measure of the national interest for use in policy evaluation high up in the Defense community. Well-known, authoritative statisticians have endorsed this approach because of its objectivity. (There may sometimes be problems, however, in deciding what sign to put on the resulting value measure.)

4. Use Bayesian methods to articulate the values which you or your boss already possess. This approach assumes that the formulation of utility functions, like the formulation of models, is largely a matter of translation from verbal or subjective understanding into mathematics. (This concept of translation is also discussed in the chapter on econometric methods.) Bayesian approaches have become very popular in industry, thanks to the efforts of Prof. Howard Raiffa (1968) and Brown (1974) of the Harvard Business School. They have also been refined and modified for use in government decision-making, as in the work of Decisions and Designs, Inc., of Tyson's Corner in Virginia(1978). These methods usually involve questions which an analyst asks of a higher-level decision-maker; such questions may sometimes encourage a dialogue, in which the decision-maker may make discoveries about his or her own values. Typically, these questions ask that you visualize two sets of possible scenarios, with probabilities attached, and decide which set you would really want to face.

As a practical matter, good engineers know how to start this kind of dialogue, when optimizing a system design; they start out by analyzing what kinds of tradeoffs are likely to exist, and asking their superiors or clients what they would prefer if faced with such tradeoffs (e.g., cost versus reliability). It is important, however, to be sure that the measures of these concepts are appropriate; for example, the kind of "reliability" which sells cars may not be the same as what the engineers usually think of when designing cars.

5. Try to translate the official missions and functions of your organization into a mathematical measure. This may be especially easy in corporations where people are located in profit centers, with a definite quantitative performance measure. Likewise, it is relatively easy in organizations like the Department of Energy which publish a long-term policy plan complete with a statement of objectives. Understanding and empathy with the authors of such plans may make the translation better, in some cases, although there may be problems if your own management has a different interpretation or a desire to pursue different goals (e.g., supporting the special group responsible for their employment).

After the translation is done, there will still be a problem in moving from the long-term goals of the organization to the more specific variables you are facing. If the decisions you are analyzing are only a subset of the decisions faced by your organization, then you are really dealing with an optimization problem within an optimization problem. You cannot ignore the greater optimization problem which your own problem is embedded in. For example, if you are designing a power plant for a profit-making company, you have to figure out the relation between attributes of the power plant and the resulting profit. This is really part of the larger decision problem, in which the company chooses power plant characteristics, load management strategies, export strategies, and prices over time; the relation between one power plant and the resulting profit is not a well-defined relation, unless all the other power plants and the larger marketing strategy are taken into account. If your organization deals with the larger optimization problem explicitly, and maintains models and forecasts which show how local changes impact the long-term objectives, you can use that information explicitly; otherwise, you have to make crude guesses yourself about how the greater optimization problem works. (These guesses should be discussed, of course, with other people.)

6. Try to be democratic. Decisions about values can be passed on to the people most affected, on grounds that their satisfaction is the long-term goal of policy or on grounds that they have a

better understanding of their part of the greater optimization problem. The classic way of doing this is to set up a market mechanism, in which markets select prices which can then be used as a measure of value. Other methods have been developed, however, for situations where a market mechanism is difficult to set up. For example, the Census Bureau, in contract research for the Farmers Home Administration (FmHA), used a method called the Group Value Measurement Technique (1979) to determine the values of people living in rural areas; FmHA then used these values to guide it in allocating loan money. GVMT combined some of the question-and-answer concepts of the Bayesians with some of the small group psychology used in Delphi Analysis.

7. Try to translate a philosophy of government into a mathematical measure, and then cope with the greater optimization problem as with alternative number 5 above. For example, some students of Julian Simon have suggested that all decisions should be guided by one maximization problem: the problem of maximizing the quantity of human life.

Many economists have argued that the free market should be the ultimate arbiter of values in a free society. In this philosophy, consumer surplus (in various forms) is the proper measure of benefit to use in making all decisions. The market is not assumed to be perfect in practice; its imperfections -- externalities due to market failures -- can be measured, and added in with one's estimates of consumer surplus.

Some economists think of GNP growth or poverty or a similar measure as an ultimate standard of value. Followers of the Darwinian tradition would argue that some measure of long-term survival best reflects what we are all maximizing anyway, with varying degrees of competence and self-awareness. Many others have argued that cultural or spiritual values should also be cranked into all decisions; however, the difficulties and complexities of that argument go far beyond the scope of this book.

This approach has at least two advantages over approach number 5. First, it bypasses the utter ambiguity of certain statements of missions and functions. Second, it avoids the conflict of objectives which results when different agencies use different utility functions. A basic textbook on military operations research by Hitch shows how different utility functions (value weights) in different units can lead to situations which are inferior by both measures of value; ironically, the same kind of reasoning suggests that defense programs as a whole should be judged in part for their nondefense implications, such as contributions to unclassified basic research or contributions to making poor people employable outside the military. On the other hand, the greater optimization problem here is far more difficult than with alternative 5.

4. AN EVALUATION OF THESE APPROACHES

A mathematical or ideological purist would try to decide which of these seven approaches is correct, and elaborate its requirements in detail. Such purists would usually endorse the seventh approach, which emphasizes the responsibility of each individual for the long-term ethical implications of his or her own work.

Many philosophers and psychologists would agree with this view in general terms, but would argue that the seventh approach is unworkable until we first learn to express basic, long-term values which are human rather than political in nature. This expression involves something like the fourth approach, applied to basic values rather than short-term preferences, and helped by the use of psychological theory in support of self-understanding. Political ideology is relevant only as a crude, initial approach to solving the greater optimization problem.

In practice, no one working for a large organization has complete freedom or responsibility for the decisions of that organization. Academic analysts usually have no control at all over the decisions which they analyze. Therefore, it is not realistic to recommend a purely individualistic

approach, like the seventh approach above. (It is realistic, however, to deplore the third approach, which totally abdicates all social responsibility and leaves the results, in effect, to chance.) Analysts in the real world must try to combine all seven approaches (leaving out the third). In other words, they must pursue an integrated approach.

How is it possible to reconcile such different approaches? This section cannot provide a complete answer, but it can provide some suggestions.

In a well-functioning organization, the utility measures which you end up with should not differ too radically between one approach and another. For example, if you ask your boss his preferences, they should be consistent with the missions and functions of your organization, so that the first approach and the fifth approach should give similar answers. If the specific problem you address involves aspects of the agency mission which he is unfamiliar with, he should still prefer that you account for such things after you have shown him that they exist. He should respond better to analysts who empathize with the larger optimization problems he is facing than to employees who are totally cynical about him, as long as the former are sincerely striving for two-way communication. Likewise, the people who construct the agency policy plan or monitor agency actions should be responsive to feedback from analysts who detect long-term impacts which are not yet part of the plan, as long as the feedback is transmitted constructively. In general, it should be possible for each analyst to take the seventh approach -- the ethically responsible approach -- and to interact with other parts of the organization on that basis; the difficulty, in practice, is to learn which parts of the organization are really analyzing the greater optimization problem, at what level, as opposed to those people who are taking their views from other sources.

This general approach does differ with the purely hierarchical view of organizations. It is based on a view of intelligent systems, which concludes that no organization can operate like an intelligent system unless there is a web of information flow and feedback based on substantive, logical connections across all parts of a system (Werbos, 1986 and 1987). The correct utility function for a subsystem is simply the sum of relevant concepts, weighted by dynamic feedback coming from analysis in other parts of the system. Students of organization theory have long observed that there is an informal network of information flow quite different from the formal hierarchy, in effective organizations. This is also reflected in the concept of matrix management, which tries to make these cross-cutting flows of information more explicit. Democratic pluralism can have similar benefits. In the limit, the theory of intelligent systems suggests a need for realistic collective foresight, in which different parts of an organization contribute to the development of coherent future scenarios and probability assessments, explicitly enough that mathematical methods can be used to estimate the sensitivity of the results to local results and assumptions. Whatever its other flaws, the Global 2000 Report (Barney, 1980) was a start in this direction.

Deep conflicts of value can eliminate the possibility of this kind of cooperative organization, especially if everyone becomes fearful and secretive to protect personal interests. Conflicts of viewpoint interfere more often with cooperation in practice, to a greater extent than they need to. If analysts are trained to truly appreciate their level of uncertainty, and to respond to empirical reality testing, they should be better able to appreciate different points of view, at all levels in an organization. Likewise, the integrity of information flow can be maintained by providing information which is as insensitive as possible to the point of view; for example, multiple scenarios can be run representing different viewpoints. (This does require some faith by the analyst that accurate, objective information will improve decisions, more so than distorting the evidence in order to sell a point of view; corporate culture varies greatly in this matter.)

These organizational issues are very important to the actual value of optimization work in the real world. For now, however, it is enough to stress that the seventh approach above is a key part of any rational, responsible formulation of utility measures. The remainder of this

section will attempt to apply this approach, somewhat loosely, to derive the priorities (utility measures) which I personally would consider most important in energy policy.

5. SUBSTANTIVE PRIORITIES FOR ENERGY DECISIONS

When specifying a measure of utility for energy decisions, the time dimension is particularly crucial. Conventional cost-benefit analysis begins by defining a basic measure of utility for each time period; total utility is the sum of utility across all time periods, discounted on the basis of some interest rate. This subsection will begin by arguing that the discount factor should reflect the common-sense notion of caring about the long-term future. Next, it will argue that long-term energy problems require certain priorities, no matter what value systems are advocated, if the premise of foresight is accepted.

The Issue of Foresight

This book is mainly directed to people who live in countries where Christianity or Communism are accepted by the bulk of the people as the ultimate standard of human values. Neither of those philosophical systems provide any justification for attributing less importance to future lives than to present lives. In fact, both systems once led to attacks on "usury," on the practice of charging any interest at all on loans. Similar consequences would follow from any philosophy which demands equal attention to all lives. (Other value systems will be discussed below.)

Basic economics classes in the West often cite dams built in the Soviet Union as a classic example of the irrational decisions which can be made when interest rates are not allowed. They stress that the realities of economic life have forced Communists, as well as Christians, to accept the practice of discounting the future. However, they usually do not go on to ask what a rational decision-maker would do about these economic realities if, in fact, he still accepted the ethical assumption of foresight. They do not explain what to think about bankers who demand a 6 percent real return in their transactions, even as they put the bulk of their money into trusts for great-grandchildren who are not yet born. Commitment to the future might be expressed, naively, by the idea that future GNP or future population should be accepted as a standard of value, without discounting. This results in a utility function which diverges to infinity (we hope), and is not useful in real-world optimization methods or planning. I would argue that this is merely a poor translation of the original ethical principle, and not an invalidation of it. Howard Raiffa has shown that the logarithm of income or GNP is probably a more realistic measure of the utility of money to people. (Strictly speaking, the logarithm has some undesirable properties as well; however, it is essentially equivalent to worrying about the expected rate of growth in GNP. The sum of the growth rates may also diverge to infinity; however, actions in the present usually have increasingly small effects on growth rates in the distant future.)

If a person used the logarithm of income as a utility measure, then he would not act as if a dollar today is only as valuable as a dollar tomorrow; instead, he would discount future dollars according to the rate of growth of real income which he expects from today to tomorrow. In most of American history, this would imply an interest rate close to the rate of growth of labor productivity, which has been in the range of 3 percent. In actuality, real interest rates (prior to risk premiums) moved between 2 and 4 percent for most of American history, until just a few years ago. In other words, the operation of the free market in the United States actually did produce interest rates fully reflective of ethical foresight.

In recent years, this has changed. Interest rates are higher. These changes are the result of deliberate manipulation of the money supply by government, forced mainly by government fiscal policy. The government, by spending much more than its revenue, is signalling to the economy that we face a short-term emergency which requires a short-term rise in interest rates. In actuality, the emergency is purely one of political will in cutting the federal deficit. This

situation may be seen as a classical case of market failure, due to inappropriate government action; interest rates no greater than 4 percent real should be used in long-term decision making, and prime interest rates at that level should be sought, so long as productivity growth does not exceed 4 percent.

This conclusion about interest rates has major implications for energy policy. For example, the latest National Energy Policy Plan Projections show a world oil price of $46 per barrel (in constant dollars) for 2005. These projections are probably not on the pessimistic side; in fact, the projection of U.S. oil production is considerably higher than that of private forecasters, according to Monthly Data Reports from the Office of Transportation Systems in DOE. If $46 in 2005 is discounted at a 4 percent rate, the resulting oil price is well over $20 per barrel, suggesting that oil is very much underpriced at this writing.

The consequences of this are large, but complex. For example, one would still not want to invest $20 per barrel to produce oil which will be sold at $13, even if the oil is underpriced. However, if the oil can be stored in existing storage facilities, which would otherwise have been empty, such an investment would be desirable. The consumption of oil from storage should be discouraged by taxes which reflect the true long-term value of the oil; however, the purchase of oil from hostile nations at a bargain price may still make sense. In summary, different weights should be placed on a barrel of oil, depending on where it is or where it could go. Policy regarding oil may be better understood as a nonzerosum game (as described in Schelling's Strategy of Conflict, 1960) than as a one-player optimization exercise.

The conclusions here about interest rates have important consequences for research and development policy as well. Economists like Denison have used empirical data to show that economic growth is due to research and development, to innovation, and to education, far more than it is to simple capital accumulation (U.S. Bureau of Labor Statistics, 1983). The short-sighted approach represented by high interest rates is especially damaging to research and development (and education), because of the long lead-times involved. For example, most Americans are aware that the recent explosion of a space shuttle was due to problems in the solid rocket boosters. Few of them remember that NASA opposed the use of solid rockets in the beginning, and proposed an alternative design which was expected to have one-third less operations/maintenance costs and even less after a buildup of experience; they were overruled by the Office of Management and Budgets, which used a 15 percent interest rate to argue that it would be better to cut 3 percent from the projected development costs. (This was widely reported in national news magazines in the period between Nixon's reelection and impeachment.)

In general, the argument here is that interest rates are a partial expression of basic values, not vice-versa. There is no magic in economic theory which invalidates commonsense concerns about energy supplies, energy security, and economic growth in the longer-term. Likewise, there is no grey magic in economic theory which reduces the importance of new possibilities for faster growth and new frontiers. A proper assignment of utility functions and interest rates should always reflect the full importance of the long-term future.

To complete this discussion, I will say a few words for those people who do not accept the ultimate validity of any egalitarian social or religious ideology. In the United States, most of these people seem to fall into two groups -- individuals who fully accept the Darwinian view of mankind, and even admit that it would apply to themselves, and people who follow certain kinds of spiritual/mystical path.

Among the Darwinists, there is no simple consensus today about the details of human motivation. However, there is a consensus that human behavior and human thought patterns have evolved to maximize some measure of survival. There is also a consensus that evolution depends on long-term survival, across many generations, and not on short-term survival; it is absurd to imagine a natural organism applying an annual discount rate to its survival. (Uncertainty may lead to a fixation on short-term problems, in some situations, but utility

functions refer to basic values before a strategy is developed for dealing with uncertainty.) From a Darwinian point of view, an organism which does not successfully pursue long-term survival is incompetent. Such incompetence may sometimes result from a lack of learning experience relevant to self-awareness or to the use of words and numbers to express the true concerns of the organism. (Actually, the degree of incompetence is relative. There is nothing wrong with a monkey who puts on a Cardinal's suit because it looks pretty, with no thought of inhibiting its sexual behavior; however, monkeys are not as competent as humans in solving problems, because they have not yet learned the power of true words and symbols.)

There is less consensus present among followers of nonChristian spiritual/mystical paths. Some of these paths seem to advocate an absorption with the present; for example, Taoists talk about "going with the flow" and Zen Buddhists talk about living in the present. Nevertheless, such traditions also tend to emphasize the importance of confronting present problems, in their totality, and treating them as learning experiences. The problem of long-term survival and growth is a present problem, just as surely as the problem of finding food to live; after all, the consequence of not finding food is in the future, not the present, even if the hunger is in the present. Likewise, the development of tangible human cooperation in solving complex long-term problems can be a very challenging and interesting learning experience.

Energy, Growth, and Survival

The formulation of utility functions is often treated as a purely subjective issue, with little objective, scientific content. Yet economists have identified two situations in which the proper value-weights for current decisions depend entirely on the objective opportunities, and not on the long-term value weights. These are called turnpike situations and crossroads situations. Both situations assume an overriding concern for the long-term future, as described above. Both situations imply, in theory, that different nations which share such a concern should be able to agree on global priorities on the basis of complete common interest, as they arrive at a more common, realistic understanding of the objective opportunities.

Most turnpike theorems apply to situations of balanced growth, where the same types of goods can be produced on an ever-increasing scale with no change in market shares. I will not discuss these further, because I consider them unrealistic for energy markets.

A crossroads situation exists when the long-term rate of economic growth is not fixed, but can be changed for millenia to come as the result of decisions in the more immediate future (Werbos, 1979). The literature on global futures suggests that we do indeed face three long-term possibilities:

- extinction of the human species
- stable equilibrium, permitting neither extinction nor sustained growth
- sustained growth for the indefinite future

If we could be certain that any one of these three will occur in the next few centuries, then we would not have a crossroads situation. The need for government programs would be marginal at best. If we do face a crossroads situation, however, the only long-term value of present government actions is their impact on the probabilities of the three situations. In other words, these probabilities define the appropriate utility function for the greater optimization problem discussed above.

The existing literature on global futures presents very diverse views of these probabilities. There is an optimistic literature, which argues that we can be certain of sustained growth, and that government actions are indeed unnecessary. There is a pessimistic literature, which argues that we can be certain that sustained economic growth is impossible. (The pessimists tend to

argue for government action to prevent an economic overshoot and collapse, which might well trigger nuclear conflicts and extinction if severe.) On the face of it, this combination of viewpoints suggests that no one is truly certain here; it suggests that we are facing the classic tendency of elite groups to understate their degree of uncertainty. (Howard Raiffa, in studies of subjective probability assessment, has shown how prevalent this tendency is.) This tendency is strongly reinforced when the best analysts and models are encouraged to defend the likelihood of "base case forecasts" or the realism of rigid future targets; such forecasts or targets, by definition, do not provide the range of uncertainty or the probabilities which are the proper focus of decisions.

A careful analysis of energy issues reinforces this picture. On the one hand, there is reason to doubt the initial claims by Meadows and Forrester, the classic pessimists of the 1970's, that resource limitations must force a halt to economic growth by 2050 or so. The "resource" quantity assumed in Meadows' model does correspond fairly well to current estimates of recoverable fossil fuel (Office of Technology Assessment, U.S. Congress, 1982); however, there clearly do exist other energy sources, such as nuclear energy (including breeder reactors and possibly fusion) which essentially wipe out the argument.

On the other hand, one cannot rule out the possibility of Meadows and Forrester becoming right after the fact, because (in my personal view):

o Recoverable fossil fuels are probably not much larger than they assume, despite the exaggerated claims of some optimists. For example, the amount of shale oil recoverable by known (though expensive) technology is on the order of 200 billion barrels, not enough to compete with coal.

o Nuclear technology might well go through a cycle of disuse followed by worldwide growth so rapid that deliberate explosions and unintended accidents could prevent further economic growth (or even cause extinction). This is arguably the most likely future if present trends continue, including a leaky system of international controls.

o Energy conservation may buy a few years worth of time in solving these problems. However, recent in-depth studies have shown that the past and potential of conservation have been exaggerated; conservation does not fundamentally change the long-term energy balance problem. Three-fourths of the energy use in the U.S. is in industry and transportation, for which careful analysis shows very little elasticity of demand (circa 0.2), and a persistent connection between energy growth and economic activity. (Efficiency in transportation could be doubled, but Meadows and Forrester accounted for the possibility of doubled efficiency across all sectors.) The limits are even more severe for small-scale technologies like wind, biomass, and conventional geothermal power.

o Large-scale energy probably can be provided at costs low enough to support rapid economic growth by a combination of earth-based solar power (for peak-load power in some regions) and space-based solar power. While the technology here looks very promising, the research and the required infrastructure have not been moving as quickly as they might; given the long lead time in development, it is possible that this technology will not become available on a large enough scale until after Forrester's dire predictions have already occurred.

o Meadows and Forrester did not analyze the more near-term problems of oil and natural gas supply. These problems, while solvable, involve similar long lead-times in developing new technology, with similar possibilities for trouble. In fact, oil markets could lead to conflicts of interest in the MidEast as early as the year 2000, conflicts so serious that they may affect the probability of human extinction.

o There is no magic bullet in free market economics or in Communist planning that can solve these problems automatically. Whatever system we use to organize our thinking and research, we have to think seriously about the substantive problems and be sure that the essential research does not "fall between the cracks" in order to minimize the probability of serious problems. Optimists often cite all the gloomy predictions of the past hundred years which did not come true; however, if they would look back thousands of years, instead of decades, and reconsider the work of Toynbee and Spengler, they would see how the pessimists usually did turn out to be right in time, in all the great civilizations of the world. Toynbee also points to possibilities for renewed vigor in a civilization, if the challenge of the times is fully faced up to.

The purpose of this discussion is not to argue that the pessimists may be right. Rather, it is to point towards those pressure points in the global energy system which will have the greatest impact on the probability of sustained growth in the future. A more rigorous approach to these issues might involve a formal global model extending to the year 2100, accounting for a wide variety of assumptions and scenarios, and calculating how the probabilities change with respect to each of the key assumptions and policies. These sensitivities, in turn, could provide the utility weights or goals for near-term policy decisions.

In general, one would expect this approach to recommend a high priority for accelerating the profitable development of new technologies such as low-cost solar cells, space-based solar power, and vehicles powered by fuel cells consuming methanol. (These technologies now seem quite realistic, from a technical point of view, because of recent progress; however, few practitioners in the industries affected believe that the nontechnical barriers will be overcome in a timely manner.) One would expect policies to overcome chicken-and-egg problems in the deployment of these technologies. One would expect a nuclear policy with three priorities, all supported at a very high level of effort: (a) to minimize nuclear deployment in most of the world for years to come (in part by subsidizing coal development); (b) to maintain nuclear development in highly secure areas, with international cooperation, enough to provide a test-bed for new technologies developed to reduce safety and security problems; (c) to maximize the degree of international control over the dissemination of the technology. Some attention would also be paid for those few fossil technologies -- like primordial natural gas or exotic forms of geothermal-- which offer some hope, however slim, of an energy source large enough to change the global energy equation.

Any other government actions involving energy should be driven by the need to maintain the economic infrastructure during the transition period. Ordinary market economics (with the usual consideration of environmental and information externalities) should be a sufficient guide for such actions. However, if so many resources are put into the infrastructure that nothing is left to get us through the transition, we will be in deep trouble. The criterion to use in making this tradeoff should be the probability of a successful transition, when all factors are accounted for. This probably requires a small, balanced sacrifice of short-term growth in order to increase the chance of long-term growth.

Comparison With the Existing Use of Models

Current energy policy does not provide a detailed, mathematical translation of all these long-term goals down to the actions required to advance them. There are many reasons for this, including:

o The difficulty of understanding and quantifying key causal relations, such as the impact of greater use of nuclear power in the Third World on the probability of a precedent-setting explosion, and the further impact of such a precedent. Analysts and contractors have a strong tendency to revisit the questions they best know how to answer, rather than questions which are difficult but important and which require a deep

reconsideration of their views. This seems to apply most of the time, even when the important questions are stated in a painfully explicit manner in a statement of work.

o The natural human tendency to react rather than act, the tendency to focus energy on current sources of pain and pleasure (such as this week's oil price). It is also natural for humans to overcome this tendency, but only through a difficult process requiring persistence and self-awareness. We have a long way to go in this process. High levels of mutual trust within a society make it possible to advance from violent ways of resolving conflict, to democratic methods which seem impossible or "impractical" in less advanced societies. In a similar way, it is possible -- but not easy -- to overcome the common tendency to degrade analysis by treating it as a propaganda tool rather than a serious inquiry directed at common interests.

o The legitimate value of energy modelling for a wide range of uses other than long-range strategy formulation.

o The common overemphasis on base-case forecasting, which tends to limit the indepth analysis of critical factors which lead to alternative outcomes. Even when such analysis has been carried out, it is often left unpublished or relegated to obscure documentation reports, when the forecast itself is viewed as the main product. Many groups outside of EIA do not even publish their documentation.

Models are used for many purposes at DOE, at many levels. In an agency where many people have quantitative skills, the distinction between analysis, modelling and science is much less meaningful than one might imagine; all three involve similar problems of coherence or integration, overcoming bias, focusing on the right questions, and communicating what was learned. Similar mathematical modelling techniques are used at DOE for:

o Describing energy markets

o Describing the dynamics of the atmosphere, to predict environmental impacts

o Describing nuclear reactors, for use in safety analysis

o Et cetera

There are two major sections of DOE which produce integrated forecasts of energy supply and demand through their own efforts -- the Energy Information Administration (EIA) and the Policy Office. In addition, the Carbon Dioxide Research office in the Office of Energy Research has sponsored the development of a tiny global energy model at the Oak Ridge National Laboratory Institute for Energy Analysis, under Edmonds and Reilly. National Acid Precipitation Assessment Program (concentrated in the Fossil Energy section) has sponsored the development of a large, integrated system of energy demand models at Argonne National Laboratory, as part of a system to predict air pollution. The Conservation Office has sponsored the development of process models of energy demand, to help them evaluate conservation options. The Energy Emergencies group once supported work with EIA on an Energy Disaggregated Input-Output model, and in-depth demand analysis focused on the issue of vulnerability to oil supply shortfalls. Lawrence Livermore Laboratories maintained a well-documented long-term integrated model for their own use, similar to EIA's old LEAP model, until a few years ago.

Figure 1 provides an overview of the integrated (multi-fuel) models used at EIA and the Policy Office, as of a year or two ago.

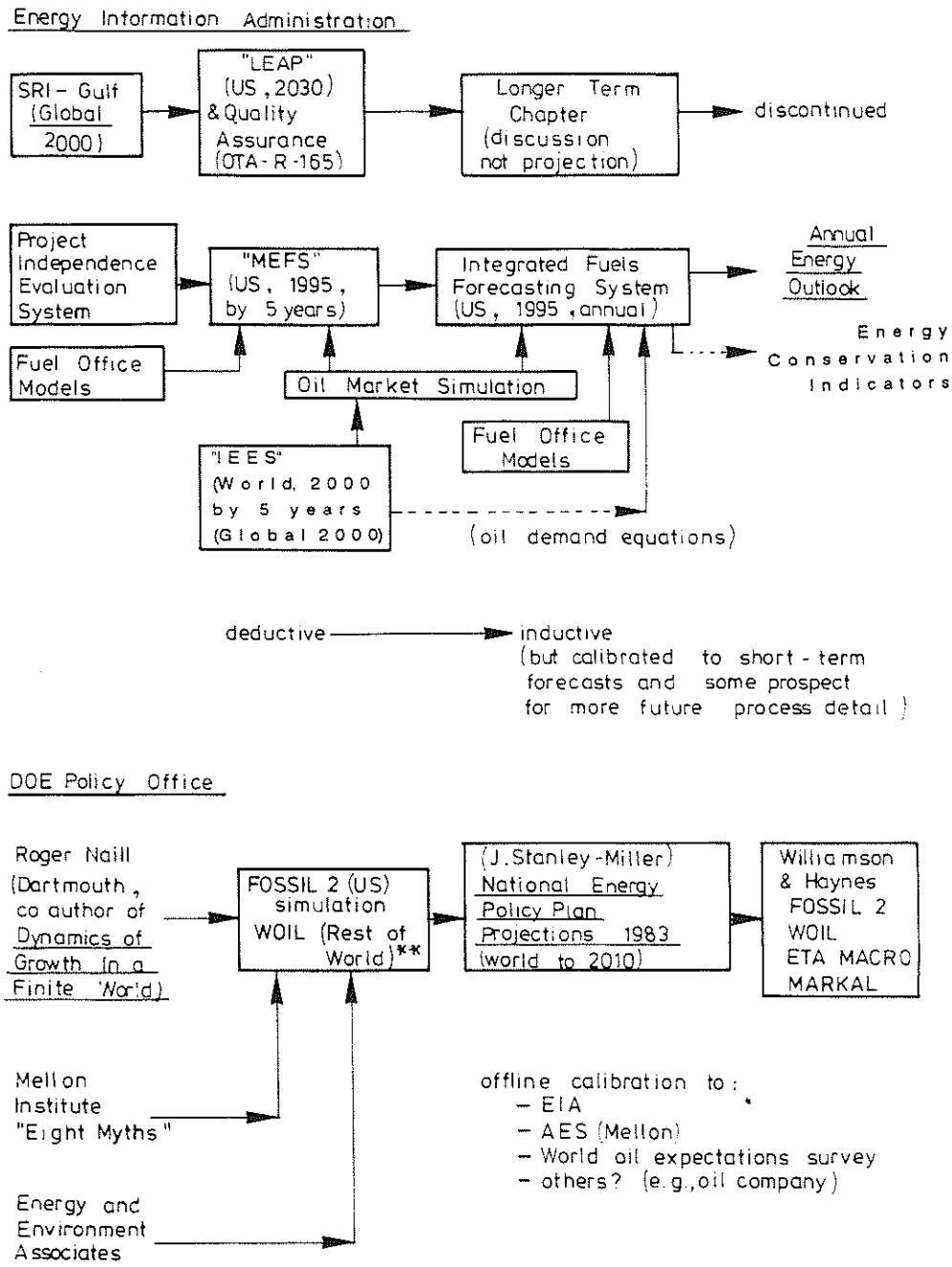


Figure 1. Roadmap of Models, Past and Present

EIA Models. Under the chart for EIA, we stress that EIA has moved from an emphasis on "deductive" models to an emphasis on "inductive" models. The early models, such as the Project Independence Evaluation System, relied heavily on linear programs and large-scale engineering models. They were motivated by an effort to understand the oil dependency problem of the U.S. in the 1970's. They were large-scale models, relatively complex, and difficult to calibrate against historical reality. The Integrated Fuels Forecasting System (IFFS) was developed to provide annual forecasts, more rooted in historical data and econometric methods. For a few years, the document Energy Conservation Indicators was an important byproduct of the analysis process;

it plotted those historical indicators which turned out to be most revealing, based on experience in trying to quantify cause-and-effect relations via econometric modelling.

At present, EIA is planning to carry this trend still further, through a new modelling system being developed for the Annual Energy Outlook 1987. Documentation should be available from the National Energy Information Center by the time this book is published.

The EIA modelling effort still involves direct participation by dozens of full-time modellers, across several Divisions of EIA. By contrast, the Policy Office currently relies on a small in-house staff and on contractors for its integrated modelling. This staff was moved to the Office of Oil and Gas in the Policy Office in September, 1978. Most of the modelling efforts within the Policy Effort focus on natural gas deregulation or, to a lesser extent, other fuel-specific issues. A much larger effort went into the initial development of the two integrated models -- WOIL and FOSSIL2 -- which are standard Systems Dynamics models, based on the methodology developed by Jay Forrester (Meadows, 1972, 1974). For a time, there was an effort to compare the results of WOIL with those of ETA-MACRO (Stanford) and MARKAL (Brookhaven), as part of the updating process, but that work has since been dropped.

The recent Energy Security Report to the President (Office of the Secretary of Energy, 1987) relied on EIA and its IFFS system rather than WOIL or FOSSIL2 for its projections. Nevertheless, the Policy Office did sponsor an analysis of energy demand in the Third World (at Lawrence Livermore Laboratory) which compares quite well with any other work on that subject. The analysis was mainly based on econometrics, but it took full account of detailed data on concrete forces at work in the Third World. Many decisions have yet to be made about the future, the goals, and the viability of these modelling efforts.

Other Integrated Models. Beyond DOE, Senator Gore and Congressman Gingrich have introduced a bill to mandate integrated modeling across all sectors -- energy, food, population, and environment, etc. -- on a government-wide basis. Critics of the bill have argued that the results of such efforts will be worthless base-case forecasts, used mainly as propaganda, based on very shallow, obscure, ivory-tower assumptions. Proponents of the bill have argued that it contains provisions for second-guessing by Congress and others (much as the Congressional Budget Office second-guesses the forecasts of the President's Council of Economic Advisers and forces attention on key assumptions), and for participation by substantive experts and policy-makers in the various Agencies. Depending on the vision and corporate style of the program managers, a global modeling project could be either a powerful stimulus to long-range thinking and debate or an oppressive deterrent to critical re-evaluations and creative imagination. At present, Congress is paying more attention to other issues; the bill is going nowhere because of sheer indifference.

The Department of Defense has also tried to develop integrated models from time to time. Some of the earlier efforts, relying heavily on Washington Beltway contractors, had very little empirical value in predicting international conflict or civil wars (Werbos, 1978), the variables which relate most directly to the immediate missions of the Department of Defense. More recent efforts, being upgraded by Meadows for the Strategy Division of the Joint Chiefs of Staff/J-5, incorporate submodels of all sectors, based on systems dynamics. There will soon be an effort to establish a connection from the socioeconomic sectors to the forecasting of conflict, perhaps by grafting on models developed by Choucri and North or Bremer (GLOBUS), using an integrating system such as ARISTOTLE developed by Mesarovich. The Department of Defense, unlike the Energy Information Administration, has a clearcut jurisdiction in trying to predict nuclear conflict (whether in war, civil war, or terrorism) as a function of nuclear fuel cycles and other global conditions; however, the technical difficulties of understanding that link are quite intimidating, and perhaps beyond the scope of present research budgets and approaches now used by political scientists.

REFERENCES

Barney, G., Ed. (1980). The Global 2000 Report to the President: The Technical Report. Washington D.C.: GPO Stock No. 1980 O-256-752.

Brown R., et al. (1974). Decision Analysis: An Overview. Holt, Rinehart and Winston.

Firestone, J. (1979). "Project Evaluation and Value Interpretation", International Journal of Urban Systems, p.141-152.

Meadows, D., et al. (1972). The Limits to Growth. Washington D.C.: Potomac Associates.

Meadows D., et al. (1974). The Dynamics of Growth in a Finite World. Cambridge, Mass.: Wright-Allen Press.

Raiffa, H. (1968). Decision Analysis: Introductory Lectures on Making Choices Under Uncertainty. Reading, Mass.: Addison-Wesley.

Schelling, T.C. (1960). Strategy of Conflict. Cambridge, Mass.: Harvard U. Press.

Werbos, P. (1987). "Building and Understanding Adaptive Systems: A Statistical/Numerical Approach to Factory Automation and Brain Research", IEEE Transactions on Systems, Man and Cybernetics.

Werbos, P. (1987). "EMF 8: A Modeller's Perspective", in The Changing World Energy Economy (Proceedings of the 1986 North American Conference of the International Association of Energy Economics), David Wood, Ed. Cambridge, Mass.: MIT Center for Energy Policy Research.

Werbos, P. (1987). "Oil Dependency and the Potential for Fuel Cell Vehicles", SAE Paper 871091. Warrendale, Pa.: Society for Automotive Engineers.

Werbos, P. (1987). "Learning How the World Works: Specifications For Predictive Networks in Robots and Brains", in Proceedings of the Annual Conference on Systems Man and Cyberneticss, Andrew P. Sage, Ed. New York: IEEE.

Werbos, P. (1986). "Generalized Information Requirements of Intelligent Decision-Making Systems", in SUGI 11 Proceedings. Cary, N.C.: SAS Institute. (A revised version, available from the author, is easier to read and contains more discussion of psychology.)

Werbos, P. (1986). "Documentation of the Transportation Energy Demand (TED) Model", DOE/EIA-M013. Washington D.C: National Energy Information Center (NEIC), U.S. Department of Energy (202-586-8800).

Werbos, P. (1979). "Changes in Global Policy Analysis Procedures Suggested by New Methods of Optimization",Policy Analysis and Information Systems. Vol. 3, No. 1.

Werbos P., and J. Titus. (1978). "An Empirical Test of New Forecasting Methods Derived From a Theory of Intelligence: The Prediction of Conflict in Latin America", IEEE Transactions on Systems, Man and Cybernetics.

Handbook of Decision Analysis. (1978) Tyson's Corner, Va.: Decisions and Designs, Inc.,.

Office of Policy, Planning and Analysis. (1983). DOE, Energy Projections to the Year 2010. DOE/PE-0029/2.

Trends in Multifactor Productivity. (1983) Chapter 4. Bulletin 2178. U.S. Bureau of Labor Statistics, (GPO stock no. 029-001-02758-3).

Office of Technology Assessment. (1982). Global Models, World Futures and Public Policy: A Critique, Appendix C. Wasington D.C.: Office of Technology Assessment, U.S. Congress.

Fuel Cell News. (1987). "Interest in Membrane Fuel Cells Rebounds", Washington D.C. 20006: Industrial Fuel Cell Association, Suite 200, 1627 K St. N.W.

Space Power, various issues.

Annual Energy Outlook. DOE/EIA-0383. Washington D.C.: NEIC, recurrent.

Energy Conservation Indicators. DOE/EIA-0441. Washington D.C.: NEIC, recurrent.

Eight Myths..., Mellon Institute (Applied Energy Services), Virginia.

Office of the Secretary of Energy. (1987). Energy Security: A Report to the President of the United States. DOE/S-0057. Washington D.C.: DOE.

Energy Vol. 15, No. 3/4, pp. 187–201. 1990
Printed in Great Britain

0360-5442/90 $3.00 + 0.00
Pergamon Press plc

1.6. APPLICATION OF MODELS IN THE PROCESS OF LEGISLATION[+]

JOHN P. WEYANT

Energy Modeling Forum, Department of Engineering-Economic Systems, Stanford University
Stanford, California
U.S.A.

Abstract - To what extent have quantitative models been used in the formulation of policy? The author addresses this question by examining two issues confronted by the Ninety-fourth Congress, oil price decontrol and synthetic fuels commercialization. He describes the influence of results from the large-scale models used to analyze the proposed policies and identifies conditions that apparently accounted for the successful use of these models.

1. INTRODUCTION

A number of investigators have attempted to assess the influence of formal modeling on the policy process.[1] Fromm, Hamilton, and Hamilton (1974), who surveyed 316 key persons involved in federally funded modeling projects, drew a typical conclusion: "It would appear that at least one-third and perhaps as many as two-thirds of the models failed to achieve their avowed purposes in the form of direct application to policy problems." The recent proliferation of policy models requires a more up-to-date assessment of their utility, however.

The analysis of both oil price decontrol and synthetic fuels commercialization involved the extensive use of computerized mathematical models. I describe how and where the results of the models were used in debates about these issues and try to infer where and why the models may have influenced political decision making. I attempt no absolute evaluation of the models employed. Decision makers may use the results of model because they or their staffs believe that the models is valid or because those results support their own positions on a particular issue.

The process by which the results of analytic studies influence policy is complex and multifaceted-and the influence can be subtle. For example, suppose a complete analysis is done of two policy options, A and B. The analysis leads to a bill in the U.S. House of Representatives favoring A, which subsequently becomes law. The vote for A over B may have been independent

---

[+]This paper is reprinted from Policy Analysis 6 (1980) pp. 211-234 with the permission of the regents of the University of California. It was originally titled "Quantitative Models in Energy Policy." The author is grateful to the National Science Foundation and the Electric Power Research Institute for support and to Martin Greenberger, William Hogan, Ron Howard, Alan Manne, and Walter Schroeder III for numerous helpful suggestions

[1]Survey approaches include Cline (1961), Shubik and Brewer (1972), Fromm, Hamilton and Hamilton (1974), Naylor and Shauland (1976), GAO (1978) and HRI (1976). Case study approaches include Brewer (1973), Greenberger, Crenson and Crissey (1976), GAO, (1976) and Mar (1979).

of the analysis, but Congress could perhaps have been led away from an option even better than A or B simply because option C was never subjected to an "in-depth" analysis. It is thus important to try to infer not only the direct influence the results of a model may have on the way a member of Congress votes on a particular issue, but also how the results of the model influence anything that could identifiably affect the resulting legislation either directly or indirectly.

I therefore established a general framework for the analysis of each of the case studies discussed here that should allow most of the potential influence of model results on the resultant public laws to be identified (see Figure 1). Within this framework, I examine the use of models in the formulation of the planning document of the relevant executive branch agency, the administration's position, and the bills initiated by Congress.

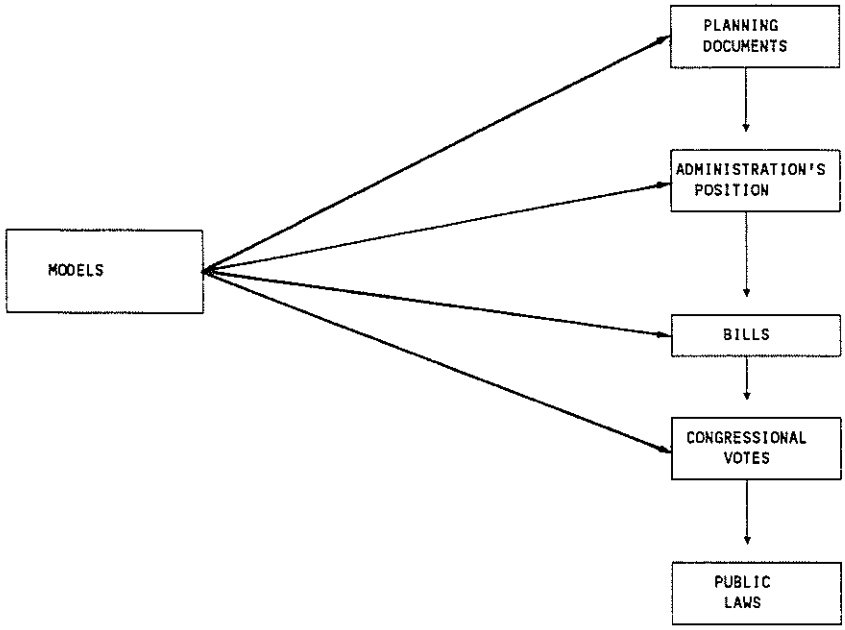


Figure 1. The results from formal models can influence the public policymaking process at several levels.

2. OIL PRICE DECONTROL

In August 1973, to minimize the inflationary impact of world-wide oil price increases, the Cost of Living Council devised a system of price controls for the petroleum industry: the price of domestic crude oil produced from a given property, when production was at or below the level of production from the same property in the same month of 1972 (termed old oil), was subject to a price ceiling equal to approximately the 15 May 1973 posted price plus $1.35. This resulted in a national average ceiling price of about $5.25 per barrel. All other crude oil produced (termed new oil) was excluded from the price controls. As an additional production incentive, each barrel of new oil released a barrel of old oil from the price controls. Thus a balance between the promotion of increased supplies and the moderation of prices was struck by adopting a "two tier" pricing system.

In response to a severe shortage of crude oil and its products in November 1973, the Emergency Petroleum Allocation Act (EPAA) of 1973 was enacted to distribute the remaining

products equitably and at equatable prices. This act created on 26 December 1973 a Federal Energy Office with the authority to regulate crude oil prices. On 7 May 1974, Congress enacted the Federal Energy Administration Act of 1974, Public Law 93-275, which formed the Federal Energy Administration (FEA) as an agency separate from the Executive Office of the President. FEA was given the responsibility of regulating the price of crude oil according to the Cost of Living Council's scheme under the authority of the Emergency Petroleum Allocation Act. This authority was to expire as of 31 August 1975, however.

Soon after its creation, the FEA established the Project Independence Task Force, directed by Eric R. Zausner. This task force was charged with exploring and analyzing pathways by which the United States might become energy independent as rapidly as possible, and a major part of its work was the development of the Project Independence Evaluation System (PIES) model, a set of energy supply and demand models linked by and integrating model. This model was a regionalized model of the energy sector and formed the basis for much of the November 1974 final report of the task force.[2]

In late 1974 President Gerald R. Ford was formulating energy proposals to include in his 1975 State of the Union Message, and on 14 December 1974, a group of high-level administration energy advisers met at Camp David. Attendees at this meeting included Alan Greenspan (chairman of the Council of Economic Advisers), Frank Zarb (administrator of the FEA), Secretary of the Interior Rogers C.B. Morton, and Eric Zausner. Results of several scenarios run through the PIES system were presented to the group to aid it in its deliberations. Alternatives the group developed were laid before the president at Vail, Colorado, at a meeting during the Christmas holidays. As a result of these and other inputs, the president, in order to stimulate the domestic supply of - and to decrease the domestic demand for - crude oil, proposed complete decontrol of all crude oil prices in his State of the Union Message on 15 January 1975 (Ford, 1975).

Congressional jurisdiction for crude oil pricing rests primarily with two committees in each of the two bodies of Congress. In the House, the Ways and Means Committee has jurisdiction over tax legislation proposed in conjunction with any crude oil price decontrol plan, while the Subcommittee on Energy and Power of the Interstate and Foreign Commerce Committee handles all other aspects of any such decontrol proposal. On the Senate side, the Finance Committee handles the tax provisions, whereas the Committee on Interior and Insular Affairs is responsible for the other impacts of the proposed legislation.

Throughout the Ninety-fourth Congress the administration used two types of analysis to support its decontrol proposals. First, in order to project how much a particular decontrol plan would contribute towards energy independence, the FEA supplied oil prices, supplies, and demands from its PIES model. Then the projected changes in domestic crude oil supply, oil imports, and oil prices due to that decontrol program were input to the quarterly macroeconomic forecasting model of Data Resources, Inc. (DRI) to obtain projections of the program's short-term impacts on the economy (including GNP, unemployment, and inflation).[3]

The Role of the Models in the Congressional Policy Debate

The House Subcommittee on Energy and Power held hearings on oil price decontrol during February 1975 to evaluate the administration's proposal for the instantaneous decontrol of crude oil prices. Testimony by administration witnesses focused primarily on model results showing that

---

[2]For an interesting account of the work of the Project Independence Task Force see Tietenberg (1976). For a summary description of this modeling system see Hogan (1977). The final report is FEA (1974).

[3]For a general description of the DRI macro model see Eckstein, Green and Sinai (1974).

decontrol would significantly decrease this country's dependence on foreign oil, but that the effect would not even begin to be significant until 1985 because of delays in the response of both supply and demand to higher prices. Further, although the short-term macroeconomic impacts of decontrol over the ensuing ten quarters (the time period over which the DRI model forecasts) were not catastrophic, they were not insignificant. At that point instantaneous decontrol was projected to result in a $ 30 billion increase in energy costs in return for an approximately 16 million barrel per day (or 10 percent) reduction in oil imports in 1976 (U.S.House of Representatives, Committee on Interstate and Foreign Commerce, Subcommittee on Energy and Power, 1975).

In addition, the economy was in a recession at the time and the increase in energy costs would have been unfair to low-income groups in the absence of compensating tax legislation. Consequently Congress concluded that instantaneous decontrol would be too risky and instead proposed legislation to extend the EPAA from 31 August to 31 December 1975. This legislation (H.R. 4035) was vetoed by the president on 21 July 1975.

Observing that its proposal for instantaneous decontrol was not doing well in Congress, the administration on 17 July 1975 proposed gradual decontrol of domestic oil prices over a thirty-month period and submitted supporting analysis of the macroeconomic impact of the plan prepared by the FEA (1975). This time around, however, the staff of the Subcommittee on Energy and Power, particularly Walter Schroeder III, uncovered a technical flaw in the FEA analysis.[4] At the request of the members of the subcommittee, the staff had been evaluating the analytical work presented in support of the administration's various proposals. Rather than doing an indepth methodological review of the PIES system, a task clearly beyond its resources, the staff began by comparing the FEA's projections of the macroeconomic effects of sudden decontrol with those made by others (U.S., House of Representatives, Committee on Interstate and Foreign Commerce, 1975). The fact that the impacts the FEA had projected were uniformly less than those made by anyone else initiated a more in-depth look at the assumptions upon which the FEA forecasts were based. This uncovered a misperception on the part of the FEA about what these assumption actually were.[5]

The FEA had been comparing the macroeconomic projects for the thirty-month decontrol plan with a base case that it thought presumed continued controls as per the EPAA. However, interaction between the subcommittee staff (principally Schroeder) and FEA staff (primarily Alvin Cook in the Office of Economic Impact) uncovered the fact that the base configuration of the DRI model that the FEA was using actually assumed decontrol over a sixty-month period. This led to a considerable understatement of the macroeconomic impacts of the phased thirty-month decontrol plan proposed by the FEA. These results were described in a congressional committee print (U.S., House of Representatives, Committee on Interstate and Foreign Commerce, 1975). Table 1 shows the significant differences between the original FEA estimates and those calculated by the Subcommittee on Energy and Power using the correct base case. FEA officials blamed this error on the excessive work lead being placed on their staff during this period, while some observers blamed the fact that the DRI macroeconomic model was not itself developed at FEA,

---

[4]Schroeder was brought to the staff of the subcommittee in May 1975 from a three-year stint at the urban system lab at MIT. A primary reason for his being brought on board was to scrutinize the morass of data and analysis the administration was submitting in support of its energy proposals.

[5]The influence of congressional staffs on the work of the Congress appears to be both significant and growing. This phenomenon is illuminated by Malbin (1977) and Scully (1977). The influence of the congressional staffs is also discussed by Eckhardt and Black (1976), where the lack of a constitutional sanction for the staffs and alternatives to the current staff system are considered. Coincidentally, Robert Eckhardt, a Democratic member of Congress from Texas, is a member of the Subcommittee on Energy and Power.

as was the PIES model. Those who had input the sixty-month decontrol assumption at DRI would be more likely to have deleted it from the reference case used during the analysis of the president's plan.

Table 1

Estimates of the Economic Impacts of Oil Price Decontrol by the FEA and the House Subcommittee on Energy and Power Differed Significantly.

	FEA	Subcommittee
Unemployment	+0.20 (200,000 workers)	+ 0.80 (800,000 workers)
Real GNP	-.84%	- 2.86%
Wholesale prices	+2.40%	+ 6.13%
Consumer prices	+1.30%	+ 2.04%
Housing starts	-4.68%	-15.92%
Car sales	-3.31%	- 9.65%

Note: Impacts as of the fourth quarter of 1977.

Regardless of the reason for the mistake, many members of Congress considered the revised macroeconomic impacts of the proposed plan to be still too great a strain on an economy that was just beginning to show signs of recovery, especially in the absence of appropriate windfall tax legislation. Thus House Resolution (H.Res.) 605, disapproving President Ford's plan to decontrol domestic oil prices over a thirty-month period, was adopted by the House on 22 July 1975. H.Res. 641, disapproving a similar plan by President Ford calling for a thirty-month decontrol period, was adopted by the House on 30 July 1975. The following exchange between FEA Administrator Zarb and Energy and Power Subcommittee Chairman John Dingell on 28 July 1975 concerning the president's thirty-nine-month decontrol proposal was typical of the tone of the debate at that time:

> Mr. DINGELL. So it has a highly beneficial effect on the economy during the first few months. But the problem is that our econometric models largely go on the basis of 10 quarters. You have indicated in the accompanying papers a table of unemployment which you anticipate might transpire. But can you tell us for the four quarters of the last year in terms of unemployment and economics, what the impact would be, please?
>
> Mr. ZARB. Mr. Chairman, as you know, post-1977 is not included in our model. The issues that were raised concerning this effort are issues that related to the economic recovery of this Nation. If we have a recovering economy, and if fiscal and other policies are sound in the last 15 months of the program, then it is our conclusion, without the benefit of the computer model, that those impacts will be negligible.
>
> I point out, Mr. Chairman, a number of times that there is no free way to answer this question. This phasing seems to accommodate the economic cycle as it now is perceived.
>
> Mr. DINGELL. Essentially, what you are saying is that you are flying blind as to the last few quarters of the program. That brings me to the point that most immediately concerns me. I think I can tell you that I can support decontrol. I can even support this proposal as an alternative to sudden decontrol.
>
> But as I understand the President's comments during his remarks and also yours, you wish to have a total package, one which involves all parts. One of the President's proposals had been the creation of a windfall profits tax. Am I correct?

Table 2

Information Obtained from the FEA Model of the Short Presented to Participants in the House-Senate Conference During the Formulation of the Energy Policy and Conservation Act (EPA) of 1975.

Program	Imports in 1985 (Millions Barrels per Day	Macroeconomic Impacts[a] GNP (% Difference from Base)	Unemployment Rate (Difference from Base)	Consumer price Index (% Difference from Base
Base Case (Continued controls, $2 tariff, no cap)[b]	6.87	-	-	-
Instant decontrol ($2 tariff, no cap)	3.82	-4.00	+1.4	+2.5
39-month decontrol ($2 tariff, $10 cap)	4.45	+1.10	-0.3	-0.7
Instant decontrol (no tariff, no cap)	5.13	-2.30	+0.8	+1.5
Continued controls ($2 tariff, $8 cap)	12.54	+1.95	-0.6	-1.6

[a]In the fourth quarter of 1977.
[b]A tariff is a per barrel tax on imports. A cap is the maximum price for which a barrel of domestic crude can be sold.

The EPCA amended the EPAA to establish a forty-month program of continued price controls on domestic crude oil.[6] EPCA established as the bench mark for regulatory purposes a "weighted average" or "composite" price for crude oil. The composite was initially set at $7.66 per barrel. The president was authorized, within certain limits, to increase the composite price to account for inflation, and thereby maintain the composite price in real dollar terms, and to provide an incentive to increase production. The limitations on the president's authority to increase the composite price included a 3 percent annual limitation on the production incentive factor and a 10 percent overall limitation on combined annual increases based on the inflation adjustment and the production incentive factor.

This plan for gradually increasing the price ceiling on crude oil was not one of the model scenarios explicitly presented to the conferees who developed it, but those that were presented provided guidance on the tradeoff between increases in the rate of decontrol of crude oil and decreases in short-term economic growth and employment. Subsequently, the analysis performed during the formulation of EPCA was cited as a major influence on the oil pricing regulations it mandated in the House conference report on the Energy Conservation and Production Act (ECPA, 1976) which became public law 94-385 on 14 August 1976. ECPA removed the 3 percent limitation on the annual crude oil production incentive factor (the 10 percent overall limitation on the overall price increase remained). Most members of Congress believed that the economy was well enough into a recovery by mid-1976 to warrant an increase in the incentive factor, especially if the rate of inflation were low, as it would have to be for the overall price increase limitation to be observed.

[6]Note that although this regulation continued controls over forty months, it did not decontrol over that period. Thus, when the forty months were over, the disparity between the world and domestic prices for crude oil was a great as it was in late 1975.

The impact of the FEA's modeling efforts can be summarized in terms of the framework for analysis described in the previous section. In the words of Zausner "the PIES model was the plan" - that is, the Project Independence Report and the next FEA planning document, the National Energy Outlook (FEA, 1976), were in large part descriptions of the outputs of sets of scenarios that the FEA's policy analyst thought would provide the most insight for the energy policy debate.

The impact of the FEA's modeling on the president's decontrol program was largely negative. The administration naturally would have been anxious to claim that a decontrol plan would rapidly decrease this country's dependence on foreign oil with little adverse impact on the nation's ailing economy. The FEA's analysis showed, however, that the lead times in the response of domestic oil supply and demand would result in only a small decrease in imports by 1985 but with a discernible cost to the economy almost immediately. This tradeoff was not politically acceptable to the majority of members of Congress.

Close coordination between the FEA and the House Subcommittee on Energy and Power staff during the conference led to the enactment of EPCA. Here the FEA and DRI models were used to help the conferees pick what they believed to be an optimal rate of increase in the crude oil price ceiling - one that would encourage conservation and the development of domestic crude oil supplies without severe economic disruption.

How much the modeling work of FEA and the congressional staffs influenced the oil pricing regulations in EPCA and ECPA is not certain. At the very least it appears to have provided a useful accounting structure within which the regulations versus deregulation debate could be conducted, and to have helped educate members of Congress, particularly those on the relevant committees, as to the important effects of each possible pricing policy. In addition, many cases were run during the conference on the EPCA, with extensive interaction between the modelers and the conferees. New proposals for decontrol were developed immediately, after the impacts of old proposals were projected. Although members of Congress disagreed about the precise numerical values for the import reductions and macroeconomic impacts of each alternative, the modeling work helped to identify the dimensions of the debate and to put reasonable bounds on the relevant parameters. Thus the modeling results provided a framework for the development of the decontrol legislation.

3. SYNTHETIC FUELS COMMERCIALIZATION

Another topic discussed during President Ford's energy policy development meetings in December 1974 was the possibility of developing/commercializing synthetic fuels (synfuels) to replace oil imports and dwindling domestic oil and gas supplies.[7] As in the case of decontrol, the primary motivation for this proposal was import reduction. Up to this point, the synfuels issue had not received nearly as much attention as the oil pricing issue. Ford's advisers stressed the need to propose a specific goal for synfuels production in the not-too-distant future as the payoff for what would have to be a relatively large level of federal support. Largely at the suggestion of representatives form the U.S. Office of Management and Budget (OMB), the president established, in his 1975 State of the Union message, a national goal of one million barrels of synfuels capacity on line by 1985.[8] Unlike decontrol, the synfuels proposal was distinctly a Ford proposal, not a carry-over of the Nixon proposal. The specific goal of a million barrels a day was probably due to the desire of Ford and his advisers to propose something new or different,

[7]The synthetic fuels referred to here are liquid, high-Btu gas, and low-Btu gas (all derived from coal), oil shale, and methane derived from biomass.

[8]The OMB representatives included Frank Zarb, who had just left OMB to take over as administrator of the FEA, William McCormick, and James Walker.

rather than to a thorough study - with or without the use of formal models - of the potential limits on achieving that goal. This is where the synfuels case study begin.

Ironically, the OMB representatives who in early January had originally suggested the on-million-barrel-a-day goal for synfuels production were beginning to question the advisability of that goal by the end of January. Preliminary analysis suggested that the current generation of synfuels technologies would have to be commercialized in order to meet the goal, and, since these technologies were unlikely to be economic any time before 1985, a large government subsidy would be required to meet the goal. On the basis of these preliminary calculations, OMB proposed a large-scale study of the possibility of a synthetic fuels program. The directive for this study came from the president's Energy Resources Council and the analysis was to be performed by a large interagency task force managed by OMB. Since OMB's primary function is to oversee all aspects of the federal budget, its role as lead agency in a large-scale study of a potential budget item was an unusual occurrence. The task force officially began its work on 19 February 1975.

Shortly after its inception the task force decided that it should analyze not only the president's million-barrel-a-day (the "medium program'), but three alternative goals for that year as well: zero barrels (no program), a 350,000-barrel-per-day "information-level program," and 1,700,000-barrel-per-day "maximum program' representing an all-out effort. During March the policy group in the task force proposed that a large-scale energy model be used as an accounting framework for the group's analysis. An accounting framework was required to ensure that the independent projections of the various task force study groups satisfied certain basic consistency conditions (such as mass balance conditions) for the energy system as a whole. An analysis of alternative energy models by the National Science Foundation's Office of Energy R&D Policy identified the SRI/Gulf model as being most appropriate for use in studying the synfuels commercialization issue (see Rattien and Bell, 1975). The model had been developed by Edward G. Cazalet of Stanford Research Institute (SRI) under contract to Gulf Oil for the explicit purpose of analyzing potential investments in synthetic fuels. The fact that the SRI model is regionalized and that it includes explicit consideration of the industries that would have to supply equipment to a synfuels industry made it well suited for the analysis of synfuels commercialization.[9] Furthermore, it was as comprehensive and well known as the two major energy models in use in Washington: the FEA's PIES model and the Brookhaven models being developed through the U.S. Energy Research and Development Administration (ERDA).

The Synfuels Task Force relied on the SRI/Gulf model for its deterministic benefit-cost calculations. The model is structured around an energy network containing various energy efficiencies, resources available at particular prices, and capacity constraints associated with specific energy technologies. The SRI/Gulf model contains considerable regional detail. It reports en-use data for nine census regions and output data for twenty supply regions in the United States. Supplies and demands are balanced simultaneously at each node at each point in time over the fifty-year planning horizon.

Initially, the SRI model was to be used only for sensitivity analysis, since some of the scenarios it had run for Gulf Oil were to be redone using inputs from the task force and its expert consultants.[10] It is also noteworthy that, because of the proprietary character of models developed by private corporations, several members of the task force were skeptical about using a model not developed by the government. Models developed by private corporations tend to be less accessible than models developed within the federal government. In general, this can slow the dissemination of the model's basic concepts, as well as increase the time it takes to run a new

[9]For a general description of the SRI/Gulf model, see SRI (1976) or Cazalet (1977).

[10]The results of these sensitivity analyses are summarized in SRI (1975).

scenario. (Use of privately developed models may also results in errors of the sort that occurred when the FEA used the DRI macroeconomic model in the case of price decontrol.)

As is often the case when key inputs to a model are varied systematically to assess the sensitivity of its projections, the analyst is left with a large number of alternative scenarios, each with different implications for the advisability of the program being studied. It turns out, however, that the group that constructed the SRI energy model specializes in a methodology known as decision analysis, which seeks to combine uncertain information into a single-value measure by explicitly specifying probability distributions for key parameters.[11] The assumed timing for the resolution of the uncertainties is another central focus of the techniques of decision analysis. Almost as soon as SRI became involved in the work of the task force, Cazalet, the SRI project leader, proposed that his group do a decision analysis of the synfuels commercialization issue using inputs (including relevant probabilities) from the task force. At the same time, James Walker of OMB, the leader of the policy analysis group in the task force, was also interested in seeing a decision analysis performed on the synfuels commercialization issue (this was manifest in the OMB white paper on the Synfuels Commercialization Program; Walker, 1975). By early April 1975 the two proponents of decision analysis had convinced William McCormick of OMB, the director of the task force, to go ahead with the decision analysis proposal.[12] The decision analysis was formulated by Ronald Howard of Stanford University and implemented by Cazalet, Steve Thani, Dale Nesbitt, and Warner North of SRI. Under extreme time pressure, with and extraordinary amount of interaction between the SRI analysts and the members of the task force, the synfuels commercialization decision analysis was performed in six weeks, in time to be included in the June 1975 draft report of the group.[13]

After considerable interagency debate, the task force agreed to a nominal set of judgmental probabilities.[14] Based on these estimates, certain measures of performance (such as consumers' surplus, producers' surplus, expected embargo loss, and socioeconomic costs) were calculated for each alternative program level. The values of these measures of performance were then added in order to provide an aggregate measure of performance that was used to compare alternative program goals. The expected discounted net benefits of the maximum, medium, and information-level program all turned out to be negative. Table 3 shows the expected benefits for the four program levels.

Table 3
SRI Decision Analysis Estimated that all Three Non-zero-Goal Synfuels Programs Would Have Negative Net Benefits

	Expected Discounted Net Benefit (billions of 1975 dollars)				
Program Alternative	Consumer Surplus	Producer Surplus	Embargo Protection	Environmental and Socioeconomic	Total
No program	0	0	0	0	0
Information Program (.35mm bbl/day)	1.07	- 2.71	0.43	0.44	- 1.65
Medium program (1 mm bbl/day)	3.29	- 3.74	1.13	-1.14	- 5.41
Maximum program (1.7 mm bbl/day	4.55	-15.77	2.23	-1.99	-10.93

[11]In fact, at SRI the group was called the Decision Analysis Group.

[12]McCormick later became ERDA's advocate for synfuels.

[13]The initial account of this work is reported in SRI (1975).

[14]The American Petroleum Institute later criticized the task force for using synfuels cost estimates that were much too low. See Peterson (1977).

The role of the Models in the Congressional Policy Debate

Given that all three of the non-zero-goal programs were estimated to have negative net benefits, a synfuels commercialization program was not worth pursuing. This conclusion was in conflict with the president's publicly announced gal and was not accepted by many members of the task force, who pushed hard for the task force to favor the million-barrel-a-day program in its draft report. The analysts who performed the decision analysis, as well as its supporters within the task force, objected very strongly to this recommendation. Those who were in favor of the million-barrel-a-day goal believed that the modelers were taking the results of their study too literally, something they were able to argue much more effectively by changing the task force's recommendation to a 350,000-barrel-a-day information program. That is to say, they replaced the goal of production per se with the goal of acquiring better information of synfuels technologies that could be used to make better decisions in the future, despite the fact that the decision analysis had explicitly considered the information value of a "small" program to be outweighed by its cost. The recommendation of the task force did not directly conflict with the president's previously announced goal. It was formulated in terms of a 350,000-barrel-a-day information program initiated immediately, with the option in 1979 of increasing the goal to pursue given information then available.[15]

The development of a synthetic fuels commercialization program became an issue for Congress in the fiscal year 1976 ERDA authorization bill H.R. 3474. After receiving the bill from the House, the Senate added an amendment (Section 193) providing $6 billion in loan guarantees for commercial development of solar energy, geothermal sources, and synthetic fuels. Of that amount, $2.5 billion was specifically set aside for high-Btu gas from coal. Although the Synfuels Interagency Task Force had included other incentives in the alternative programs it had considered, loan guarantees were the key ingredient, and the $6 billion figure was about enough to cover the 350,000-barrels-a-day program the task force had recommended. Thus, although the bill included provisions for the financing of geothermal and solar technologies, synfuels proponents argued that most of the $6 billion in loan guarantees would go towards synthetic fuels, perhaps because an already completed analysis of that particular program level was available.

The House Subcommittee on Energy Research Development and Demonstration (Fossil Fuels) held ten days of hearings on Section 103 during September and October of 1975. The work of the Synfuels Interagency Task Force was prominent in these hearings. The hearings also revealed numerous issues that had not been addressed in either the task force report or Section 103. No provision had been made for cooperation with state and local officials. Coordination with federal, state, and local environmental regulations remained unspecified. No account was taken or provision made for social and economic impacts on local communities. There was no mechanism for congressional review; ownership of patents under such an arrangement was unclear. Furthermore, attachment of such a loan guarantee to an authorization bill seemed to many House members an attempt to circumvent normal budgetary process. Additionally, in order to expedite the commercialization program, ERDA chose to support Section 103 in the amended House version of its authorization. Since it was only part of the recommended incentives program, and since the draft version of the incentives program in Section 103 did not take social, economic, and environmental issues directly into account, ERDA's position was attacked as being piecemeal and lacking comprehensiveness.

A report on these and other findings was made to the House Committee on Science and Technology on 30 October 1975. Thereafter, the House conferees met and decided to go to

[15]This recommendation is included in the final draft of the report submitted by the Synfuels Interagency Task Force to the president's Energy Resources Council (SITF, 1975).

conference with the Senate and to try to work out a compromise on Section 103 of the FY 1976 ERDA authorization.

After extensive meetings by the conferees, compromise language was adopted that addressed all of the issues raised by the subcommittee. When the ERDA authorization bill was taken to the floor, however, the House delated Section 103, on 11 December 1975 by a vote of 263 to 140. It is of interest that during the hearings on the synfuels issue several congressional staffers had questioned the SRI analysts directly and learned that no program was the optimal strategy according to their decision analysis.

Soon after the beginning of its second session, two new loan guarantee bills were placed before the Ninety-fourth Congress. S. 2869, the Synthetic Fuels Act of 1976, was introduced in the Senate on 23 January 1976. This bill essentially recapitulated the amended version of Section 103, deleting only loan guarantees for the development of oil shale.[16] H.R. 12112 was introduced in the House on 25 February 1976. The major change this bill initially promoted was authorizing $2 billion in loan guarantees for the full range of technologies as a supplement to the FY 1976 authorization. It also set a time limitation beyond which no more loan guarantees could be furnished.

After extended hearings by three House committees (Banking, Currency, and Housing; Interstate and Foreign Commerce; and Ways and Means), a compromise bill was finally submitted in the form of a substitute in which an agreement was reached on funding. The compromise bill allowed $3.5 billion in loan guarantees and direct loans and $500 million in price supports. The bill was killed, on a procedural vote rather than on substantive issues, when it was brought to the floor on 23 September 1976 - losing by a single vote, 193 to 192.[17] The close vote is indicative of the doubts that remain over still unresolved economic and environmental issues, as well as the more basic issue of the degree to which the federal government should intervene to promote commercialization of a program intended to rely primarily on private investment and development. This issue was especially important in the case of synfuels development because of the public's distrust of the oil industry.

In summary, it is clear that the SRI decision analysis, which involved extensive use of the SRI/Gulf energy model, played a significant role in the report of the Synfuels Interagency Task Force to the president's Energy Resources Council. This report led directly to a reduction in the presidential goal for synthetic fuels from one million barrels a day by 1985 to 350,000 barrels a day by 1985. The $6 billion in loan guarantees requested in the proposed amendment (assuming most of it was targeted for synthetic fuels commercialization) to the FY 1976 ERDA authorization bill not coincidentally was just enough to cover that important component of the 350,000-barrel-a-day program. The defeat of this proposal could have been due, in part, to the fact that the SRI decision analysis supporting the task force report had actually shown that no program was the best choice. Certainly, the defeat of the synfuels commercialization legislation in 1976 could have been brought about by this modeling result.

---

[16]The Congressional Budget Office issued a report in January 1976 suggesting that a program smaller than 350,000 barrels per day might provide benefits in excess of its costs may have influenced the proposals being developed at this point. See U.S., Congressional Budget Office (1976).

[17]A GAO report issued in August of 1976, which contended that alternatives to synfuels like conservation and solar were a more cost effective investment than synfuels, may have influenced the vote, although Bruce Peterson (1977) criticized it for not evaluations (GAO, 1976).

4. CONCLUSIONS

From the evidence presented here we can infer that the influence of the modeling in the oil pricing case and in the synfuels case was much greater than the influence formal modeling has typically had in the past. What can we conclude from these cases about the conditions that lead to the effective use of formal models in policy-making?

1. High-level Organizational Position of Model Builders and Advocates -- The high-level position of those who built, advocated, and used the quantitative models employed at the FEA was a major factor in the impact these efforts had in the oil pricing cases described above. Eric Zausner, who directed the Project Independence Task Force and gave most of the technical testimony on the PIES projections during the hearings on these issues, was promoted from assistant administrator to deputy administrator during 1975. His knowledge of the technical details of the FEA's quantitative studies was probably unprecedented for a government official at such a high level. His ability to coordinate, motivate, and promote the key members of the PIES modeling and policy analysis groups (William Hogan, Bruce Pasternak, James Sweeney, and David Nissen) was instrumental in obtaining policy-relevant results from the PIES system in a timely fashion. Similarly, the high-level position in the Synfuels Interagency Task Force of model advocate James Walker and the belief of the task force chairman, William McCormick, in quantitative models were key factors in the emergence of the SRI decision analysis as the centerpiece of the task force's final report.

2. Inclusion in the Analytical Framework Adopted of All Impacts Thought to Be Important by the Decision Makers -- The FEA's PIES system was designed to project the response of both supply and demand to higher oil and gas prices. Further, a short-term macroeconomic model was used to project the macroeconomic impacts of the price increases. These are the impacts that were at issue in the oil price decontrol case. How much economic growth had to be sacrificed in the short run to decrease our dependence on foreign oil and to decrease the magnitude of the natural gas shortage in the lorn run?

Similarly, in the synfuels commercialization case, the fact that the SRI/Gulf model is regionalized and that it includes explicit consideration of the industries that would have to supply equipment to a synfuels industry made the model ideally suited for the analysis of synfuels commercialization.

3. Compatibility between the Options and Criteria Considered in the Analysis and Those Used in the Policy Debate -- A problem with most analytical studies in the energy policy field to date has been their failure to consider the options involved in the policy debate and to employ the criteria actually used to evaluate those options. This is largely problem of timing; congressional debates unravel over a period of a few weeks or months, while analytical studies can take a year or more to complete.

The FEA's work in conjunction with the congressional staffs during the oil pricing case was unique in that scenarios were run literally overnight in response to the latest proposals. In fact, the runs that were made during the House-Senate conference on EPCA actually bracketed the crude oil pricing system that was eventually adopted.

In the synfuels case, some problems arose during the congressional debate because the work of the Synfuels Interagency Task Force had considered a package of incentives, including loan guarantees, price supports and subsidies, whereas the congressional version included only the loan guarantees. Additionally, the $6 billion in loan guarantees in the proposed bill were for solar and geothermal technologies, not just for synthetic fuels technologies. This made it hard for some members of Congress to relate the conclusions of the task force to the advisability of the proposed bill.

The successful application of formal modeling in these case studies should be heartening to analyst and policymaker alike. The model results were not always viewed as "substantive evidence," but they did enter into the policy debates. The efforts of many people and organizers will be required, however, to achieve the ultimate goal of generally improved communications between researchers and decision makers in the energy field. Only then can the full potential of models as aids to decision making be realized.

REFERENCES

Brewer, G.D. (1973). Politicians, Bureaucrats and the Consultant: A Critique of Urban Problem Solving, New York: Basic Books.

Cazalet, E.G. (1977). "Generalized Equilibrium Modeling: The Methodology of the SRI/Gulf Energy Model," mimeographed, Palo Alto, CA: Decision Focus Incorporated.

Cline, R. (1961). "A Survey of Mathematical and Simulation Models as Applied to Weapons System Evaluation," ASD - Technical Report 61-387, Washington, D.C.: U.S. Government Printing Office.

Eckhardt, R. and C.L. Black, Jr. (1976). The Tides of Power; Conversations on the American constitution, New Haven, Conn.: Yale University Press.

Eckstein, O., E.W. Green and A. Sinai (1974). "The Data Resources Model: Users, Structure and Analysis of the U.S.Economy," Economic Studies no. 15, mimeographed, Lexington, MA: Data Resources.

Ford, G.R. (1975). "The State of the Union," in Weekly Compilation of Presidential Documents, Federal Register Office, Volume II, No. 3, Washington, D.C.: Government Printing Office, p.49.

Fromm, G., W.L. Hamilton and D.E. Hamilton (1974). "Federally Supported Mathematical Models: Survey and Analysis," RANN Document Center, Washington D.C.: National Science Foundation.

General Accounting Office (1976). "An Evaluation of Proposed Federal Assistance for Financing Commercialization of Emerging Energy Technologies," A Report to the U.S. Congress by the Controller General of the United States, Washington, D.C.: GAO.

General Accounting Office (XXXX). "The Evaluation and Use of the Transfer Income Model-TRIM-to Analyze Welfare Programs", Report to the Congress, No.PAD-78-14, Washington, D.C.: GAO.

General Accounting Office (1978). "Ways to Improve Management of Federally Founded Computerized Models." Report to the Congress, No LCd-75-111, Washington, D.C.: GAO.

Greenberger, M., M.A. Crenson and B.L. Crissey (1976). Models in the Policy Process: Public Decision Making in the Computer Era, New York: Russell Sage Foundation.

Hogan, W.W. (1977). "Project Independence Evaluation System: Structure and Algorithms", in Mathematical Aspects of Production and Distribution of Energy, volume 21 of Proceedings of Symposia in Applied Mathematics, Providence, R.I.: American Mathematical Society.

Holcombe Research Institute (1976). Environmental Modeling and Decision Making: The United States Experience, New York: Praeger.

Malbin, M.J. (1977). "Congressional Committee Staffs: Who's in Charge Here?," The Public Interest 47.

Mar, B.W. (1989). "RANN Models of Regional Environmental Systems", Appendix B, Washington, D.C.: National Science Foundation, forthcoming.

Naylor, T.H. and H. Schauland (1976). "Experience with Corporate Simulation Models - A Survey", Long Range Planning.

Peterson, B.L. (1977). A Critique of Two Assessments of the Synfuels Commercialization Program, Washington, D.C.: American Petroleum Institute.

Rattien, S. and J. Bell (1975). "Critical Review of Synfuels Projections: Models and Methodologies", mimeographed, Washington, D.C.: National Science Foundation, Office of Energy R&D Policy.

Scully, M.A. (1977). "Reflections of a Senate Aid," The Public Interest 47.

Shubik, M. and G.D. Brewer (1972). "Models, Simulations, and Games - A Survey," Rand Report R-1060-ARPA/RC, Santa Monica, CA: Rand Corporation.

Stanford Research Institute (1975). "Synthetic Fuels Commercialization Program: SRI Energy Model Sensitivity Analysis Results," mimeographed, Menlo Park, CA.

Stanford Research Institute (1975). "Decision Analysis of the Synthetic Commercialization Program," mimeographed, Menlo Park, CA.

Stanford Research Institute (1976). "A western Regional Energy Development Study: Economics," mimeographed SRI project 4000, Menlo Park, CA.

Synfuels Interagency Task Force (1975). Recommendations for a Synthetic Fuels Program, report to the president's Energy Resources Council, stock no. 041-001-001112-3, Washington, D.C.: U.S. Government Printing Office.

Tietenberg, T.H. (1976). Energy Planning and Policy: The Political Economy of Project Independence, Lexington, MA: Lexington Books.

U.S. Congressional Budget Office (1976). Background Paper No. 3, Washington, D.C.

U.S. Federal Energy Administration (1974). Project Independence Report, stock no. 4118-00029, Washington, D.C.: Government Printing Office.

U.S. Federal Energy Administration, Office of Data and Analysis and Office of Economic Impact (1975). "Energy and Economic Impacts of Alternative Oil Pricing Policies," mimeographed, Washington, D.C..

U.S. Federal Energy Administration, Office of Economic Impact and Office of Analysis (1975). "The Economic Impact of the President's Proposed Program of Phased Decontrol", Washington, D.C.: FEA.

U.S. Federal Energy Administration (1976). National Energy Outlook, Washington, D.C.: Government Printing Office, Stock No. 041-018-00097-6.

U.S. Federal Energy Administration (1976). Project Independence Report, Washington, D.C.: Government Printing Office.

U.S. House of Representatives, Committee on Interstate and Foreign Commerce, Subcommittee on Energy and Power (1975). The Implications of the President's Proposals in the Energy Independence Act of 1975, 9th cong., 1st sess., 17, 18, 20, and 21 February 1975, Washington, D.C.: Government Printing Office.

U.S. House of Representatives, Committee on Interstate and Foreign Commerce (1975). "A Summary of Economic Analyses of Sudden Decontrol," 94th Cong., 1st sess.

U.S. House of Representatives, Committee on Interstate and Foreign Commerce (1975). Analysis of the President's July 17, 1975 Program to Decontrol Domestic Oil, 94th Cong., 1st sess.

U.S. House of Representatives, Committee on Interstate and Foreign Commerce, Subcommittee on Energy and Power (1975). The Potential Economic and Social Implications of the Sudden Decontrol of Domestic Old Oil, 94th Cong., 1st sess., 14, 17, 23 July and 8, 9 September 1975, Washington, D.C.: Government Printing Office, Serial 94-49, p.544.

U.S. House of Representatives, "Energy Conservation and Production Act (ECPA)" (1976). House Report No. 94-1392, 94th Cong., 2nd sess.

Walker, J. (1975). "Synfuels Commercialization Program," U.S. Office of Management and Budget.

Energy Vol. 15, No. 3/4, pp. 203–206, 1990
Printed in Great Britain

0360-5442/90 $3.00 + 0.00
Pergamon Press plc

1.7. POLICY MODELING: AN OVERVIEW+

JOHN P. WEYANT

Energy Modeling Forum, Department of Engineering-Economic Systems, Stanford University
Stanford, California
U.S.A.

Abstract - This article will explain how data analysis and modeling can be used in planning in the volatile environment in which the industry currently operates. The discussion will be quantitative and systematic in terms of trying to integrate the various components of the picture, i.e., OPEC supply and demand in developed and developing countries.

In this discussion we will use oil price forecasting as an interesting lens from which to view this process of using data analysis to analyze important issues. There are many diverse opinions regarding the oil price trend over the next decade. The modeling theory presented here has as a basic premise that much can be learned from comparing different projections.

Ideas on how industry can utilize analysis by using a scenario approach will be presented. In addition, this article will illustrate how to incorporate and quantify uncertainty in the planning process.

Inherent Uncertainty

Basically, data analysis and modeling are already useful to a certain extent in the industry. More, however, could be done. There should be more communication about the uncertainty inherent in individual forecasts and about the uncertainty that surrounds all forecasts.

Oftentimes, people in planning positions expect and demand more accuracy in forecasts than can be obtained. One has to understand that uncertainty is part of the world, and that it is an integral element in planning. Statisticians cannot be blamed for not always making accurate forecasts. There is a great deal of uncertainty inherent in the world and data analysis and modeling cannot be blamed for that uncertainty.

Diversity of Opinions

Since 1981 I have been involved in a project called "The International Energy Workshop." Anyone who has projections about the world energy scene can submit a forecast and revise it every six months. There are between 60 and 80 different groups participating.

In the Workshop projection of oil prices (Figure 1), the real price of oil in 1980 has an index value of 100. Other prices in real terms are relative to that level. For calibration

---

+Reprinted with permission from Petroleum Management (7), pp. 35-37, were it was entitled "Using Data Analysis and Modeling in Planning."

purposes, the 1980 price of oil in 1984 dollars is about $42. If the 1984 index is converted to 1984 dollars, the value goes to $28.

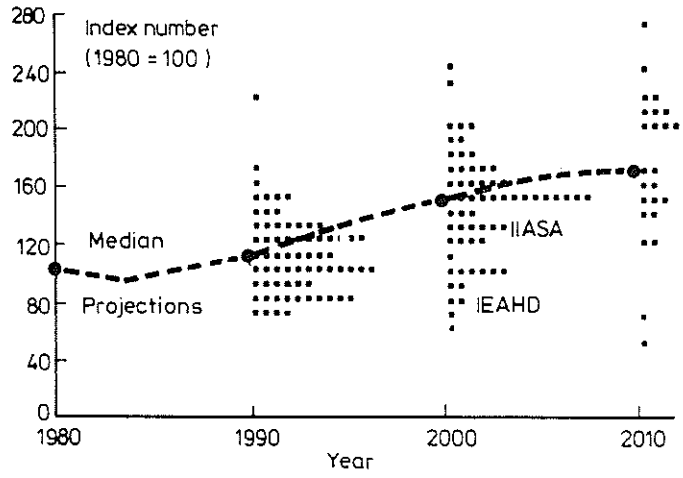


Figure 1. International Price of Crude Oil

The median projections made in 1983 indicated that the price of oil in 1990 will be lower than it was in 1980. The 1990 projections have come down to the low $30s. The projections for the year 2000 are at about $60. The index number is 150 in 2000. The index number of 150 implies a $60 price in 1984 dollars. So, half the groups project a price in 2000 above $60 and half below. During the last couple of years the median projection has come down to about $50.

One forecast shown in Figure 1 is the International Energy Agency's High Demand forecast (IEAHD). This projection was not a systematic matching of supply and demand, but rather, was designed to show just high demand. Thus, the price was set at a level that reflected 20 million barrels of excess demand on the market in 2000. So, the I.E.A. projection was much lower than the median projections. However, if the surplus in the market is dissolved, the median is produced.

Another group, the International Institute for Applied Systems Analysis (IIASA), published a report that showed 1.2 times the 1980 price, which is about $40 or $50 a barrel, through 2000. This was based on the optimism that IIASA had that synthetic fuels would be widely available and would be less than $50 a barrel. Putting the U.S. and Canada together, the analysts projected a supply of 10 to 12 million b/d of synthetic fuels in North America by the year 2000.

So, one can see again the diversity of opinion in forecasts. My general conclusion about using analysis is that although it is impossible to make very accurate projections about the future of the world oil market, much more than nothing is known.

A typical model of the world oil market that has OPEC supply, non-OPEC supply and demand in developing and undeveloped nations. It is designed to balance supply and demand. Such a model is based on some reasonable assumptions such as 3% economic growth until the end of the century and a price elasticity of delivered energy in the aggregate of 1.0.

Outside the Persian Gulf and the centrally planned economies, it is assumed that there are 1.5 trillion bbls of oil-in-place awaiting discovery and that the target level of Persian Gulf production is 20 million bbls per day.

The findings are consistent with other poll projections. With the reference assumptions, projections are for $31/bbl in 1990 and $50/bbl by the year 2000.

Minimum/Maximum Scenario

Next, high and low values for different input assumptions were incorporated. The sensitivities from 1990 to the year 2000 reflect changes in assumptions about economic growth; the price responsiveness of demand to changes in prices; the level of resources outside of OPEC, centrally planned economies and the Persian Gulf; and the target level of Persian Gulf production.(The Persian Gulf is treated as a cartel.)

The model is sensitive to these inputs and outputs. Minimum/maximum scenarios give the best and worst possibilities (Figure 2). A particular project can be evaluated using highs and lows. In fact, many large corporations use the high/low scenario analysis approach in deciding whether or not to invest in a given project.

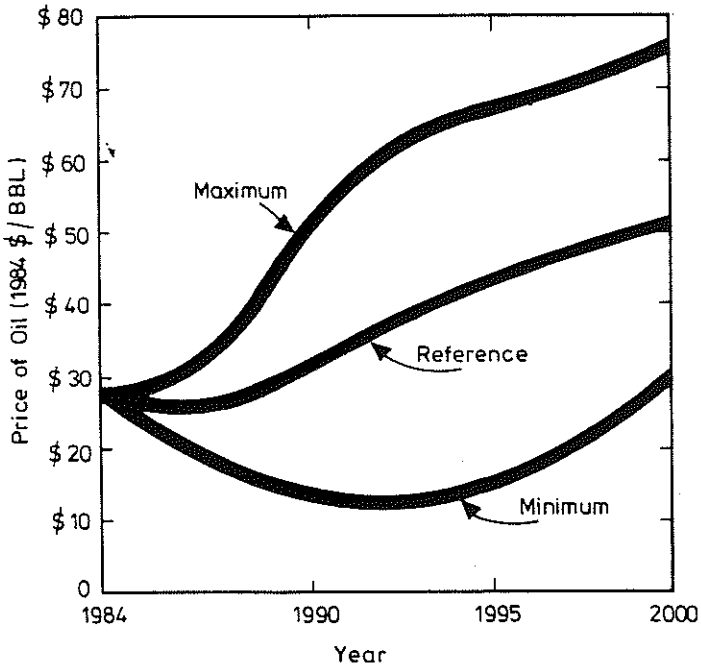


Figure 2. Range of Oil Price Projections

Incorporating Uncertainty

Taking into consideration the major uncertainties in oil price forecasting, it was possible to perform a sensitivity analysis (Figure 3). This is done by putting in extremes in the various inputs. For example, an economic growth of 3% with a probability of 50% was used. This is part of projecting uncertainty into planning.

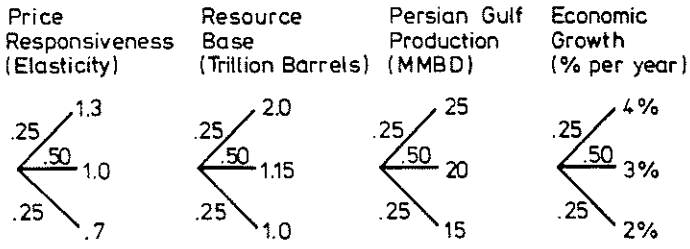


Figure 3. Key Uncertainties in Oil Price Forecasting

Results, after the probabilities were inputted, ranged from $15 to $50 a barrel for 1990 (Figure 4). This was the range from the minimum to the maximum price projection. Thus, there is a 50% probability of the price being below $25 and a 25% probability of the price being above $35. If one wants to include 75% or 80% of the probability, you would have to go from $20 to $40 a barrel.

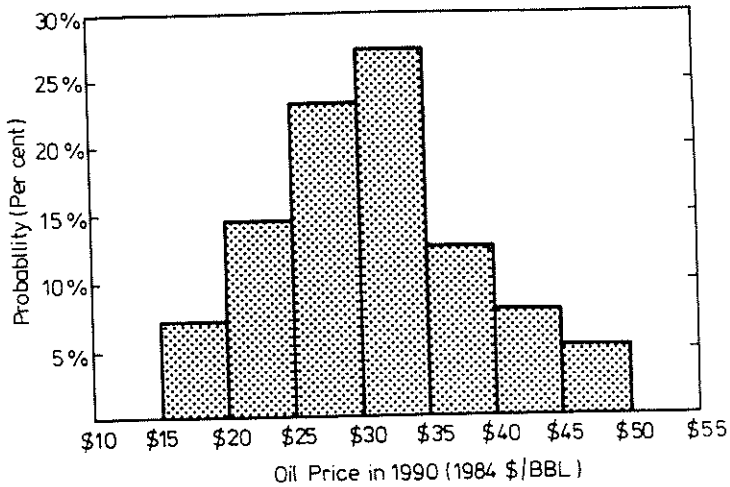


Figure 4. Distribution of Oil Projections

Highs and lows on the year 2000 chart (Figure 5) do not grow much with respect to the 1990 chart. There is about a 10% chance that prices will be below $40 in the year 2000 and a 15 to 20% chance that prices will be above $60/bbl.

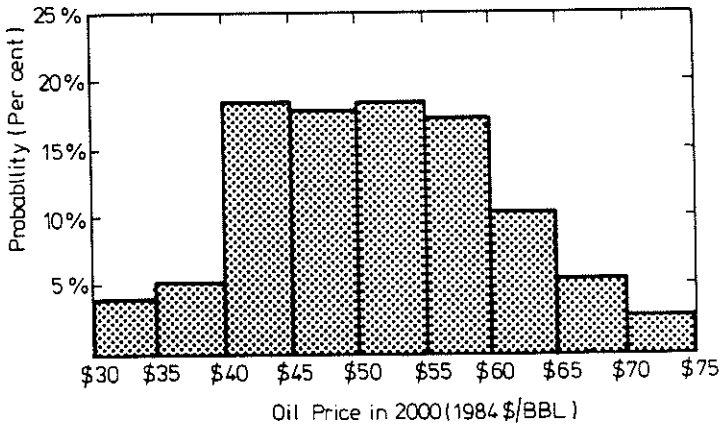


Figure 5. Distribution of Oil Price Projections for 2000

Conclusions

It would be wise never to rely solely on point forecasts. Another important factor in forecasting is to identify uncertainties. Analysts can take probability distributions on the key determinants of the world oil price and translate those probabilities into an understandable format. It is also important to quantify outcomes. Contingencies should be prepared on the upside and downside, and they should be compared.

Finally, any result from an analysis that does not coincide with intuition should be assumed to be wrong. Then, based on subsequent analysis, a mistake may be uncovered or intuition revised.

SECTION 2

DEMAND FORECASTING METHODOLOGIES

Energy Vol. 15, No. 3/4, pp. 207–211, 1990
Printed in Great Britain

0360-5442/90 $3.00 + 0.00
Pergamon Press plc

2.0. INTRODUCTION

ANDREW J. LIPINSKI

Institute for the Future
Menlo Park, California
U.S.A.

Although located in this chapter, the following section serves as an introduction to forecasting methodologies in the entire book.

Choosing appropriate methodology for the task in hand is not easy. Typically a forecast makes use of a model, or several models. The models in turn may require the same or different forecasting methodologies. Both the methodologies and the models employing the methodology will be described, the model exemplifying a particular methodology.

In the introduction we review the factors that guide the choice of the methodology - and of the model - in an effort to place the individual contributions within a common framework. The framework has been selected to be consistent with the theme of this book, namely, that forecasting and modeling are not the end objectives. Rather, forecasting and modeling are tools to assist decision making.

The Purpose of Forecasting

In the broadest meaning of what it means to "assist", forecasts are meant to educate the decision maker or the decision making team. Some of the contributions in this chapter emphasize the educational aspect of forecasting. For example, sections 2.2 and 2.3 portray the complexity of energy flows in the commercial sector of the economy while section 3.8 illustrates connection between demand, supply and price of crude oil.

However, the more fundamental purpose of forecasting is to present decision makers with information about the future values of a handful of variables that will be needed to evaluate the "goodness" of a decision. Assessing the goodness means choosing the option, policy or strategy that is likely to result in a set of outcomes most consistent with the decision makers' preferences. The role of the forecasting model is shown in Figure 1.

Several important recommendations naturally follow if one adopts such decision oriented forecasting:

- The decision makers value model should be communicated to the forecaster prior to building the forecasting model. The forecaster would then be able to concentrate on the few variables that are sufficient to portray the outcome for the decision maker.

- By building the appropriate "connections" the forecaster can ensure that the forecast of the future state of the outcome variables can respond to the settings of the "decision" variable or variables. Thus different states of the decision variable(s) (options or policies available to the decision maker) could result in different forecasts. The decision maker should be able to "play" the options and observe their effect on the outcomes.

  Because almost invariably the exogenous variables and often the forecasting model itself are most likely to describe an uncertain situation, the forecaster ought to incorporate or attach a probability assessment to the forecast. How likely is it, in the judgment of the forecaster, that this particular forecast might be true?

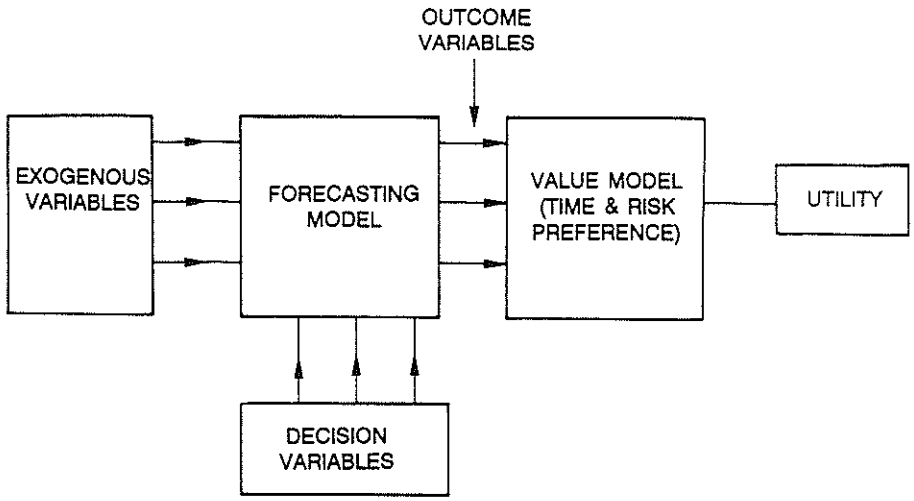


Figure 1. Place of Forecasting Model in Decision Making

The Role of Judgment in Forecasting

There are no "objective" forecasts and there are no "objective" methodologies because, one way or another, both forecasts and methodologies include subjective judgments about:

- Choice of the model or a combination of models that best capture the "essence" of the situation or of the decision problem.
- Choice of the exogenous variables on which the outcome (i.e. the future state of the key variables) will partly depend.
- Selection of historical data from which one might be able to draw inferences about the future behavior of some or all variables.
- Reliability and hence the applicability of historical data to that particular situation.
- Choice of functional representations connecting the state variables, the state variables and the modeling variables and between the variables within the model.

Credibility of a Forecast

To enhance credibility of the forecasts in the eye of the decision maker(s), the forecaster should indicate why a specific combination of a particular model and forecasting methodology was chosen for the task and flag the key judgements made.

Choice of Methodology

Typically, more than one methodology is available to meet the forecasting objective. Several will be described in the book. The choice of methodology will used will depend on:

- Time available for project completion.
- Resources (money, manpower, facilities).

- Nature of the situation to be captured by the forecast.
- Decision maker's familiarity with a particular approach.
- Subsidiary objectives of forecasting - for example, education of the decision maker, objectives or preferences of the forecaster, etc.

Forecast as a Source of Advice

A particular forecast, offering the decision maker information about the future state of variables relevant to a decision, is but one source of "advice" from among the many that are available or offered. In that sense, a forecast is like an expert judgement and therefore is merely one of the many sources of advice on which the decision maker relies. The model becomes an "advisor" and must compete in credibility with other, human advisors. It follows that several forecasts or forecasting models can be used in combination with several human advisors and that the decision maker faces the task of assigning the "likelihood function" to the forecast of a model in exactly the same way that he would have to assign the likelihood function to the forecast made by a human advisor. For all possible outcomes, what probability would the decision maker(s) assign that the output of the forecasting methodology might indeed be true? If the oil prices were to double, what, in the decision maker's judgement, is the probability that the model using this particular methodology would produce the forecast that it did? And if they were to fall instead, or remain steady, what then would be the probability that the model would produce the forecast that it did?

This is why credibility and "transparency" of the forecasting method is so important - it allows the decision maker to assign a higher relative weight (compared to other sources of advice) to a credible methodology and understandable forecast. Notice that credible methodology and an understandable forecast should be the determinants of the relative weight, and not what the forecast has said and whether it agrees with what the decision maker would like to hear.

The Understanding of the Situation Behind the Forecast

A forecast deals with some future situation of interest. The possible range of understanding of that situation is enormous, from celestial mechanics which is so well known as to be virtually deterministic (except for possible surprises such as the intervention of comets) to well proven engineering processes, to less predictable behavior of the economic systems, to poorly understood behavior of the political systems and, continuing, to the almost incomprehensible behavior of the individual who may be a key actor in the system under consideration and finally, to the characteristics of a technology that has not yet been invented.

As the understanding increases, the uncertainty shrinks, except in probabilistic modeling where there may be little uncertainty about the properties of the probabilistic system but the system still produces uncertain outcomes. The roulette wheel is well understood yet still produces (or ought to produce) an outcome characterized by a uniform probability distribution. Generally, well understood systems can be described in terms of fewer variables, mostly because, being well understood, it is possible to exclude variables of lesser importance.

The most challenging problem of forecasting is one in which the structure of the situation being forecast is likely to change within the time span of interest. This may require the forecaster to change from a more deterministic to a more probabilistic model (or vice versa). Alternatively, at a certain time within the span of the forecast the forecaster might have to introduce two or more structures side by side not being sure which is going to be the "correct" model of the situation. And more vexing still is the possibility that the structure of the problem will smoothly change as time goes on, except that the forecaster does not know either the direction or the amount of change. Yet, unfortunately for some forecasters, this is sometimes how structures behave in real life.

The Role of Modeling in Forecasting

In completely judgmental forecasting no model is necessary. But most forecasting methods introduce a model of some sort. From a structural point of view, models have been categorized in terms of their complexity, dynamics and ability to capture uncertainty.

All but one corner of the three-dimensional characterization of Figure 2 offer models to choose from. The corners and the examples of a model belonging to the particular corner are:

1. An equation (Simple, deterministic, static)
   Example: Area of a circle.
2. Linear programing (Complex, deterministic, static)
   Example: Best location of a warehouse.
3. An equation of motion (Simple, deterministic, dynamic)
   Example: Acceleration of a bullet.
4. System dynamics (Complex, deterministic, dynamic)
   Example: Flood control systems.
5. A Bernoulli function (Simple, probabilistic, static)
   Example: Coin toss.
6. Markov chain (Complex, probabilistic, static)
   Example: Probability of a subsequent purchase.
7. A stochastic process (Simple, probabilistic, dynamic)
   Example: Failure of a bearing after hours of use.
8. No model available (Complex, probabilistic, dynamic)
   Example: Most political-economic systems.

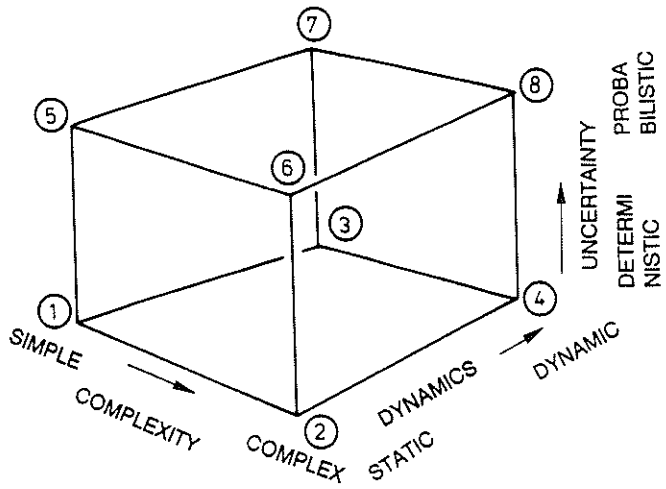


Figure 2. Categorization of Models

Models described in this book fall into all the categories listed above. As one would expect most of those models are complex and dynamic and, as one would also expect, several forecasters use sets of models of different types in making their forecasts.

Guide to Forecasting Methodologies Described in Chapter 2 and Chapter 3

In section 2.1 Paul Werbos introduces econometric techniques. Econometric methodology is by far the most prevalent in modeling the behavior of complex systems. Econometrics is an example of dynamic, deterministic modeling. To provide information about the future uncertainty, econometric models must be combined with a probabilistic model. If the uncertainty lies mostly in the exogenous variables (the assumptions on which the model depends) then the probabilistic model should precede the econometric one. If some of the critical assumptions within the model are uncertain (such, for example as the future price elasticity of demand) then the model could be run several times, using the Monte-Carlo technique with the weighing corresponding to the probability assessments of different price elasticities.

Econometric techniques assume that the reliable historical information is: 1) available, and 2) a good guide to the future. If either of these assumptions is questionable then one could use a judgmental approximation to econometric modeling, such as the cross-impact models described in section 3.6 or the system dynamics approach described in section 3.5.

In section 2.2 Walter Labys and Hiroshi Asano describe the methodology involved in creating process models. Process models attempt to optimize a particular variable (say, the cost of a product) subject to several constraints (such as production capacity). In its simplest form, process models are complex, deterministic and static. When coupled with an econometric model, they become dynamic. Process modeling is an example of the decision maker specifying the variable of interest (e.g. product cost) and the objective (lowest product cost). There is no need for including decision variables in the input - the model adjusts several internal variables until the desired result is achieved.

In section 2.3 David Reister develops the idea of hybrid modeling one step further, combining the best features of process and econometric modeling. The primary objective here is to educate the policy maker by showing how energy flows in a complex system.

In section 2.4 Bertrand Chateau and Bruno Lapillone describe the end-use models. The end-use model they describe combines the objective of educating the policy maker with explicit acknowledgement of uncertainty. The uncertainty is encoded by means of "scenarios", a "what-if" device that allows the decision maker, at least in principle, to assign probabilities to different scenarios and to weigh the end results according to the probability assigned to the scenarios. The decision maker could explore the effect of alternative policies (e.g. those of energy saving) under alternative situations of the end-use of energy.

In section 3.6 Andrew Lipinski shows how the cross-impact modeling offers a substitute for an econometric approach where historical information is absent. Because in cross-impact models the interrelationship between variables have to be elicited from experts rather than computed through regression, cross-impact technique is limited to simple assessments that human could make and that is describing interactions between pairs of variables. Variable A can be a function of variable B, but not of B and C. Even then the burden of eliciting the information is severe and automation of the elicitation process is advisable.

In section 2.5. the cross impact model described in section 3.6 is combined with a probabilistic model. Andrew Lipinski describes the attempt to forecast price differentials between the types of crude oil that could result from the unstable situation in the Persian Gulf. The uncertain situation in the Gulf is captured by a probabilistic tree model. The conditional probabilities are assured by groups of experts. Here we also encounter the use of a "surface" model that reduces the findings of the cross-impact model to one equation. This equation is then run through the thousands of branches of the probabilistic tree.

Energy Vol. 15, No. 3/4, pp. 213–236, 1990
Printed in Great Britain

0360-5442/90 $3.00 + 0.00
Pergamon Press plc

2.1. ECONOMETRIC TECHNIQUES: THEORY VERSUS PRACTICE

PAUL J. WERBOS

Energy Information Administration[+]
Department of Energy, Washington D.C.
U.S.A.

Abstract - This paper introduces the basic concepts used in econometric modeling, and describes five prescriptions to avoid common real-world pitfalls in that style of modeling. The paper begins by comparing econometric modeling with other forms of modeling used in energy modeling and engineering. It describes what an econometric model is, and how to build one. It then gives a detailed explanation of many facets of the five prescriptions: pay attention to uncertainty; don't expect a free lunch when devising specifications; pay attention to prior information; don't expect to draw conclusions without adequate data; and check the historical track record of your model. The issues of generalization and robustness over time receive special attention; they are important in practice, and subtle in theory. Finally, the paper discusses model development in practice, building upon experience with PURHAPS, a model I developed for the Energy Information Administration (EIA).

1. BACKGROUND

Economic theory, in the United States, usually begins with simplifying assumptions like free markets, perfect competition, no externalities, and perfect foresight. After years of study, the advanced student is told how to modify this theory to address real problems in the real world, which are often quite different from the theory in important ways. Some students never quite make the adjustment.

Econometrics is very similar. This paper will introduce the novice to the basic assumptions and methods of econometrics, and then discuss problems which come up in modifying the theory to fit the real world.

Broadly speaking, there is no sharp dividing line between econometric models, engineering process models, statistical models, simple time-series models, systems dynamics models, etc. All these types of models are systems of equations designed to forecast or simulate whatever we want to forecast or simulate. The real difference lies in how we obtain information or parameters to plug into the models.

Some classes of models tend to rely on a priori information or indirect information about what we are forecasting; models of this sort include "pure" process models, classical systems-

[+]This paper expresses the views of the author, not those of EIA or DOE, though it was reviewed at EIA prior to submission. As this paper goes to press, the author's address has changed to: Room 1151, NSF, Washington D.C. 20550.

dynamic models and expert systems. Other classes are based on empirical data about exactly the kind of variables we are trying to forecast; this includes econometric models, time-series models, statistical models, (identified) control-theory models, and artificial neural networks.

In the energy business, the a priori models tends to be very complex, because people include lots of lower-level detail to create a feeling of earthy realism; however, the parameters are usually based on judgment or guessing, and it is hard to be sure the model will track actual trends. The good empirical models tend to be simpler, because they are usually limited to variables which are observed on a time-series basis; however, they are strongly rooted in empirical reality, if done right.

The a priori models sometimes seem easier to understand, at first, because they mimic concrete, well-known engineering processes (at least in part); however, because they contain so much detail, it is not always easy to know what causes the bottom-line forecasts to come out as they do, and the role of human behavior is often neglected or oversimplified. Empirically-based models are the reverse: the overall behavior is easier to understand, but the detailed reasons behind the trends -- both historically and in the forecast -- may require further analysis. A good researcher will learn how to combine both prior information and empirical information into a model, as this paper will discuss.

The relations between different kinds of empirical model are subtler.

On some level, there is no real difference between a statistical model, an econometric model, and a model developed by using the identification techniques of control theory; all three rely on the same core of theory, which this paper will discuss. "Simple time-series models" and artificial neural networks depend on the same theory as well, but they try to automate the process of coming up with a functional form; in effect, they assume that the user does not really understand the structure of the system he is studying, so that a computer can do the job as well as a person. This is a good assumption in some cases (as in recognizing patterns among thousands of variables which no one fully assimilates into his or her intuition), and a poor assumption in others (as in the study of physical phenomena for which the dynamics are well-understood).

2. GOALS OF THIS PAPER

In the United States and many other nations, econometrics is a major academic discipline, based on the idea that a careful analysis of historical data can be a good starting point for analyzing or projecting the future. Like any major discipline, econometrics has a long history, full of false starts, new perspectives, and hundreds of applications, some good and some bad.

This paper will present those concepts and rules of thumb which we have found most important, in practice, in a government organization concerned about the quality of its forecasts. No one can expect to become a first-class econometrician after reading one article; there are simply too many tricks and traps to learn. However, we will try to explain the key concepts, and cite books which elaborate on their application. We also hope to pinpoint those misunderstandings which are common among experienced practitioners, and we apologize to them that there is not enough space here to explain all the details. Unfortunately, these misunderstandings have often led to the creation of models which totally misrepresent the dynamics of the variables which they are supposed to predict.

This chapter will begin by saying what an econometric model is and how -- mechanically -- to build one. Next, it will discuss five major prescriptions for the correct use of econometric tools. Then it will discuss the use of these tools in practice at the Energy Information Administration (EIA). It will conclude with a very quick overview of the PURHAPS model, one of the econometric models I have developed for EIA.

3. WHAT IS AN ECONOMETRIC MODEL?

Strictly speaking, an econometric model is no different from any other forecasting model -- it is made up of any set of equations or formulas which can be used to predict the future. For example, consider the following simple model to predict population:

$$POP(t+1) = c \cdot POP(t) \quad (1)$$

This says that population in year t+1 is equal to a constant "c" multiplied by the population in year t. If you obtain your estimate of the constant c by asking your boss what c should be, or by studying textbooks on theology or ideology, then we would call this a judgmental model. If you obtain your estimate of c from small-scale case studies of controlled populations, we might call this an engineering model. If you have an historical time-series of data on population, for the state or nation whose population you are forecasting, and if you estimate c from that time-series in a rigorous way, then we would call this an econometric model. In principle, then, there is no such thing as an econometric model; there are only econometric methods for estimating parameters such as "c" in general models. A pure econometric model is simply a general model, in which all of the parameters have been estimated by econometric methods, based on empirical data.

Econometric methods were initially developed for use in economic forecasting. However, there is nothing in our discussion which will restrict their use to economics. Econometric methods have often been applied directly to forecasting social and political systems (Werbos, 1974; Werbos, 1977; Werbos and Titus, 1978). Human minds and computers which truly imitate human minds must also have a built-in capability to learn cause-and-effect relations by somehow analyzing a time-series of sensory experience; we have shown how econometric methods may be embodied directly into the wiring of such systems (Werbos, 1987a; Werbos, 1986a).

In general, people who use historical data or trends or track records to help them make decisions are making inferences about cause and effect. Like it or not, they are engaged in a form of statistical inference. Even if they say they are merely testing an hypothesis, or a relation, and not formulating a model, the fact is that they are estimating a model; the potential for error and uncertainty is merely less visible and harder to correct when they deny this fact. (Of course, some managers would prefer to hide such uncertainties from their superiors. If a superior really cannot understand econometrics, there is an art to using econometrics properly and then translating the results back into English, using graphics and discussions of percentage growth rates and historical analogues.)

4. HOW CAN AN ECONOMETRIC MODEL BE BUILT?

The first stage in building a model is to review the available data and concepts, as we will discuss further in the section on "Practice".

Next one must choose a computer package to work in, to implement the econometric methods. EIA generally prefers to use the SAS package (SAS, 1985a; SAS, 1985b) on its large computer, because of its superior flexibility and data-handling capabilities; however, Troll (1981) has also been used, because of the sophisticated econometric tools it contains (some developed under contract to us). On microcomputers, SAS is also available, but is relatively expensive at present; Lotus is widely used, and new packages from Wharton Econometric Forecasting Associates (of Philadelphia) and elsewhere may be used more in the future. We use SAS to estimate parameters such as "c" in equation 1, and to evaluate the overall degree of fit of equations such as 1 and alternatives to 1; then, when all the equations are estimated and selected, we usually program the forecasting itself in FORTRAN. Actually, SAS and Troll have

the capability to simulate the model -- to generate forecasts - as well; we have used that capability only rarely, because our models have usually been too big to fit into those systems.

The next step is simply to use the package chosen. To estimate equation 1, for example, you would first locate a time-series of data on the variable POP, and load it into a SAS dataset using the SAS command DATA (SAS, 1985a). Then you would use a SAS command such as GLM (SAS, 1985b) to estimate "c" in equation 1, and to evaluate the error which this equation would have led to in forecasting the past. You could also use SAS to estimate alternative equations, and their errors, and you could select between equations based on their error. At that point, you have the equation, and you need only code it into a forecasting program.

The most common way to estimate a complex model, in econometrics, is to use "regression" or "least squares." Using regression, we estimate each equation of the model separately, one after another. For each equation, the regression command finds those values for the parameters which lead to the smallest possible error over the historical period you have data for. "Error" is defined as the sum over all observations of the square of (actual minus predicted). Regression also reports what the error is for the equation as estimated.

In actuality, most computer packages have two main regression commands available -- a linear regression command and a nonlinear regression command. (See Wonnacott and Wonnacott, 1977, Chapters 13 and 15, for more explanation.) To avoid complications, most economists use linear models such as the following two-equation model:

$$Y(t) = a \cdot Y(t\text{-}1) + b1 \cdot X_1(t) + \ldots + bn \cdot X_n(t) + c \qquad (2a)$$

$$Z(t) = c1 \cdot Y(t) + c2 \cdot Y(t\text{-}1) \qquad (2b)$$

In equation 2a, Y(t) is the "dependent variable" -- the variable being predicted in that equation. Y(t-1) and $X_1(t)$ through $X_n(t)$ are the independent variables of that equation.

The term "c" is the "constant term" or "intercept;" note that equation 2b has no intercept. The parameters of the model are the constants a, b1 through bn, c, c1 and c2. The "endogenous variables" -- Y and Z -- are the variables being predicted somewhere in the model. Y(t-1) is a "lagged endogenous variable" (because Y is endogenous and because t-1 represents a "lagged" value, a previous year's value.) X_1 through X_n are "exogenous" because they are not endogenous.

This example is linear, because the dependent variable in every equation is predicted as a linear combination ("weighted sum") of the independent variables, plus an optional constant term. To estimate each equation in SAS, you need only use the linear regression command (GLM or something similar) once for each equation. You can be sure of quick results, and you do not have to give an initial guess for the values of the parameters. Each time, you only have to tell SAS the name of the dependent variable and the names of the independent variables. You also have to tell SAS whether you want a constant term in the equation, and whether you want SAS to print out all the diagnostic statistics anyone has ever thought of.

At first glance, equation 2a may appear somewhat abstract and unrealistic. Economic relations in the real world are often more complex. For example, even in a simple model of fuel oil demand (QOIL) as a function of residual oil prices one would not want to use the simple equation:

$$QOIL = a \cdot PRESID + b \cdot PDIST + c \cdot DISTSHARE \qquad (3)$$

If you used this equation, by regressing QOIL on PRESID, PDIST, and DISTSHARE, you would expect to find that "a" and "b" are estimated as negative numbers, expressing the idea that higher prices lead to lower demand. However, with "a" and "b" negative, there will always exist a price so large that demand becomes negative, which is an absurd forecast. Likewise, the effect of

changes in PDIST should depend on how large the distillate share is; PDIST and DISTSHARE have an "interaction effect." For these reasons, a better specification would be:

$$LOG(QOIL) = a + b \cdot LOG(POIL) \tag{4a}$$

$$POIL = PDIST \cdot DISTSHARE + (1 - DISTSHARE) \cdot PRESID \tag{4b}$$

Equation 4a says that QOIL is a function of the weighted average price of fuel oil, POIL. It says that a given percentage change on POIL leads to a proportionate percentage change in QOIL; the factor of proportionality is just "b", the price elasticity of demand. (To see this, differentiate 4a or see Wonnacott and Wonnacott, 1977, Section 15-3). Equation 4a can be estimated easily in SAS by first using a DATA step to calculate:

$$LOGQOIL = LOG(QOIL)$$

$$LOGPOIL = LOG(PDIST \cdot DISTSHARE + (1 - DISTSHARE) \cdot PRESID,$$

and then calling the regression command and asking it to regress LOGQOIL on LOGPOIL. Equation 4b contains no parameters at all to estimate; it is called an "accounting identity" (as opposed to the "behavioral equation" 4a). Equation 4a is linear in the parameters a and b, but not in the original variables QOIL and POIL. Most econometric models are linear in the parameters but not in the original variables. Most of them also use tricks like the above to express economic relationships.

If equation 4a had actually been nonlinear in its parameters, then nonlinear regression could have been used. Nonlinear regression requires a lot more care and patience, depending on what computer package you use, but there is usually a way to make it work. Likewise, there are alternatives to regression which would require you to estimate the entire model as a system, together; to use these alternatives, you would have to type both equations into a single model file or command block.

Aside form their linearity, the models in equations 2 and equations 4 have two other simplifying features. First, they are "recursive". In economics, this means that they are really just simple formulas; you can calculate a forecast by plugging in values for the exogenous variables and lagged endogenous variables, and using the equations one after another like a formula or a recipe. Most econometric models are actually simultaneous, as in the following example:

$$LOG(SUPPLY) = a + b \cdot LOG(GNP) + c \cdot LOG(PRICE) \tag{5a}$$

$$LOG(DEMAND) = d + e \cdot LOG(GNP) + f \cdot LOG(PRICE) \tag{5b}$$

$$SUPPLY = DEMAND,$$

where GNP is exogenous and where the model is used to forecast a PRICE that makes SUPPLY and DEMAND balance. to make a forecast, you cannot jus plug in GNP and PRICE on the right-hand side; you cannot, because you don't yet know that PRICE is. You have to solve this system of equations, as a set of three simultaneous equations in three unknowns. In fact, if you insert these three equations into Troll (1981), Troll will take care of this problem and give you a set of forecast which solve the equations.

Notice that it would be very dangerous to estimate a system like this by ordinary regression. If SUPPLY did equal SEMAND in all historical years, then you would get exactly the same set of parameters (d,e,f) when you use regression on 5b as you did (a,b,c) when estimating 5a; you would not really have two different equations. Even if the equations were very slightly different, you could not rely on what you get when you subtract one from the other (as required in solving

them). In these kinds of situations, it is important to estimate the model as a system (Wonnacott and Wonnacott, 1977, chapter 22).

These situations arise in energy modeling, but the problem is usually not significant, mostly because we deal with dispersed system involving lagged responses. The systems estimation methods often lead to worse results, because e of their complexity, because the use of instrumental variables introduces random noise, and because of problems with "robustness" (discussed below). On the other hand, the simultaneity problem can be serious with fuels like LPG and other minor forms of oil, whose markets are very limited and respond quickly to price; our goal, in those cases, is to look for something like a "reduced form" model for each such fuel (e.g. in equations 5 first solve, to get log(Price)=g+h·LOG(GNP), and then estimate g and h).

Even the model in equations 5 still has one further simplification: all of the variables are assumed to be available for all historical years in you data base. (SAS will overlook a few missing values here and there, however.) It is possible to build econometric models which do not have this property, because they include "time-varying parameters" or "hidden variables;" however, this is not common at present, and the tools to estimate such models are hard to come by. One can work around this problem, to some degree; for example, if the population growth rate, "c" in equation 1, varies over time as a function of women's education (WED), then one might postulate that c=a+b·WED, and rewrite equation 1 as:

$$POP(t+1) = a \cdot POP(t) + b \cdot WED(t) \cdot POP(t) \tag{6}$$

Finally, for completeness, it should be emphasized that variables in an econometric model are not always simple time-series. Many authors will perform regressions on a data base of different observations at the same time, such as data from different states, and then use the results to predict the future. This is called forecasting based on cross-sectional analysis, and the results are usually unreliable at best, both in the short-term and in the long-term. For example, one of the first econometric equations ever studied was the classical consumption function:

$$C(t) = a + b \cdot Y(t) \tag{7}$$

where C is national consumption and Y is national income. In cross-sectional analysis, "a" was significantly larger than zero, and there seemed to be a large saturation effect in consumer spending. But in time-series, "a" was quite close to zero. For purposes of forecasting changes over time, the time-series version is the right one to use. In general, variations across space tend to be different from variations across time, and we have seen this lead to problems over and over again. (For example, see the discussion of "locational bias" in Werbos, 1983, Chapter 4.)

An ideal model should be able to account for variations over time and space both; however, without data from different times, it would be foolish to assume that one has an ideal model. Still, one can collect "pooled" data, which vary over time and space both, as we have often done (Werbos and Titus, 1978; Werbos, 1983). To use such data in packages like SAS can be slightly tricky, when you estimate a model containing lagged variables. In arranging our data (Werbos, 1983), we found it necessary to include a dummy year, 1973, to precede the years for which we had pooled data (1974-1981), and we inserted the SAS missing value code for all 1973 data. Observations 1 through 8 represented 1973 through 1981 in the first state, while 9 through 16 represented the second state, and so on. (Without this, the SAS "LAG" function would not have given us valid time lags.)

5. FIVE FUNDAMENTAL PRESCRIPTIONS

This section will provide a kind of back-door introduction to the theory underlying econometrics, by trying to explain five prescriptions for avoiding gross errors which are common even among professionalists.

- o Pay attention to uncertainty
- o Don't expect a free lunch when choosing specifications
- o Pay conscious attention to prior information
- o Don't expect to draw conclusions without adequate data
- o Check the historical track record of your model (This may be the most important)

Pay Attention to Uncertainty

None of the models above -- from equation 1 to equation 7-- say anything about uncertainty. They are all forecasting models, recipes for making base case projections. Even though there are many different schools of thought in statistics and econometrics, they all agree that uncertainty needs to be addressed explicitly as a central part of the analysis.

Broadly speaking, there are two major schools of thought here:

- o The purist school, which has done an admirable job of simplifying and unifying our understanding of statistical methods, and devising new and better and more elegant methods.
- o The utilitarian school, which has made live complicated and tricky all over again, by focusing on the intractable problems which occur in real-world forecasting. (This is quite different from the quick and dirty school, which pays more attention to deadlines than to quality problems either in theory or in the real world.)

Both schools have a great deal to contribute, but we incline towards the utilitarian school.

From the purist's point of view, regression simply cannot estimate equation 1 as it stands, as if it were a meaningful model of population growth. If you regress POP(t+1) on POP(t) with no constant term, then the model you are really estimating is:

$$POP(t+1) = c \cdot POP(t) + e(t) \quad (8)$$

where e(t) represents a random disturbance, governed (generated) by a normal probability distribution. We sometimes call e(t) "error," but statisticians like to think of it as something out there, in the real world, rather than an "error in the sense of "mistake.". Often we call e(t) "white noise," to make this view explicit. Equation 8 is a "stochastic model," because the assumptions about the random disturbance have been made explicit as an integral part of the model.

When we look at the noise term explicitly, we can see immediately that there is something implausible about the model in equation 8. Equation 8 assumes that the noise comes from the same probability distribution in all years, implying that we should expect the same general size range for the noise in all years. If population grows by a factor of 10 in the period under study, this could be a very poor assumption about the noise; as a practical matter, this assumption would lead to an estimate of "c" dominated by the experience of the last few years, disregarding the earlier data. It is more plausible to expect that the noise will represent a certain percentage of the population, and that its size range will grow in proportion to the population, as in the model:

$$POP(t+1) = c \cdot POP(t) + e(t) \cdot POP(t) \quad (9)$$

In this equation, the overall noise terms -- $e(t) \cdot POP(t)$ -- grows in size in proportion to population; in other words, e(t) -- which now represents noise as a fraction of the population -- still comes from a fixed probability distribution (imposing a fixed size range).

Equation 9 cannot be estimated directly in regression. However, now that we have a complete stochastic mode, it is legitimate to divide both sides by POP(t), as we could with any algebraic equation; this yields the equivalent model:

$$POP(t+1)/POP(t) = c + e(t) \tag{10}$$

This can be estimated by regressing POP(t+1)/POP(t) on no independent variables plus a constant terms; in practice, this is just a matter of estimating the average value (mean) of POP(t+1)/POP(t). It is more conventional, however, to use a similar but slightly more plausible alternative to equation 9:

$$LOG(POP(t+1)) = c' + LOG(POP(t)) + e(t) \tag{11}$$

which is equivalent to:

$$LOG(POP(t+1))/POP(t)) = c' + e(t) \tag{12}$$

More generally, equations like equation 8 -- which assume a constant size range for error when a constant size range is not plausible or does not fit the data -- are said to have a problem with "heteroscedasticity." This is a common problem, and algebraic transformations (like the above) are commonly used to overcome it. Sometimes, however, algebraic transformations are not a workable solution. For example, when the dependent variable is LOG(QOIL/QGAS), as in the standard "logit" specification for fuel choice, there is a heteroscedasticity problem which can only be resolved by resorting to weighted regression, which explicitly treats the size range of e(t) as a function of other variables; the theory is given in Pindyck and Rubinfeld (1976), and applied in the PURHAPS model (Werbos, 1983, p.12,64). (This correction would have been desirable, but far less necessary, if we had worked with a simple time-series showing no order-of-magnitude variations in fuel shares.)

Besides heteroscedasticity, there are other possible problems with the theory that e(t) is random and normal across time. For example, e(t) may be correlated with its previous value, e(t-1). When the standard Durbin-Watson test (available in SAS and other packages) gives a score much different from 2.0, it is conventional to use a different regression command -- regression with an autocorrelation correction -- to estimate the model under the assumption that $e(t)=r \cdot e(t+1)+a(t)$, where a(t) is random; if r -- the "autocorrelation parameter" -- is not significantly different from zero, one can go back to using conventional regression.

Recently, many statisticians have begun to recommend a more careful study of the model residuals, e(t), to see if they fit more complex "Box-Jenkins" models (Box and Jenkins, 1970). In theory, certain classes of Box-Jenkins models can represent the idea that forecast errors result from a combination of noise in the real world and noise in measuring what is happening in the real world. These kinds of models can reduce forecasting errors, but tests done for real-world multiple-equation models (Werbos, 1974; Werbos and Titus, 1978) suggest that it would be better to focus on the long-range track record of a model, as we will describe below. (Engineers have another way of estimating such stochastic models, but their formulation, unlike the statisticians' formulation, contains excess parameters and can almost never be uniquely estimated.)

All of these recommendations are based on the following fundamental theorem, an application of Bayes' Law, which underlies all inference from empirical data (in statistics or in other fields):

$$Pr(Model|Data) = Pr(Data|Model \cdot Pr(Model)/Pr(Data) \tag{13}$$

This states that the probability of a model being true, after we have observed a certain history of data, is the product of three terms. One of these terms -- Pr(Data) -- is the same for all models, and has no effect on our relative choice between models. Another -- Pr(Data/Model) -- refers to the probability that we would have observed what we did in the data if the model were true. For stochastic models, like equations 8 through 12, this term can be calculated directly by calculating what e(t) would be in all the years of data (assuming a given estimate of the parameters as part of the "Model), and then using the normal probability distribution to calculate the associated probabilities. This term is called the likelihood function; it is a function of the parameter estimate, the model, and the data. The remaining term -- Pr(Model) -- represents the probability that a model would be true, a priori, before any statistical data are examined.

Most purists agree that it would be unscientific to account for Pr(Model) explicitly in statistical estimation. They argue that all possible models (and all possible sets of parameter estimates) should be treated as equally probable a priori. They argue that modelers should always estimate these models by finding parameter values which maximize the likelihood function. Most existing statistical packages do in fact maximize likelihood exactly (as in regression) or approximately (as in iterative methods which imitate the full information maximum likelihood command for estimating systems of equations).

Bayesian statisticians have argued that economists have important information, prior to statistical analysis, about the relative probability of different models and parameter values. We would agree, but would argue that the economists' information is very complex; it would be better to use the computer to produce a complete, graphic description of the likelihood functions -- the information found in the data -- and then count on the human being to account for his prior information after the statistical analysis is complete. This puts a heavy burden on the person doing the statistics, since it is not enough to just print our estimates of one final equation; it is essential to consider the range of uncertainty for all the parameter estimates, and to consider different ways of looking at the data.

Utilitarians (like us) go further, and argue that simple statistical models are never "true" in any absolute sense. They argue that your choice of estimation method should depend on the application of the estimates or forecasts. The overemphasis on definite, base case forecasts is a product of naive decision-makers, who have yet to understand well-known procedures for coping more honestly with uncertainty (Brown et al., 1974). Indeed, one may argue (Werbos, 1979) that probabilities, rather than expected outcomes, should be the main focus of long-range planning anyway; however, the efficient implementation of this principle involves many complexities (Werbos, 1987a; Werbos, 1986a). The utilitarian Raiffa has found that elite Americans tend to understate ranges of uncertainty by a factor of 3 or so, perhaps because they do not account for the limitations of the assumptions they use. This suggests a need for great care in using mathematical models built on expert judgement rather than empirical fact. Raiffa's followers, such as Rex Brown, have developed many techniques to train, improve and organize probability assessment by human judgement; nevertheless, the problem of bias remains difficult and fundamental. It is important that modelers help decision-makers think more clearly about alternative scenarios, rather than aggravate these biases. Even though it is very difficult to estimate probabilities objectively -- when technological and political forces are primary sources of uncertainty -- it should be possible to convey the nature of uncertainty in a useful way, and explain alternative viewpoints.

Utilitarians also tend to look for estimation methods which are likely to give more accurate forecasts even when it is hopeless to formulate a model which is "true" in an absolute sense; such methods are called "robust estimation methods," The problem of heteroscedasticity leads to a simple (though unconventional) example of robust estimation. Consider the simple model:

$$\text{ENERGY-USE}(s,t) = c \cdot \text{PRODUCTION}(s,t) + e(s,t) \qquad (14)$$

where energy use is projected by state (s) and by the year (t). A purist would tell us to replace e(s,t) by e(s,t)·PRODUCTION(s,t), and they divide by PRODUCTION(s,t) to get a regression equation. If we do this, we are guaranteed that the percentage error in predicting energy use will average out to zero (i.e., positive and negative errors will balance out). If we had kept equation 14 as is, we would be guaranteed that the actual error in predicting energy use will average to zero; in other words, greater attention would be paid to bigger states. If the goal is to predict total energy use, then the latter is preferable; it would lead to more uncertainty in our estimate of c, in theory, but it would also guarantee that we are estimating that version of c which is right for our application. If we admit that equation 14 is only a simplification, then we have to accept that the version of c which minimizes one error measure will be different from the version which minimizes another. In technical terms, there is a tradeoff here between statistical efficiency (i.e., random uncertainty in our estimate of c) and statistical consistency (estimating the right c). Tradeoffs of this sort are quite common, and often require some sort of ad hoc compromise.

Don't Expect a Free Lunch When Choosing Specifications

Choosing the equations of a model is a difficult process, whether the model is econometric, judgmental, or engineering-based. The process is essentially the same for all three, except that econometricians normally restrict themselves to using variables for which they have data. Econometricians often start from a general theoretical model and translate it into its implications for observable variables; there is no need to represent the entire mechanism by which variable A affects variable B if the ultimate impact is represented correctly. Also, when doing econometrics, you usually consider several alternatives, and use empirical results to decide which version to select in the end. In fact, you typically try out new alternatives after you have studied the results and looked for explanations of what is going on.

There are some analysts who offer you a hope of forecasting without resort to this difficult process. They often suggest that "simple time-series analysis" or "simple econometrics" can be an alternative to the labor and uncertainty which comes with explicit models. In actuality, this is an illusion (though the explicit models of econometrics are simpler than most engineering models). For example, "simple Box-Jenkins analysis" (Box and Jenkins, 1970) offers more complicated models of noise than regression assumes; it essentially offers yet another complex correction to explicit models (Werbos, 1974). The vendors of "simple" analysis typically apply statistical methods to a simple forecasting model, such as:

$$Y(t+1) = a + b \cdot Y(t), \tag{15}$$

where Y is the variable you are trying to forecast. Admittedly, this model is sometimes worthwhile. Admittedly, simple models in general tend to be more robust than complex models, ceteris paribus. Some salesmen have suggested that this approach can be applied to electricity demand, to save utility planners from the pain of using models which require forecasts of local industrial growth, which are fraught with uncertainty. However, this economic uncertainty is real, and unavoidable; a forecaster can hide the uncertainty from his clients (which does them a disservice), but the economic uncertainties are there and will affect electricity demand. If economic growth is known to be central to electricity demand, then it should be reflected in the model. In general, the choice of a model should be based on a careful analysis of what is known about the variable being modeled, and what is shown in the data; there is no magical way to escape this process.

Forecasting problems in private industry are sometimes so complex that analysts cannot devise an adequate specification, even when data are plentiful. In such situations, a full-scale "neuron network" system may be useful. The best neuron network systems (Werbos, 1987a; Werbos, 1988; Werbos, 1989) are essentially equivalent to a massive automated search through all possible specifications -- linear and nonlinear -- to find that specification which minimizes some combination of forecast error and model complexity. The prior knowledge of the analyst

is not used at all (except in the construction of the data base.) At EIA, we have yet to encounter such situations.

In one recent situation (IFS, 1986), we had access to a massive data base on fuel-switching in which we didn't know what to expect. Elaborate cross-tabulations in SAS were very useful in helping us form hypotheses, and then formulate and estimate econometric models.

All of these examples emphasize that a modeler should take the time to devise specifications carefully. Some students are willing to do this, but expect to be given exact rules on what kinds of specifications to use. Once again, they are looking for a kind of fee lunch, and they can find a few misleading papers in journals which give them the rules they are looking for. In actuality, the choice of specification should be based on a translation of your prior knowledge (Pr(Model) in equation 13) into mathematical equations; there is no set rule for what the equations must look like, but there are guidelines for how to do the translation. A good econometrician should have some familiarity with the guidelines for translation (Brown et al., 1974; Forrester, 1961) which have been developed for models in general. Econometricians have developed further guidelines, but they are too numerous to cover here; still, please do consider what happens to your model when the independent variables take on extreme values, and do consider whether the forecasts would change the way you want them to in response to a small change in the inputs (as a function of other inputs). Also consider whether the specification really could represent alternative points of view (e.g. large and small price elasticities) through different parameter estimates.

This notion of translation between human knowledge and mathematics is so vital that it merits several examples.

First of all, translation from English into mathematics may be compared with translation from Chinese into English. In chinese, one can make statements like "man see horse." In English, this could mean that "a man saw a horse", or that "every man sees a horse sometime in his life", or that "those three men are looking at a horse", or that "this man will see a lot of horses", etc. In order to translate from chinese into English, one has to decide what tense to give the verb "see", what number or article to put before the word "man", etc. A good translator will make these decisions based on a careful understanding of the context in which the statement appears. Even then, several interpretations may still make sense; in that case, the translator may go back to the author of the Chinese statement, and ask which alternative would be used. Note that the translator can state the alternatives to someone who only speaks Chinese; the Chinese language permits ambiguity, but does not require it.

An irresponsible translator would not try to understand the Chinese original; instead, he would follow a mechanical rule, such as assuming the present tense in every sentence which does not explicitly refer to the future or the past. Irresponsible translators can easily produce paragraphs in English which look downright silly (as in the instruction manuals which come with certain imported products). In translating from English into mathematics, one can produce silly mathematics just as easily, if one is not careful about the role of time in the equations.

Second, consider a question which the Energy Modeling Forum brought up in 1985; "How much fuel-switching has there been between oil and gas in response to prices in manufacturing?" Two modelers came up with completely different answers to this English-language question, based on the same set of data (Annual Survey of Manufactures) at the State level from 1974 to 1981. One modeler (David Reister of Oak Ridge) translated the English-language question as follows: "In any given year, was the market share of natural gas in manufacturing as a whole much greater in those States where the gas price was a smaller fraction of the oil price?" The other (myself) translated it as: "In any given industry, was the change in market share from one year to the next much greater in those States and years where the change in the price ratio was also great?" These are two different questions, and it is not surprising that they yield different

answers. At EIA, we have tested both kinds of specifications, and were not surprised that the latter type led to much smaller forecasting errors.

Translation back from the statistics into intuitive terms is just as important. For example, a few years ago we reviewed a major paper on new car Miles per gallon (mpg), which developed a model which predicted that a doubling in gasoline prices would double mpg in the future. Working through their vehicle attribute equation, we discovered that this forecast depended on the assumption that new cars would be eight feet tall or eight feet long in order to get higher mpg.

Pay Conscious Attention to Prior Information

Often forecasts try to provide a "most likely" view of the future, based on all of the information available. Historical information is only part of that information. Expert judgement, private sector plans, and engineering information are also part of that information. In an ideal world, we should not have to choose between econometric forecasts versus engineering forecasts, and so on; we should develop forecasts which have the highest probability of coming true, conditional upon all three kinds of information. In an econometric model, this can be approximated by choosing specifications and altering parameters, where necessary, to reflect such information.

This kind of adjustment is a tricky process. There is a risk of confusing the final user, who may not be able to tell what comes from historical trends and what comes from adjustment. Also, when adjustments are made on the basis of judgments, political biases and wishful thinking easily enter in, and cause further confusion. Adjustments based on population wisdom which in turn depends on past history may represent a "double-counting" of history or far worse. Therefore, there is much value in having some forecasters -- such as academics -- produce pure econometric models and leave the discussion of other factors to their verbal discussion sections.

The Energy Modeling Forum (Werbos, 1987b) has given us some examples of how this kind of adjustment can be done and explained to the reader. For example, if an historical trend (reported, say, as $c \cdot year$ in an equation to predict energy intensity) can be clearly explained as the result of using a single technology throughout the historical period, and if we are quite sure that a radically different technology will come on and dominate the forecast period, then an econometric model should be adjusted to reflect our best knowledge of the new technology. Conversely, if new technologies are expected in the future, but are numerous and hard to predict exactly, and if there were also new technologies coming on line in the historical period, one is better off trusting the econometric model.

Don't Expect to Draw Conclusions Without Adequate Data

Econometric models, when estimated, make a statement about cause and effect relations. For example, in equation 6 above, if "b" were estimated as a negative number, this would say that increases in women;s education can reduce population growth by some amount. If b were estimated as a positive number, it would say the reverse. In either case, the estimated value may be a fluke, a coincidence due to a relative shortage of data. To see if this is likely, we need to examine the "standard error" of b, which is possibly the most important statistic printed out by standard regression programs. By and large (Wonnacott and Wonnacott, 1977), the true value of b will be equal to the estimated value plus or minus the standard error, in seventy percent of the cases; it will lie within two standard deviations in ninety-five percent of the cases. Old-fashioned statisticians would say that b "is not significantly different from zero" when the value $b=0$ was between these confidence limits; however, it is better simply to report what the standard errors are, to make it clear to the reader how big b still might be (given the limitations of the data).

Large standard errors may result from any of the following:

- o Lack of data. (Having four times as many independent observations cuts the standard errors in half. However, with pooled data, the observations -- e.g., from neighboring states -- may not be entirely independent, and the true standard errors may be larger than the reported ones.)

- o Lack of an adequate specification. (Cutting historical error in half cuts standard errors in half.)

- o Correlations between the independent variables, or "multicollinearity." (This represents a qualitative lack of data -- a lack of data on situations where the independent values have different values.)

Many social scientists do not appear to realize that standard errors really do account for the effects of multicollinearity. When two independent variables do correlate very strongly with each other, there is no magic procedure to solve the problem; an honest statistical analysis will simply report that there is not enough data available to decide which variable has an impact on the dependent variable. When there is strong multicollinearity (as hinted at by large standard errors), but no really strong correlation between two variables, one should suspect three-way patterns of correlation; to locate these kinds of patterns, one can perform an eigenvector analysis of the correlation matrix, and look for the eigenvector whose eigenvalue is closest to zero. (Belsley, Kuh and Welsch have discussed these kinds of diagnostics.) When there are many variables involved, and when forecasts will be made for situations where the independent variables continue to correlate with each other, methods like "ridge regression" may be better than ordinary regression when multicollinearity is suspected (Dempster et al., 1977). When a model must be estimated, but the data are inadequate, one must generally fall back on prior information (and flag the resulting uncertainty).

Check the Historical Track Record Of Your Model

There are many economists who test out alternative specifications, and publish whatever gets the highest "R–squared" score as printed out by SAS. This often leads to disaster, because R-squared scores are not comparable between equations which represent the dependent variable in different ways; for example, an equation which predicts energy per unit of output will often be more accurate than an equation which simply predicts energy, but will often have a lower R-squared score (because the dependent variable has less variance). The situation is even worse with complex statistical methods as used in fields like cost function estimation; there, the "adjusted R squares" are often aggregate constructs whose relation to forecast error may be quite tenuous.

As a first step, one can try to compare mean square error (MSE) across models, because it is reported by SAS and is more often comparable between equations. As a second step, one can simply use the alternative equations to predict the same basic variable (e.g., energy consumption), and calculate the average error; this can be done in SAS by using the "OUTPUT" option to output the regression predictions to a file, and by using the numerous SAS utilities to calculate the implied predictions of energy and their errors.

In practice, there is no substitute for trying to understand what is in the data, as directly as possible. Predictions and actual values should be plotted against time, where possible, and the differences explained. This provides a basis for going back and changing the model (or better understanding its weaknesses). Plots like these are important both in estimating a model and in explaining the model to others. Tukey of Princeton has written a book on Exploratory Data Analysis, describing additional techniques for better understanding the residuals graphically.

In the past, some econometricians have routinely used "dummy variables" (1 in some years and 0 in others) or other procedures to throw out "outliers," observations which are hard to explain using their forecasting model. (Some statisticians have also recommended maximizing the

1.5 power of error instead of the square error, which has much the same effect.) More recent authors, like Belsley, Kuh and Welsch, have stressed to need to study the outliers (and other "influential observations") rather than simply throw them out, because they may be crucial to what your model is trying to forecast and may be important as a guide to a better model. For example, the oil shortages of 1974 and 1979 were "outliers." but a model which ignores them is a poor guide to reality. Again, the hard-to-explain observations should be obvious in a plot of predictions and actuals, but more sophisticated tools exist for identifying them.

In practice, we have also found that there is no substitute for performing a "dynamic simulation" test of a model, if you are considering the use of a model which contains a lagged endogenous variable. This is quite different from evaluating the "predicted values" which come out of a standard regression command. For example, if you were estimating equation 1 over the period from 1967 to 1985, the "predicted population" for 1980 would be calculated as "c" multiplied by the actual population in 1979, in a regression package. In dynamic simulation, the prediction for 1980 is calculated as "c" multiplied by the prediction for 1979, which in turn is calculated as "c" times the prediction for 1978, and so on. The regression test would be appropriate, in theory, if predictions one year ahead were all that you care about. The dynamic simulation test would be better if you planned to forecast further out into the future, or if the real concern for policy is the eventual result several years into the future.

A purist would argue very strongly that the regression test is adequate, if one has faith in the truth of one's model. He would argue that those lacking in faith should look for better models. A utilitarian would argue that all models are oversimplifications, and that faith without tests is no way to do modeling. Experience has shown that cumulative error tends to be quite important (or even overwhelming) in models containing lagged endogenous variables, regardless of their theoretical virtues. More to the point, it has shown that such errors can be avoided either by specifications without lagged endogenous variables (if such can be found, with otherwise comparable MSE scores) or by a new form of robust estimation.

An example of the former comes from our PURHAPS model: by filtering the effect of energy prices over time, we can represent the notion of capital-embodied price responses just as effectively as do recent academic models based on lagged dependent variables (previous period energy intensity); however, our older version may be more robust. On the other hand, there are many models (especially models which assess changes induced by policy) which have much higher error levels and much crazier parameter estimates when lagged endogenous variables are not included.

To estimate models with lagged endogenous variables, several authors -- including Larry Klein of Wharton, and myself (Werbos, 1974) independently -- suggested several years ago that models could be estimated by directly picking parameters so as to minimize errors in dynamic simulation. This could be implemented in practice by doing the dynamic simulations on a PC package like Lotus, and adjusting the parameters by hand to minimize the error in dynamic simulation. This form of robust estimation may be somewhat extreme; however, we have found (Werbos and Titus, 1978; and Chapter 4 of Werbos, 1983) that it is possible to compromise between this approach and regression, and still allow for a noise term in the model (allowing for uncertainty). Dynamic robust estimation methods of this sort have cut errors in half in a number of applications (where lagged endogenous variables were important), and have even done better in short-term forecasting (Werbos, 1974; Werbos and Titus, 1978). The compromise method looks similar to "exponential smoothing" methods which are essentially equivalent to the Box-Jenkins methods discussed above; however, they weight the square error in different ways, and this difference in weighing will probably be the key feature even of more advanced methods along the same lines.

When calculating error for a system of equations, one may simply use a weighted sum of the error for different variables (including "filtered" variables). A utilitarian would argue strongly for doing this, and for weighing each dependent variable's error according to the importance of that

variable in a larger context; he would not use the classic form of full information likelihood, which is very volatile and subject to singularity problems.

6. MODEL DEVELOPMENT IN PRACTICE

Econometric methods are used for many purposes in government and industry. Many people use them for causal analysis, their original purpose as described above.

In my division of the Energy Information Administration (EIA), we use these methods to build up models of energy demand by sector in the United States economy. These models are then used as part of a larger modeling system (IFFS) which currently projects energy supply, demand, and prices by year from 1985 to 1995. The base case projections are published in the Annual Energy Outlook, along with a few sensitivity cases. IFFS is also used in a variety of special studies which come up every year; for example, it may be used to predict energy demand and supply with and without an oil import tax. EIA, like Wharton Econometric Forecasting Associates and many other forecasting organizations, maintains both an annual and quarterly forecasting model. The problems of publishing annual forecasts are similar in all such organizations, where quality is a concern, and will be discussed here in general terms.

EIA has never used econometric models exclusively to generate forecasts. The goal is always to make forecasts which represent our best guess about the future, conditional upon various assumptions about the GNP or world oil prices. The GNP forecast is taken from an economic forecaster, such as Data resources, Inc. The world oil price is projected by another division of EIA, and alternative scenarios are developed to reflect the uncertainties in this projection. Our goal is to report the best guess we can, accounting for all sources of information, including historical data, engineering ada, the trade press, etc. In some sectors, we begin with econometric models, and in others we begin with engineering process models; in either case, we try to inform or calibrate the model, accounting for information from other sources.

In developing a new model, one rarely starts from scratch. There is usually an existing model that was used in the previous year to forecast the same general concept. Generally speaking, we go through six stages (when we have the resources to do things correctly):

- o preliminary evaluation of the existing model(s)
- o detailed literature review
- o assessment of sources of uncertainty and how to reduce them
- o data acquisition, econometric analysis, and the like
- o coding and testing
- o model maintenance

Preliminary Evaluation

The preliminary evaluation usually starts from the annual review of model forecasts which occurs as part of the annual forecasting cycle. Most reviewers have certain expectations about what the forecast should look like, based on a variety of sources. If a forecast is much different from these expectations, it is scrutinized further. (Unfortunately, it is more difficult to identify poor forecasts adjusted by brute force to match common expectations; however, when there are competing viewpoints represented in the process, based on deeply held professional or political orientations, any forecast may be different from what someone expects. It is common to be told within the same week, by different people, that a given elasticity is "absurdly high" and "absurdly

low" both.) "Industry" is widely cited in this process, but in a haphazard way and in different directions.

Next, the analyst tries to explain why the forecast differs from expectations. Often, this explanation involves a plot of historical data, showing how the assumptions of the model perform compared with more popular assumptions. Simple plots and growth-rate calculations are vital to these explanations, which take up a major part of the analyst's time; translations back and forth between English, statistics, and simple plots or tables are central to the process.

A critical advantage of econometric methods at this point is that they do have an historical track record and a corresponding data base. Also, they may be simpler to understand and explain than the process models; they do not contain large files of assumptions which can carry hidden biases. Process models, on the other hand, can simply be adjusted to match the prior expectations. In the past, process models were more popular than they are now, because they could show large impacts from proposed regulations such as the Powerplant and Industrial Fuel Use Act. Models such as the Project Independence Evaluation System have come under great criticism for their optimism about oil imports, even though the authors worked hard to warn the reader in the document that the conditions for zero imports by 1985 (as requested by the policy-makers) would be very difficult to meet.

If the explanation for a forecast does not have enough information behind it, or is not communicated properly, adjustments are made, and the issue will be studied in more detail after the current forecasts are produced. (Econometric models can be adjusted and rerun at least as easily as process models, but the adjustments tend to be more visible.) Likewise, questionable or uncertain forecasts are usually revisited even if they are not initially modified. At times, a detailed review from the Quality Assurance Division may spark a preliminary evaluation. If the preliminary review of a model then suggests a major problem, the problem goes to the top of the priority list.

A great difficulty in this process is to sort out the difference between common expectations, political or other prejudices, and objective reality. When there are many layers of analysis between the raw data and the final publication, there is often a danger that the likelihood function will be multiplied many times by the same prior probability distribution, therefore biasing the publication a bit too heavily towards the priors; there is no cure for this problem other than improved communication.

In general, this process tends to be very instructive, but the details are rarely published; they tend to be viewed as too technical.

Detailed Literature Review

When there are major problems with an existing mode, the model may or may not be rewritten; that depends on the nature of the problem. The first step in reviewing the problem is a thorough review of alternative models and forecasts, and the information behind them. Uncertainty assessment is the next stage, but the literature review itself tries to find complete information for that next stage; thus, an informal uncertainty analysis is always being done.

The literature review requires an evaluation of the statistical methods used in the existing models. It requires an effort to understand what is going on, substantively, in the sector being modeled. It requires a great deal of thought about how different phenomena, in the real world, would influence different kinds of statistical analysis or model in different ways. (Among the important phenomena are things which bias different data collection efforts.) It requires skepticism and a search for evidence when judging the statements of other modelers and of "substantive experts" both. It requires a search for possible biases, as with models which make strong assumptions about the costs of new technologies. It requires a search for widely opposing

points of view. Above all, it requires an effort to understand why different forecasts of the same variables from different sources come out differently.

After the literature review is largely complete, it is usually important to write a brief memo on the issues which have emerged, and to get feedback from other analysts.

Uncertainty Assessment

The most critical stage in modeling, most often neglected, is the analysis of first-order uncertainty. First-order uncertainties are those things which explain most of the differences between the existing models (or potential models). It is sad how often government agencies have spent millions of dollars on detailed, highly precise models, without checking to be sure that the gross, first-order sources of error have been thoroughly understood.

A key part of the first-order analysis is an approach to explaining the basic, first-order trends in past history. This includes a plan to resolve or reduce the uncertainties by statistical analysis or by the most reliable, empirically-based methods available. After the first-order uncertainty has been resolved, one can move on to the second-order uncertainty, and so on; at all times, no complexity needs to be added (on the conceptual level, anyway) unless it really helps reduce the basic uncertainties.

With behavioral effects, such as price elasticities, energy demand, or fuel-switching sensitivities, the historical data provide a good way to estimate model parameters; econometric techniques are appropriate. With engineering variables, such as synfuels costs, we have supported statistical cost-calibration analysis by Ed Merrow at the RAND Corporation; early studies, taking engineering estimates at face value, had been too optimistic. The key question about technology is how much changes in its rate of development are likely to change trends or price responses in the future; often, these changes introduce uncertainties both on the upside and the downside, uncertainties which it is difficult to resolve realistically. Changes in markets are often at least as important as changes in technologies as such.

In some cases, as with industry, the technological change has been so complex and so continuous (on average) that we use an econometric approach almost completely (except with petrochemicals). With transportation, a hybrid approach is used because of the unique role of new car miles per gallon (mpg). In the residential and commercial sectors, a key problem (not yet resolved) is to reconcile the conflicting studies which claim that conservation has mostly been due to lower thermometers (which won't continue when prices stabilize) or due to structural improvements in new buildings (which will continue as housing stocks roll over).

Uncertainty assessment cannot be done unless you know which forecast you are trying to evaluate. At EIA, our main concern is to get the national base-case forecasts right, and to get an accurate response to moderate price changes and economic growth. When new policy issues arise, a new assessment is needed for the new issue; at times, this may force the development of a new model. For example, a few years ago, when issues about contracts were considered fundamental to natural gas regulation issues, a new model was developed which specializes in that area; now, however, the empirical issues of base case accuracy are beginning to get more priority again.

According to industry representatives at the Energy Modeling Forum, private corporations often have a special need for State-level forecasts or the like. Unexpected national trends probably invalidate the State-level forecasts as often as any other problem does; likewise, behavioral assumptions at the State level may be a problem. For this reason, insights developed in national analysis may be useful to private industry. However, the forecasts themselves have rarely been subject to uncertainty analysis at the State level, even when State-level data are available; therefore, they are not to be taken at face value. EIA does publish regional forecasts,

in special service reports; however, these forecasts are produced mainly because they are important to the accuracy of the national forecasts. The regional forecasts are checked, but are not evaluated in depth.

Analysis

The work at this stage should follow directly from the analysis of uncertainty. Econometric techniques are used at this stage more and more at EIA, in part because we have a comparative advantage in using data and in part because of cost considerations. (A thorough and conclusive analysis of even one technology is not cheap.)

Looking at models as boxes, we often ask what kind of real knowledge or information is really "in" the box; our econometric models contain primary information about what can be learned from history, based on the analysis stage. Because of the adjustment process, they contain a bit of secondary information that we borrow from other sources about markets and technologies. Good process models contain primary information about technologies, developed at the analysis stage; however, they are adjusted to match historical reality, and may be secondary users of econometric information. Some hybrid models or global models tend to be secondary users of both kinds of information.

Coding and Testing

This is straightforward conceptually; I wish it were so easy to do. At EIA, we code our models in FORTRAN, because even our econometric models are too complex for basic systems like Troll to handle at present. The statistical analysis is mostly done in SAS, because SAS is flexible and easy to use with complicated data bases; the DATA and PROC MATRIX components of SAS are especially powerful.

Model Maintenance

The uncertainty analysis and review described above continue for all models, even after they are established, on a regular cycle.

General Observations

After this exercise is completed, it often seems that the insights gained in modeling are more important to policy than are the forecasts themselves, if the job was done right. Good Bayesians know that the future is highly uncertain, and that any policy which is based on a definite expectation of a "baseline future" is likely to be a poor policy. People at EIA and elsewhere vary greatly in their willingness to describe and explain the full range of uncertainties, as opposed to defending the base case or presenting a small number of conservative sensitivity cases in line with popular expectations.

The steps needed to get good forecasts are also needed to refine one's understanding of cause and effect, as I pointed out in the discussion of maximum likelihood theory above. When cause-and-effect analysis is done outside of the forecasting context, it is often based on quick, casual statistical models or analysis by eyeball; it often derives incorrect conclusions because it fails to account for issues we learn about in forecasting. Unexpected futures are also important to cause-and-effect analysis, as they are to forecasting. It is a continuous challenge for modelers to find ways to exchange such insights and make them fully available to decision-makers.

Of EIA's model based reports, the Short-Term Energy Outlook is often cited as the most respected. These quarterly forecasts are based on simpler models, and are more completely econometric than the annual forecasts are; the model assumptions and historical track record have been communicated effectively in recent years, and there has been feedback from high

levels of the government. The communication is more difficult for the annual model, for at least three reasons:

- o the annual model contains major process-model components (mainly on the supply side)
- o the published year-ahead forecasts are not a good basis for testing track records, because they are calibrated to the quarterly forecasts.
- o the annual model makes use of more detailed data sources, some of which have been discontinued, changed in scope, or collected at irregular intervals; while the models account for this, it is not possible to put all the relevant information on one neat plot that goes up to 1985

A key problem in comparing published long-term forecasts over time is to control for GNP assumptions and world oil price assumptions. For example, optimistic assumptions about economic growth, taken from other government components, can push up energy demand forecasts; energy use in industry tends to be very volatile with respect to the rate of growth, and few government forecasts ever show major recessions. EIA has published plots showing the track record of our model in predicting world oil prices as a function of OPEC capacity utilization; the track record has been surprisingly good, but earlier forecasts of OPEC capacity utilization led to major errors, due in part to economic assumptions.

7. EXAMPLE

The Purchased Heat and Power System (PURHAPS) is one of the models I have developed at EIA, following the approach described above. The core of PURHAPS is a system to predict purchased fuels and electricity in manufacturing in the United States, excluding refineries. This represents only half the energy covered in our industrial sector, but there is not enough space here to discuss all the rest. This part of PURHAPS contains good examples of the issues discussed above. For other -- simpler -- examples of these issues, see also our Transportation Energy Demand (TED) model (Werbos, 1986b) (which should be adjusted to reflect more recent analysis of alternative fuels (Werbos, 1987c).

EIA management decided it needed a new model of industrial energy use early in 1982, after several analysts in several divisions agreed that the existing model was unsatisfactory. The existing model, a version of the Oak Ridge Industrial Model (ORIM), was driven by "econometric equations of total energy use which had not been derived from empirical data, and which could not be estimated very easily; the fuel-switching was driven by a "process representation" which did not reflect differences between the different 2-digit industries, differences which seemed critical in our reading of the literature. ISTUM2 did have an adequate process representation, and was seriously considered for use; however, the cost of calibrating all the technology data, and doing the historical calibration we would require for our purposes, was enough to rule it out.

The decision was made to start out, at least, with an econometric model, and add engineering detail only when and as there was reason to do so. I proposed that we first get the basics right, and then consider more elaborate models of primary metals, chemicals, and paper after this point is reached. In later years, the possibility of more elaborate industry models was considered but never seemed to promise much reduction of uncertainty at acceptable cost; however, the system is still considering a suggestion that we buy data from SRI International which might allow a better treatment of petrochemical feedstocks, for which the usual data appear contradictory and problematic. (Some analysts have recommended that we model this sector equating ethylene with petrochemicals, but the complexity and flux of this industry suggests that this would not buy us anything better than the simple, aggregate approach we now use.)

Our initial analysis goal was two-fold:

- To get total energy use right, at the national level, we wanted a model to track energy use in the individual industries (for which good historical data were available back to 1958) as well as possible.

- To get fuel-switching right, we wanted a model to represent State-level tradeoffs within specific industries. Other EIA demand models do not carry State-level detail, but it was felt that fuel price tradeoffs vary a great deal from State to State in industry; also, the availability of State-level data from the Annual Survey of Manufacturers was critical.

Forecasting Total Energy By Industry

Based on our initial literature review, it was felt that Dale Jorgenson's work with Barbara Fraumeni was the best starting place for the national energy equations. (Other work by Ernie Berndt at MIT appeared very promising, but too difficult to update in time for the next Annual Energy Outlook.) Dale Jorgenson was given a small contract to update his database on the price and quantity of capital, labor, energy, and materials used in industry through to 1979, and to reestimate the equations he had published.

As part of our normal review process, we tested the resulting equations with Jorgenson's own data to see how well they would have performed in the past. The results were a great disappointment to myself and to Jorgenson; we have published a curve or two of the errors, but have otherwise tried to move on to something better. Some people suggested that the Jorgenson effort had been ill-conceived; however, the purpose of scientific research is to discover new information, as we had, and an updated database was needed in any case.

The original Jorgenson equations had been estimated as a complete system, using a system of constraints and estimation methods related to 3-stage least squares. We tried to reestimate Jorgenson's energy model, using only the energy equation and multiple regression; the errors were less, and some people wanted to go ahead and use the results, but the price elasticities did not make sense over time and the problem turned out to be unavoidable with the "translog" model being used.

We then estimated a similar equation, using the same data base and still controlling for the effects of capital prices, etc., but assuming a constant elasticity of energy prices on energy demand per unit of production. This worked better, and made much more sense in historical plots. It also showed much smaller price elasticities than the theoretical economists had been publishing in recent years (based largely on the translog or on data from 1958-74).

In the following year, we studied the residuals more carefully, and found some signs of bias. We explained this bias as a result of "capital-embodied price responses," which represent the effect of slow capital turnover on energy conservation; this explanation led to a much better fit, including average errors of about 1% in all the major industries in the postembargo period. The equation has no lagged endogenous variable to allow the possibility of cumulative error in forecasting several years out. (The new Transportation Energy Demand (TED) model has slightly larger errors -- 1.8% -- in forecasting total personal travel; however, because these errors are visible in a single plot of actual travel versus predictions from 1954 to 1983, the personal travel model is probably viewed as a bit more reliable.)

Forecasting Fuel Shares

Our analysis of fuel shares also produced a few surprises.

Coal use changed so slowly over time that we needed a nonlinear model (requiring nonlinear regression) to track it at all. Coal use turned out to be complementary to natural gas, to a very

slight degree; at first, we thought that this was an artifact of our equations, but we found out that the historical fits got noticeably worse when we "corrected" the specification to cut out the complementarity.

With electricity use, we initially used a lagged endogenous variable to represent long-term price effects. However, we discovered that the parameter estimates were very unstable when we changed the equations slightly, because e of "robustness" problems; we then settled on something a little simpler.

With oil-gas switching and with electricity-fossil shifts, we predicted changes in shares as a function of changes in prices from one year to the next. Initially, we tried a simpler version with oil-gas switching, predicting shares as a function of prices, but a comparison of the two approaches verified our suspicion that the simpler version would be biased. States initially varied in fuel shares in 1974 because of differences in industrial mix which depend on other factors; an equation based on changes assumes that different States may have different fuel shares even when prices are equal, and it does lead to much smaller errors. (The MSE has been published for both versions, and is several times larger for the simpler version.)

With oil-gas switching and electricity shares both, we were worried that trends might have changed over time, or that long-term price effects might be larger than short-term price effects (i.e., that the simple specifications might have robustness problems). To test this, we tried specifications projecting changes in shares from 1974 and 1978 to year t as a function of price changes over the same period; little change was found. To test the trends, we used "dummy" variables (each set to 1 for a particular year and 0 in all others) in place of the constant term in these equations; we found little noticeable effect with gas, but we did decide to use the average trend over the last four years for electricity shares.

Sources of Further Information

Even the important details of this story would fill up too much of this book. For a concise description of the model, please consult Volume I of the PURHAPS documentation (Werbos, 1984). Volume III (Werbos, 1983) gives the story and the rationale behind the initial version of the model in much more detail (one chapter per equation). The nonmanufacturing portion of the model has changed since then. More recent information (including updates and adjustments) is given in a number of papers, mostly focusing on the substantive meaning of our forecasts (Werbos, 1987b; AEO, 1985; Werbos, 1987d).

Structure of the Manufacturing Model

The basic structure of PURHAPS is shown in Figure 1. For each of 17 manufacturing industries, 7 econometric equations were estimated, corresponding to the 7 in the chart. Each of the State-level equations is applied 51 different times in ever forecast year, once for each State; the results differ across States simply because the fuel prices, industrial production, and previous-year conditions differ across States. The State forecasts for total energy use are not simply "shared out" from the national total; in each industry, the national equation and the 51 State equations form a system of 52 simultaneous equations, each constituting a well-thought-out model, which are "solved" as a system.

The original Jorgenson model was a "KLEM" model, where "K" represents capital, "L" labor, "E" energy, and "M" intermediate materials. PURHAPS may be called a "KIM/LEO" model, because capital, materials, and time (I) are accounted for a national level, while labor, energy and output are also used at the state level.

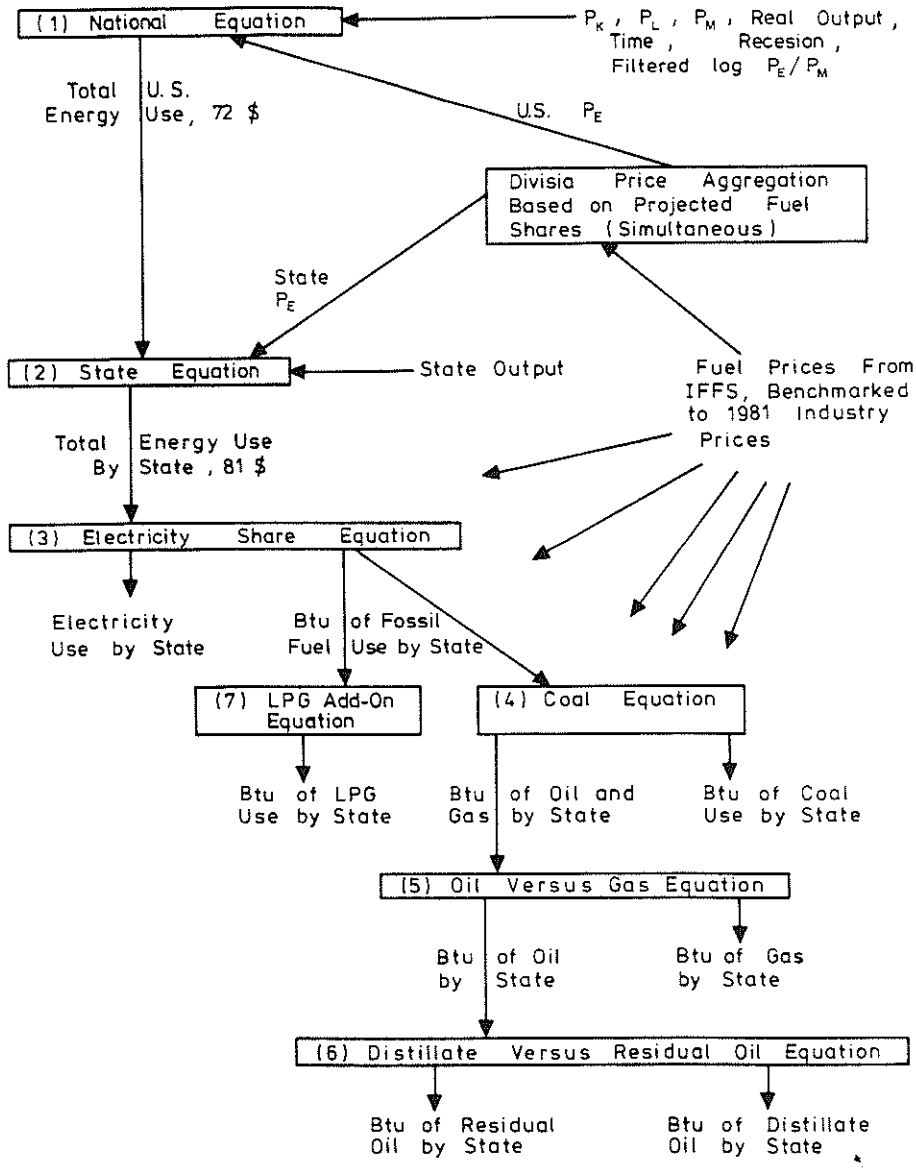


Figure 1. Structure of the Manufacturing Model

REFERENCES

Annual Energy Outlook (1985). DOE/EIA-0383(84), NEIC.

Box, G. E. P. and G. M. Jenkins (1970). Time-Series Analysis: Forecasting and Control, Holden-Day: San Francisco, Calif.

Brown, R. et al. (1974). Decision Analysis: An Overview, Holt, Rinehart and Winston: New York.

Dempster, A. et al. (1977). "A Simulation Study of Alternatives to Ordinary Least Squares", Journal of the American Statistical Association.

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

Industrial Fuel Switching (1986). DOE/EIA-0504, NEIC: Washington D.C.

Pindyck, R. S. and D. L. Rubinfeld (1976). Econometric Models and Economic Forecasting, McGraw-Hill: New York.

SAS Users Guide: Basics (1985a). SAS Institute: Cary, N.C.

SAS Users Guide: Statistics (1985b). SAS Institute: Cary, N.C.

Troll Users Guide (1981). MIT Center for Computational Economics and Statistics: Cambridge, Mass.

Werbos, P. (1974). "Beyond Regression: New Tools for Prediction and Analysis in the Behavioral Sciences," Ph.D. Dissertation, Harvard University, Cambridge, Mass.

Werbos, P. (1977). "Advanced Forecasting Methods For Global Crisis Warning and Models of Intelligence", General Systems Yearbook.

Werbos, P. (1979). "Changes in Global Policy Analysis Procedures Suggested by New Methods of Optimization", Policy Analysis and Information Systems.

Werbos, P. (1983). A Statistical Analysis of What Drives Industrial Demand. DOE/EIA-0420/3. National Energy Information Center (NEIC, 202-586-8800): EIA, Washington D.C. (Should be available at no charge, or contact the author at 202-586-1197.)

Werbos, P. (1984). Documentation of the PURHAPS Industrial Demand Model, Volume I. DOE/EIA-0420/1, NEIC: Washington D.C.

Werbos, P. (1986a). "Generalized Information Requirements of Intelligent Decision-Making System" in SUGI 11 Proceedings. SAS Institute: Cary, N.C. (A revised version, available from the author, is easier to read and contains more discussion of psychology.)

Werbos, P. (1986b). Documentation of the Transportation Energy Demand (TED) Model, DOE/EIA-MO13, NEIC: Washington D.C.

Werbos, P. (1987a). "Building and Understanding Adaptive Systems: A Statistical/Numerical Approach to Factory Automation and Brain Research", EIII Transactions on Systems, Man and Cybernetics.

Werbos, P. (1987b). "EMF 8: A Modeler's Perspective" in The Changing World Energy Economy, International Association of Energy Economists: Washington D.C.

Werbos, P. (1987c). Oil Dependency and the Potential for Fuel Cell Vehicles, SAE 871091, Society for Automotive Engineers: 400 Commonwealth Drive, Warrendale, Pa. 15096. (On request from author.)

Werbos, P. (1987d). "Industrial Structural Shift: Causes and Consequences for Electricity Demand", in A. Faruqui and J. Broehl, (eds.) The Changing Structure of American Industry and Energy Use Patterns, Battelle Press: Columbus, Ohio.

Werbos, P. (1988). "Backpropagation: Past and Future", Proceedings of the Second International Conference on Neural Networks. IEEE: New York.

Werbos, P. (1989). "Maximizing Long-Term Gas Industry Profits in Two Minutes In Lotus Using Neural Network Methods", IEEE Transactions SMC.

Werbos, P. and J. Titus (1978). "An Empirical Test of New Forecasting Methods Derived From a Theory of Intelligence: The Prediction of Conflict in Latin America", IEEE Transactions on Systems, Man and Cybernetics.

Wonnacott, T. H. and R. J. Wonnacott (1977). Introductory Statistics for Business and Economics, Wiley: New York.

Energy Vol. 15, No. 3/4, pp. 237–248, 1990
Printed in Great Britain

0360-5442/90 $3.00 + 0.00
Pergamon Press plc

2.2. PROCESS MODELS

WALTER C. LABYS

Department of Resource Economics, West Virginia University
Morgantown, West Virginia
U.S.A.

and

Universite d'Aix-Marseille
Marseille
FRANCE

HIROSHI ASANO

Central Research Institute of Electric Power Industry (CRIEPI)
Tokyo
JAPAN

Abstract - Process models represent a particular class of engineering-economic models . Their purpose is to analyze commodity and transformation processes and the economic structures that control and influence these procedures. The theory of these models is introduced, followed by an example from the industry. Other applications also are reviewed.

Process models concentrate on the transformation of fuels and other primary inputs into intermediate or final products. When used for demand determination, they explain the demand for primary products as they flow through a transformation or refining process. When used for supply determination, they explain the supply of primary products as they undergo typical preparation processes. The particular engineering characteristic of these models has also caused them to sometimes be labeled as technico-economic models. They often attempt to couple technical aspects of engineering systems with economic variables. This approach is particularly useful for modeling industrial process which normally involves different stages of process or production, e.g., see Kovisars (1976). Using the mineral industry as an example, these stages can include: mining, ore treatment (milling and concentration), reduction (smelting), purification (refining), and consumption by fabricators. Recycled material is an important process input for many mineral flows and may enter the supply flow at several stages. While Manne and Markowitz (1963) describe the early development of these models as an application of process analysis, they also report on the integration of this approach with spatial equilibrium analysis, e.g. see Fox (1963) and Marschak (1963). Not much has been reported on more recent developments of this methodology until the appearance of a survey article by Sparrow and Soyster (1980).

The actual formulation of a process model is typically based on linear programming but other forms of mathematical programming such as integer modeling can be used. What is important is that an intermediate step is included to better approximate the production function. The model also requires the usual objective function, activities and constraints. Model solution usually involves minimizing an objective function specified in terms of production costs, subject

to such constraints as the time sequence of production, plant capacities, the demand for the product, and technical relationships among the production variables.

An activity or process component can be considered the distinguishing feature of a process model. It can be considered as an intermediate step that describes how decision variables (resources in the production problem) are combined in fixed proportions by the production technologies to produce an output. This emphasis makes process models some ways similar to the input-output approach. But instead of a single technology for every process, programming methods can allow for multiple processes or activities. These activities define technologically possible alternatives in physical terms (involving different energy, material and labor requirements) and do not necessarily yield economically (or socially) efficient solutions. Economic considerations are introduced by the cost function, which shows the minimum cost of producing various levels of output, given factor prices and technologies.

Process models are also suited for conducting engineering-economic analysis of commodity supply. Such models are usually formulated in terms of production technologies, accounting for the important inputs (labor, energy, and materials requirements) at each stage of the production sequence.

This requires constructing mathematical equations that represent the various technological or engineering production possibilities. Usually the overall production flow is disaggregated into elementary process routes, and input-output parameters are derived for each stage of each route, based upon engineering data. For each process, programming techniques can be used to select from among the various production possibilities the ones that optimize a preset goal.

Process model can also be used for commodity demand analysis by permitting a linkage to be made between final product demands and derived material demands. National economic activity can be used to explain final product demand. The process model then explains the product transformation process, that is, how primary products derived from materials can be transformed into secondary and then tertiary products (or greater) until final product demand is met. Sometimes the supply of material is also studied as part of the demand process.

Perhaps the best way to explain this variety of applications is with an example drawn from the process models for assessing industrial response to time-of-use(TOU) rates, developed by Asano, Sagai, and Yamaji(1988c).

A process model which describes customers' actions cope with given electricity rates through the explicit modeling of engineering characteristics of electricity use would be a useful tool in microscopic evaluation of customers' responses to electricity prices. EPRI (Electric Power Research Institute) has developed simulation models for cement and chlor-alkari plants (1985).

For the customers whose load characteristics are strongly influenced by the technical structure of electricity utilization, microscopic engineering-economic analysis is appropriate to evaluate the responses to TOU rates.

We took an integrated steel mill for the first case study of a typical industrial customer with intensive electricity consumption. With the application of Linear Programming technique, the optimal weekly operation schedule was derived and its effects on the customer's operating costs and the utility's load profile were analyzed (Asano, et al., 1988b).

We have chosen an electrical apparatus works for the second case study, since the electrical machinery industry makes a significant component of utilities' electricity loads in Japan. With several heuristic methods to find the quasi-optimal monthly operation schedules, approximate values of the potential savings were evaluated (Asano,et al., 1988a,b).

A CASE STUDY FOR THE IRON AND STEEL INDUSTRY

Model Overview

Some portion of electricity is supplied by the industrially owned generation plants and co-operative generation companies which are in turn jointly owned by steel makers and electric utilities in the region concerned. Except the energy contained in the coke for blast furnace(BF)s, other major forms of consumed energy are gases used for heating and electricity used for motive power. By-product gases recovered from coke plants (COG), blast furnaces (BFG), and basic oxygen furnaces (LDG) are used to meet these gas demands and to generate electricity. Electricity is also recovered by the generation facilities owned by the Works including top pressure recovery turbine(TRT) attached to BF. The share of self-supplied electricity in total electricity demand, which mainly depends on the supply/demand condition of by-product gases, varies widely among steel mills in Japan.

Two typical mills with high and low electricity self-supply shares (called Works A and B, respectively) are taken for case studies. The manufacturing processes of Works A are shown in Figure 1. There exist no wide flange beams, wire rods, UO pipe,and spiral pipe plants in Works B. Works A has a cooperative power company, which burns by-product gases and supplies electricity back to Works A. Works B has a self-generating plant which burns purchased oil.

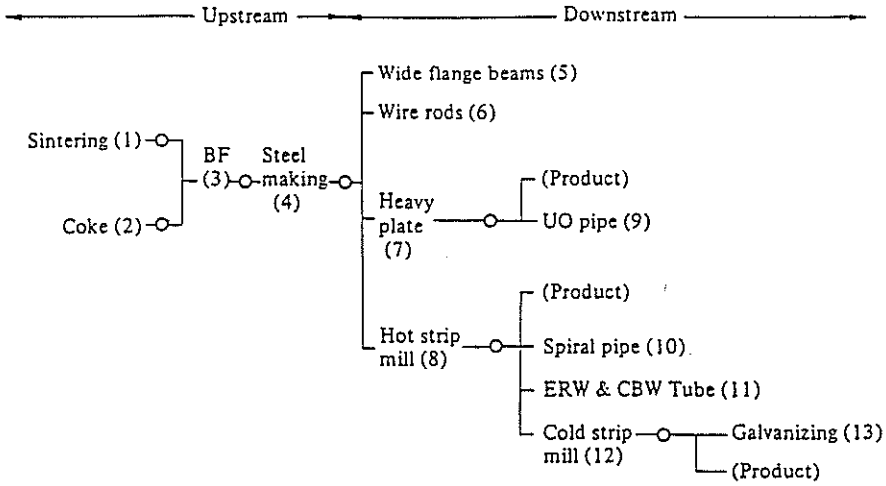


Figure 1. Manufacturing Process of an Integrated Steel Mill
(0 indicates buffer storage)

Note: ERW & CBW: electric - resistance - welded & continuous butt - welded.

Weekly demands for final products are given exogenously. An operation schedule of each process is determined to minimize the total weekly production cost. A day is divided into three eight-hour periods and a weekly operation schedule consists of 21 periods. Material flows are balanced for each period considering buffer stocks between the processes.

Energy flows are modeled to represent supply/demand of electricity and by-product gases for each period. Three types of by-product gases are treated separately and gas holders are modeled for each type of gas to describe the effect as buffer storage.

Figure 2 shows an overview of the process model for the integrated steel mill. All equations including the relations between operating levels and input factor demands(i.e.,energy demands

and labor requirements) are described in linear form so that the Linear Programming technique is applied to minimize the total cost.

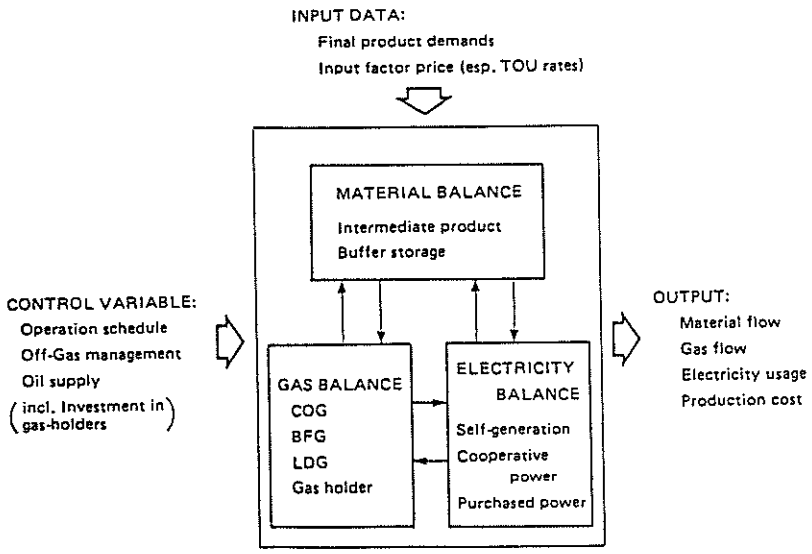


Figure 2. Overview of Process Model for the Iron and Steel Industry

In our formulation, control variables are operation schedules of each process and management of by-product gases in each period. We assume that terminal values of stock levels for buffer storages and gas holders should be equal to the respective initial values which are to be optimized.

The following types of constraints are considered: (1) material balance constraints with buffer storages, (2) electricity balance constraints, (3) gas balance constraints, and (4) miscellaneous operational constraints. Operational constraints mean that BF and steelmaking processes are always operated at a constant utilization rate, and that all processes have equal operation pattern on Tuesday, Wednesday, and Thursday. The objective function is short-run production costs including labor and energy costs.

The following parameters are given exogenously:(1)process capacities, (2)TOU rate structure, (3)wage rates, (4)gas holder capacities(COG:1910 Gcal, BFG:195 Gcal, LDG:178 Gcal as a reference case), (5)material buffer storage capacities, (6)sales price of by-product gas to the cooperative power, (7)electricity and gas consumption intensity of each process, and (8)labor requirements for each process.

Data have been prepared in close cooperation with the staff of the Works. An on-peak period is defined during 7:00-23:00(the first and second shifts) on weekdays and during 7:00-15:00(the first shift) on Saturday. The revenue neutral TOU rate for this customer is used in the simulations. With this assumption, the total electricity costs would remain unchanged if the customer does not respond to TOU rates. The base rate is a non-TOU type rate with a 12.65 yen per kWh. We set Wd2=1.05Wd1, Wd3=1.32Wd1, We1=1.10Wd1, We2=1.15Wd1, and We3=1.43Wd1, where W_{ij} stands for the hourly wage rate during the j-th shift on weekday(when i=d) or weekend(when i=e), (e.g. Wd1 is the wage rate during the first shift on weekdays).

Response to Time-of-Use Rates: Optimal Weekly Operation Schedule

Figure 3 shows consumed electricity consumption purchased from the electric utility under the TOU rate scenarios. Figure 3(a) shows that for the Works A, the pattern of purchased kWh distribution between on-peak and off-peak periods is very sensitive to the ratio of on-peak to

off-peak energy charge. As the ratio is raised to 1.014, Works A increases off-peak power and the resulting on-peak kWh share declines to zero. The total purchased electricity is constant and is about seven percent of the total electricity demand. With the optimal gas holder management, the conflict between the labor cost minimization and the minimization of power cost is resolved by achieving daytime oriented operation as well as confining the demand for purchased power in the nighttime.

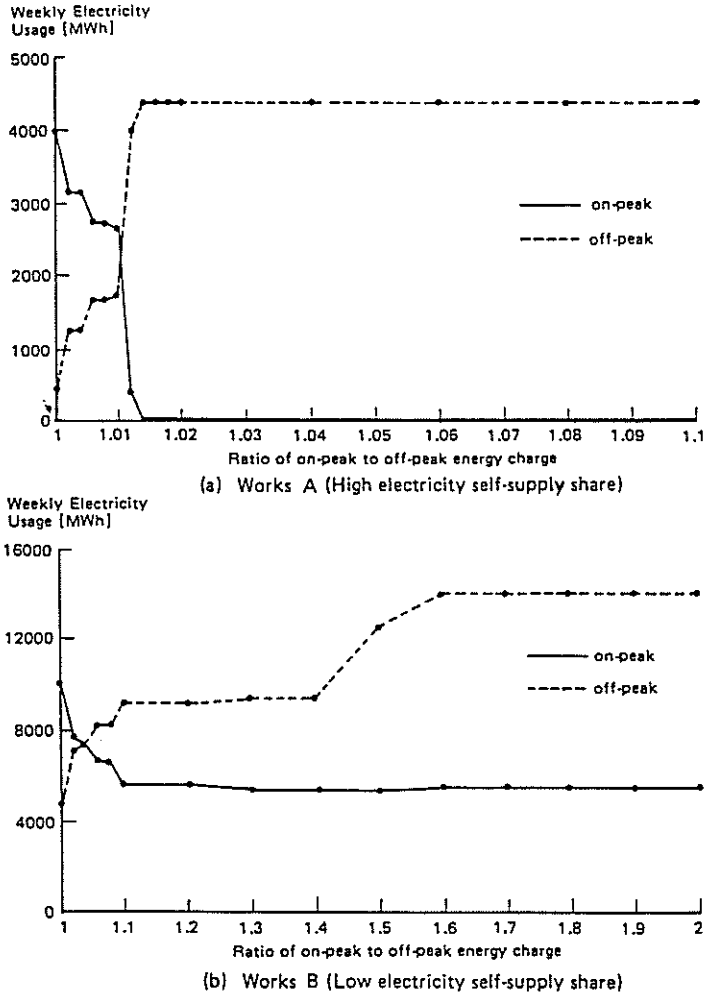


Figure 3. Effects of TOU Rates on Purchased Electricity Usage

As shown in Figure 3(b), sensitivity of Works B is lower than that of Works A. The total amount of purchased electricity in creases as the ratio of energy charges is raised over 1.4, saturating at 1.6. This increase occurs because off-peak power supplied by the utility becomes more economical than self-generation.

For Works A, as the capacity of the gas holder is considered to be critical in responses to rate structure, sensitivity analysis is performed to examine the effect of the gas holder capacity. Reductions in the total operating costs are evaluated for the 30 percent increase of the capacity of COG gas holder. As a result, the condition for TOU rate structure was obtained so that the marginal benefit in operating cost balances the marginal capacity cost of the gas holder.

A CASE STUDY FOR A HEAVY ELECTRICAL APPARATUS WORKS

Model Overview

The electrical machinery industry makes a significant component of utilities' electrical loads in Japan. A heavy electrical apparatus works was taken for the second case study. For our modeling of the production processes of heavy electrical apparatus, we used the characteristics of a works that produces turbines and turbine plant equipment, generators, and motors. We developed a model to describe the production processes of turbine generators, since large scale turbine generators are major products of the Works and some processes of a large scale turbine generator may consume about 30 percent of total electric power of the works.

Processes for rotor balance tests and final electrical tests consume a large amount of electric power, and the number of workers involved in these processes is relatively small. Thus the rescheduling of these processes to shift activities to the period with low power cost is expected to be economically justified. In our illustrative problem, six products are parallel tested in a month. Activity sets(or job sets) for each product are defined.

Formulation

Figure 4 presents an overview of the structure of a process model. In our formulation, control variables are binary decisions whether to do or not to do each job in each period. We consider the following types of constraints: (1)processing order and due date of each product, (2) limits for machine and manpower loading, (3) contracted electricity demand. Operation schedule of each activity is determined through simulations to reduce the total monthly operating cost including electricity cost and labor costs. A day is divided into four unequal length periods; 8:00-12:30, 12:30-17:00, 17:00-22:00, and 22:00-8:00. An on-peak period is defined during 8:00-22:00 from Monday to Friday.

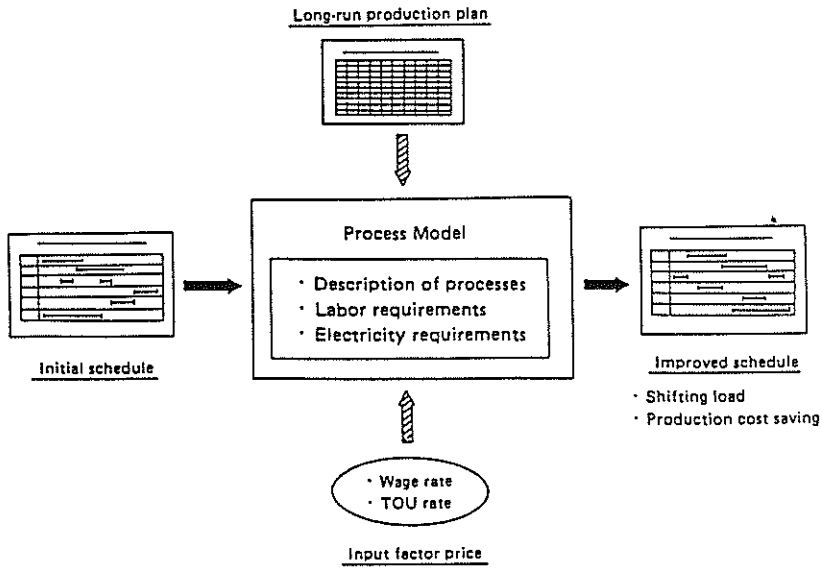


Figure 4. Overview of Process Model for the Heavy Electrical Machinery Industry

Since this type of scheduling problem is proved to be NP-complete, we can not find the best schedule within a reasonable time. Then we consider the scheduling using heuristic methods. These heuristic algorithms have several varieties in selection of the operation sequence for

improvement and also for the objective functions. A basic idea of algorithms for a stepwise improvement is expressed as follows:

Step 1) select one of the periods at which an activity is conducted in some order decided by a kind of method,

Step 2) judge the possibility of operation shift to other period taking into account the technological constraints,

Step 3) calculate the objective function for the shifted schedule

Step 4) shift the activity performance if the shifted schedule decreases the objective function value.

The characteristics derived from this basic idea are:

a) to shift activities to decrease the objective function monotonously

b) the optimality of the improved schedule is not guaranteed, since the algorithms do not take back-tracking steps.

We developed three types of algorithms. Each algorithm is summarized as follows:

Algorithm A: This is the simplest algorithm. An approach to rescheduling is to scan activities just forward or backward. Algorithm A minimizes process operating cost including energy charges and labor costs.

Algorithm B: To give an advantage to more effective activity for cost reduction, Algorithm B resequences the initial schedule according to a kW per man of each activity, i.e., a ratio of kW of electrical demand to labor requirements. The objective function is the same process operating costs as in Algorithm A.

Algorithm C: Algorithm C minimizes the maximum electricity demand for the entire planning period in order to reduce the demand charge.

Data

The model's engineering and cost information is based on the actual data of the Works. Processing time, electricity requirements, and labor requirements for each job are given. Hourly wage rate in the electrical industry is assumed based on Year Book of Labor Statistics, 1985 published by Ministry Ministry of Labor (1987). Rate scenarios 1 and 2 are assumed on the actual rates of a Japanese electric power company. TOU rates numbered 2 through 7 have higher energy charges during on-peak hours than off-peak hours. Rate scenarios numbered 8 through 11 have demand charges ranging from 1050 yen per kW to 2050 yen per kW. Energy charges are adjusted downward as demand charges are raised to make these rates revenue-neutral.

Simulation Results

Total operating cost includes monthly production costs for labor and electricity that would be affected by alternative operation schedules. Monthly electricity usage for the Works under the base rate was estimated by the model as 2290 MWh of energy use and 16.5 MW of demand during on-peak hours, and 1344 MWh of energy use and 12.5 MW of demand during off-peak hours. An electricity bill, computed using the base rate, amounts to 65 million yen per month. Under standard TOU rate rescheduling of electrical tests of three turbine-generators by Algorithms A and B brings about 7.2 million yen saving in comparison of the cost of reference operation pattern(approximately 11 percent of total electricity cost). Both algorithms produce good approximate solutions. A schedule improved by Algorithm C is found to be the most economical with 9.6 million yen saving, approximately 14 percent of total electricity cost. This cost saving is due mainly to demand charge saving. A schedule was remarkably improved by Algorithm C using output from Algorithm A(forward) as an initial schedule.

Figure 5 shows sensitivity analyses on energy charge. They demonstrate that the customer's load pattern varies significantly depending on rate structures, especially on the peak/off-peak energy charge ratio. Under the TOU rate, the on-peak kWh share was dropped from 78 percent

to 53 percent, then reducing total production cost to 19 percent. When the ratio of energy charges is between 1 and 2, the Works responded by reducing on-peak kWh share to 52.5 percent from the baseline case of 77.7 percent. When the ratio is raised to 3, effects of rate structure is saturating and the figure shows a gradual decline in total costs.

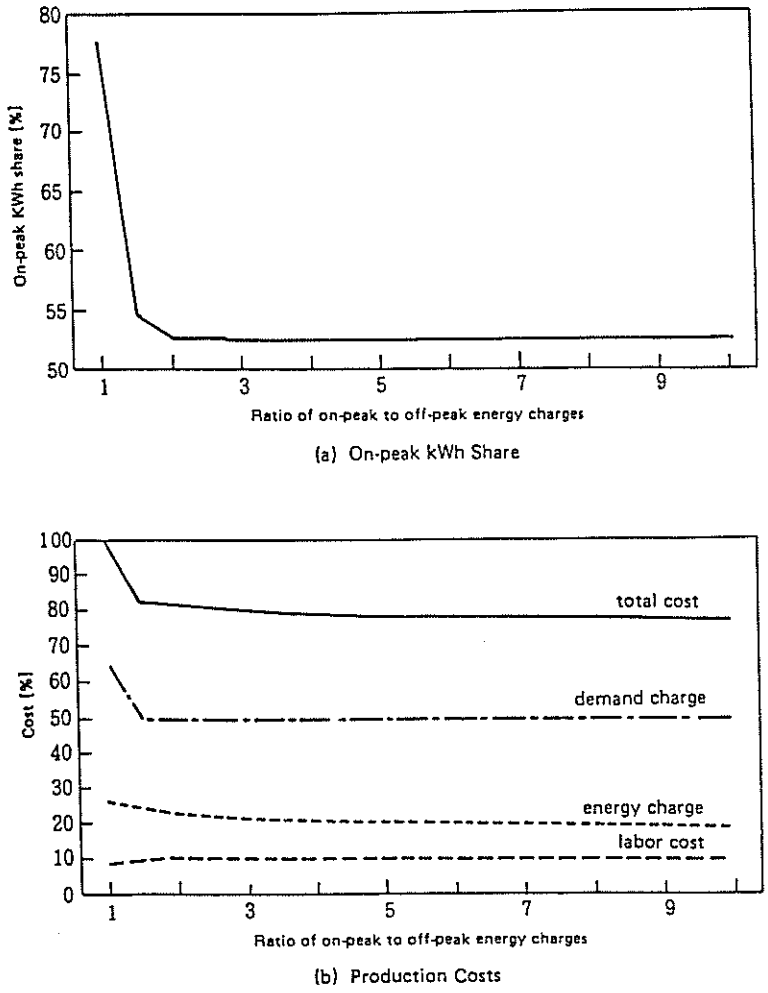


Figure 5. Sensitivity Analyses on Ratio of On-Peak to Off-Peak Energy Charge

As a typical industrial customer with high electricity intensity, an integrated steel mill was taken for the first case study. Simulations under more than the rate ratio of 1.1 indicated a potential for substantial load shifting in a steel mill. Gas holder management makes it easier to shift load from on-peak to off-peak hours. TOU rates seem to provide the mill with opportunities to lower costs. More self-sufficient of energy, more sensitive to TOU rates.

We have chosen an electrical apparatus works for the second case study. Simulation results for the heavy electrical apparatus works suggested the limits of TOU rates as a load management scheme. TOU rates were assumed to have significant load management effects only on relatively high electricity intensive processes(e.g., a final electrical test of turbine generators). Effectiveness of our approaches to estimate a potential response to TOU rates was demonstrated through these applications.

Among early examples of a process model is that of the Adams and Griffin (1972) model of the U. S. petroleum refining industry. The core of that model is a typical 200,000 barrel per

day refinery LP model. Major inputs to the model are the required product outputs, the prices of factor inputs, and the prices of by-products. The activity inequalities and the constraints are determined by the twelve major refining processes, the capacities of these processes, and the product quality requirements. The actual operation of the model involves simulation analysis where product demands, inventories and imports are solved from an econometric model. Medium- and long-term annual projections are prepared by solving the overall model in successive time periods to determine the impact of economic conditions and energy policies on the refining industry.

A more elaborate process model of the oil industry by Deam et al. (1974) extended that approach to include fuel transportation among some 25 worldwide geographical areas. Fifty-two types of crude oil and 22 refining centers were represented along with 6 types of tankers that could be selected for transport. The LP matrix for this model was quite large (about 3500 rows and 13,500 columns). The exogenous inputs to the model included future demands for products by region, refinery technology, costs of product refining, and transport of specific crudes and products. The model is solved to determine the optimal allocation and routing of crude oil and products between sources, refineries, and demand centers for some future target date.
Because the model includes the transport and refining costs of crude from specific sources, it provides a basis for analyzing the relative price of these crudes in a competitive market or in a controlled market where relative prices are set to reflect the differences in transportation and refining costs among the many sources.

Among models which expand process to include electricity demands, Baughman and Joskow (1974) adopted an approach which was primarily engineering in nature, while the Griffin (1979) approach was primarily econometric. More recently DeGenring and Jackson (1980) have balanced the coupling of these approaches by explaining full utilization with an econometric fuel choice model and capital equipment changes with engineering and cost relations. Hogan and Weyant (1983) also refer to integrated model systems as approximations to the pure ideal of a network of process models. Process models thus provide some natural elements for linking large components at engineering-economic systems.

Regarding applications to mineral engineering-economic analysis, Ray and Szekely (1973) have searched for the least-cost combination of production technologies in an integrated steel plant, subject to demand, technological, and raw material constraints. Their formulation allows the analyst to find the optimal mix of three process routes (blast furnace/basic oxygen furnace, blast furnace/open hearth furnace, electric furnace), depending on the prices of the factory inputs, such as scrap and hot metal. Such a formulation is of interest for studying the demand for materials that are used in the production sequence.

A number of other process models have been built for the steel industry. One of these was the model of Tsao and Day (1971) which is a short-run (fixed capacity) model employed to compare the optimal and actual behaviour of the U. S. carbon steel industry over the period 1955-1968.

Alternative arrangements of processes are included in the technology matrix, with optimization indicating the desirable proportion of each. In addition to the technology matrix, the model also contains a set of auxiliary activities, which represents the purchase of inputs and the sale of final outputs.

The constraint on input purchases prevents the total use of a given input (such as oxygen) from exceeding the amount purchased. Sales constraints are included to insure that production of finished products meets sales requirements.

Finally, Clark and Church (1981) describe a model that represents the production of stainless steel products (bar, sheet, strip, rod, wire, etc.) by alternative technological routes, starting with basic raw materials such as carbon and stainless steel scrap, ferrochromium, ferronickel, and

other materials including gases. Although not constructed explicitly as a demand model, this type of model can be used to analyze indirectly the demand for input materials required in the production process. In the case of stainless steel, it is possible to estimate the demand for input materials, such as chromium, nickel, and manganese (in various forms) as a function of the production of stainless steel by different processing routes.

REFERENCES

Adams, F.G. and J.M. Griffin (1969) "An Econometric Model of the U. S. Petroleum Refining Industry," Journal of the American Stastical Association 67: 542-51.

Asano,H., S.Sagai and K. Yamaji.(1988a). "Evaluation of Electric Power Demand Adjustment for the Heavy Electrical Machinery Industry". Economics & Public Utilities. CRIEPI, No.24 (in Japanese)

Asano,H., S.Sagai and K. Yamaji.(1988b). "Microscopic Analysis of Industrial Customers' Response to Time-of-Use Rates: Case Studies for an Integrated Steel Mill and a Heavy Electrical Apparatus Works." CRIEPI Report EY87002

Asano,H., S.Sagai and K. Yamaji.(1988c). "Process Models for Assessing Industrial Response to Time-of-Use Rates." Proceedings of Tenth Annual North American Conference, IAEE. Oct.31-Nov.2, 1988, Houston, Texas.

Brown, M., A. Dammert, A. Meeraus and S. Stoutjesdijk, (1983). Worldwide Investment Analysis: The Case of Aluminum, World Bank Staff Working paper No. 603, Washington, D. C.: The World Bank.

Choksi, A.M., A. Meeraus and A.J. Stoutjesdijk, (1983). The Planning of Investment Programs in the Fertilizer Industry, World Bank Research Publication, Baltimore: John Hopkins University Press.

Clark, J. and A. Church, (1981). "Process Analysis Modeling of the Stainless Steel Industry," Paper presented at the National Academy of Sciences Workshop on Non-Fuel Mineral-Demand Modeling. Airlie House, Warrenton, VA.

Deam, R.J. et. al, (1974). World Energy Model: Description and Results." Energy Modeling. London: IPC Science and Technology Press, Ltd.

Eckbo, P.L., Jacoby, H.D. and Smith, J.L., (1978). "Oil Supply Forecasting: A Disagregated Process Approach." Bell Journal of Economics. 9:218-230.

EPRI(1985). "Commercial and Industrial Response to Time-of-Use Rates: Methodology and Case Studies." EPRI EA-4206. Palo Alto, California: Electric Power Research Institute.

Fox, K.A., (1963). "Spatial Price Equilibrium and Process Analysis in the Food and Agricultural Sector," in A. Manne and H. Markowitz (eds.), Studies in Process Analysis, New York: John Wiley & Sons, pp. 215-234.

Griffin, J.M. (1979). Energy Conservation in the OECD: 1980-2000. Cambridge: Ballinger Publishing Co.

Hibbard, W.R., et al, (1979). "An Engineering Econometric Model of the U.S. Aluminum Industry," Proceedings, New York: American Institute of Mining Engineers.

Hibbard, W.R., A.L. Soyster and R.S. Gates, (1980). "A Disaggregated Supply Model of the U. S. Copper Industry Operating in an Aggregated World Supply/Demand System," Materials and Society, 4: 261-284.

Hogan, W. W. and J.P. Weyant (1983). "Method and Application for Energy Model Composition: Optimization in a Network of Process Models." In B. Lev (ed.), Energy Models and Studies. Amsterdam: North-Holland, 1983.

Judge, G.G. and T. Takayama (eds.), (1973). Studies in Economic Planning Over Space and Time, Amsterdam: North-Holland Publishing Company.

Jung, S.S., (1982). "An Investment Planning Model for the Refining and Petrochemical Industry in Korea," Ph.D. Thesis, University of Texas at Austin.

Kendrick, D., (1967). Programming Investment in the Process Industries: An Approach to Sectoral Planning. Cambridge: MIT Press.

Kendrick, D. and A. Stoutjesdijk, (1978). The Planning of Industrial Investment Programs, A Methodology, Baltimore: John Hopkins University Press.

Kendrick, D., A. Meerus and J. Alatorre, (1984). The Planning of Investment Programs in the Steel Industry, Baltimore: Johns Hopkins University Press.

Kennedy, M., (1974). "An Economic Model of the World Oil Market," Bell Journal of Economics and Management Science, 5: 540-577.

Kostad, C., D. Abbey and R. Bivins, (1983). "Modeling International Steam Coal Trade," WP No. LA-9661-MS, Los Alamos National Laboratory, New Mexico.

Kovisars, L., (1976). "World Production Consumption and Trade in Zinc - An LP Model," U. S. Bureau of Mines Contract Report J-0166003, Stanford Research Institute, Stanford, CA.

Kovisars, L., (1975). "Copper Trade Flow Model," World Minerals Availability, SRI Project MED 3742-74, Stanford Research Institute, Stanford, CA.

Labys, W.C. and D.O. Wood, (1985). "Energy Modeling," in W. Vogely (ed.), Economics of the Mineral Industry, New York: American Institute of Mining Engineers.

Labys, W.C. and P.I. Pollak, (1984). Commodity Models for Policy Analysis and Forecasting, London: Croom-Helm.

Labys, W.C. and C.W. Yang, (1980). "A Quadratic Programming Model of the Appalachian Steam Coal Market," Energy Economics, 2: 86-95.

Labys, W.C., (1987). Commodity Markets and Models: An International Bibliography, London: Gower Publishing Company.

Labys, W.C., F.R. Field and J. Clark, (1985). "Mineral Modeling," in W. Vogely (ed.), Economics of the Mineral Industry, New York: American Institute of Mining Engineers.

Manne, A.S. and H.M. Markowtiz (eds.), (1963). Studies in Process Analysis, New York: John Wiley & Sons.

Marschak, T.A., (1963). "A Spatial Model of U.S. Petroleum Refining," in A.S. Manne and H.M. Markowitz (eds.), Studies in Process Analysis, New York: John Wiley & Sons.

Ministry of Labor(1987). "Year Book of Labor Statistics, 1985."

National Academy of Science (NAS). (1982) Mineral Demand Modeling. Washington, D.C.: National Academy Press.

Russel C.S. and W.J. Vaughn (1976), Steel Production: Processes, Products and Residuals. Baltimore: Johns Hopkins Press.

Sparrow, F. and Syster, A. (1980), "Process Models of Mineral Industries," Proceedings of the Council of Economics of the AIME, New York: AIME, pp. 93-101.

Tsao, C.S. and R.H. Day (1971), "A Process Model of the U.S. Steel Industry." Management Science, 17:588-608.

Energy Vol. 15, No. 3/4, pp. 249–260, 1990
Printed in Great Britain

0360-5442/90 $3.00 + 0.00
Pergamon Press plc

2.3 THE HYBRID APPROACH TO DEMAND MODELING

DAVID B. REISTER

Engineering Physics and Mathematics Division
Oak Ridge National Laboratory
Oak Ridge, Tennessee
U.S.A.

Abstract - This paper presents an engineering-economic approach for developing energy demand models. The approach combines the economic theory of production with engineering process data. We review the economic theory of production, the Constant Elasticity of Substitution (CES) function, and the Logit function. Using the Oak Ridge Industrial Model (ORIM) as a case study, we show how the Logit function can be used to specify a fuel share model and how the CES function can be used to specify an energy conservation model. Shephard's lemma is a key result of the economic theory of production. The lemma can be used to derive input-output coefficients from a cost function. We use CES cost functions because they have good global properties. To obtain more flexibility, we use nested CES cost functions. To identify the parameters in a nested CES structure, we assume that the multilevel cost function corresponds to a multistage production process. When input-output coefficients are derived from a cost function, the tradeoffs between the factors satisfy the law of diminishing returns. Energy conservation occurs when capital, labor, or materials are substituted for energy service. In most cases, the substitution obeys the law of diminishing returns. Thus, a CES cost function can be used to specify an energy conservation model. Normally, perfect substitutes do not obey the law of diminishing returns. Since we include all of the costs for each option (the options use fuel and provide energy service) in ORIM, the options are perfect substitutes and do not satisfy the law of diminishing returns. Since the options are perfect substitutes, we use the Logit function to specify the ORIM fuel share model.

1. INTRODUCTION

The preceding chapters have discussed engineering-economic models based on econometric techniques and engineering process data. This chapter will discuss a hybrid approach that combines the best features of the two methods. Our approach is an engineering/economic methodology that uses the economic theory of production to create a model structure and uses both historical and engineering data to estimate parameters.

The methodology presented in this chapter is designed to provide policymakers and planners with an overview of technical change in an economy. The case study for this chapter is based on the Oak Ridge Industrial Model (ORIM). ORIM was developed during a two year period (1979-1980) for the Energy Information Administration (EIA) of the U.S. Department of Energy. ORIM was designed to be a part of a large model of energy supply and demand (the Midterm Energy Forecasting System (MEFS)).

ORIM is a dynamic model that forecasts energy demand by fuel, region, industry, energy service, and vintage of capital stock. The input parameters include: fuel prices and industrial output by region and industry. The data base includes information on: parameters estimated from historical data (price elasticities and rates of technological progress) and detailed engineering data (market shares for energy services, input-output coefficients for energy services and for industrial production).

ORIM is more detailed than an econometric model and less detailed than typical engineering process models. To interface with MEFS in a selected year, ORIM was required to forecast the demand for 6 types of fuel in 10 regions; a total of 60 numbers per year. ORIM forecasts energy demand for 6 types of fuel, 10 regions, 9 industries, 32 types of energy service, and 30 vintages of capital stock; a total of 518,400 forecasts. Each of the levels of disaggregation (industry, service, and vintage) represent important insights into the technical details of industrial energy demand. Although ORIM calculates detailed forecasts, the code and data are compact (the code is 1014 lines of FORTRAN and the data base consists of 431 lines of data).

The maintenance of an engineering-economic model (like ORIM) is expensive. Most of effort is required to develop the data base (The two year development of ORIM required about three person-years of effort. Approximately 70% of the effort was devoted to the development of the data base). Detailed surveys of industrial plants are required to develop the engineering data base. When the data base is developed, the model should be backcast to allow the model results to be compared to historical data (ORIM produces results for the period from 1970 to 2010).

An increase in energy cost can increase the cost of industrial output and reduce the production levels for industrial goods. The demand for industrial goods is outside the scope of ORIM. In ORIM, the level of output of each industry in each region is an input. ORIM focuses on the energy consumption required to produce a given level of industrial goods.

The remainder of this section provides a general introduction to the hybrid approach to energy demand modeling. Energy is an intermediate good that is consumed by capital equipment to provide energy services. Examples of energy services are: space heat, process heat, lighting, mechanical drive, and passenger transport. Often several options are available to provide each energy service. For example, space heat can be provided by furnaces that burn oil, gas, LPG, coal, or wood; or by electric resistance heaters; or by electric heat pumps. Intercity passenger transport can be provided by automobile, bus, train, or airplane.

All sectors of the economy use energy services. To create an energy demand model, we subdivide the economy into sectors and regions. Examples of sectors are the residential sector or the transportation sector. Sectors and regions can be quite detailed: the paper industry in Wisconsin or the commercial sector in Buffalo, New York. The number of sectors and regions that are appropriate depends on the purpose of the model and the available resources. Engineering-economic models tend to become large and expensive to construct and maintain. Discipline is required to focus the model on the key issues of interest to the client. For each sector and region, we identify the appropriate energy services.

Since energy is consumed by capital equipment to provide energy services, energy consumption patterns are influenced by both the investment decision and the operating decision. If we buy an electric resistance heater, we must always use electricity in the heater. If we buy an oil furnace, we might be able to use oil, gas, or LPG in the furnace.

For both the investment decision and the operating decision, we can subdivide each sector, region, and energy service into markets. For space heating in the residential sector in New York, the markets could include single family houses in Rochester and multi-family houses in New York City. For process steam in the paper industry in Wisconsin, the markets could include large factories with high capacity factors and small factories with low capacity factors.

For each market, we can identify options that provide the same energy service. We shall call the choice between two or more options "fuel switching," and we shall call a model that forecasts the choice between the options a fuel share model.

In addition to the fuel choice decision, each sector must decide how much energy service is required to produce the output for the sector. We shall call a reduction in energy service per unit of output "energy conservation," and we shall call a model that forecasts the demand for energy service per unit of output an energy conservation model.

Energy conservation occurs when capital, labor, or materials are substituted for energy service. In most cases, the substitution obeys the law of diminishing returns. For example, when a house is insulated, each increment of insulation causes a smaller reduction in energy consumption than the last increment.

Normally, fuel switching does not obey the law of diminishing returns. For example, two brands of regular gasoline can be substituted without diminishing returns. Since we include all of the costs for each option, the options are perfect substitutes and do not satisfy the law of diminishing returns.

The mathematical specification of fuel share models must be different than the specification of conservation models. For conservation models, we use a Constant Elasticity of Substitution (CES) function, while for fuel share models we use a Logit function.

In Section 2, we will introduce the CES and Logit functions. In Section 3, we will develop a fuel share model. In Section 4, we will develop an energy conservation model. In Section 5, we will summarize our results.

2. THE ECONOMIC THEORY OF PRODUCTION

Consider a sector that produces output (y) from a vector of physical inputs (x). Input-Output (I-O) coefficients (a_i) are the amount of input (x_i) per unit of output:

$$a_i = x_i / y \quad . \tag{1}$$

A basic economic identity is that revenues from sales minus payments to the factors of production (labor, energy, and materials) are equal to before-tax profits. If we consider profits to be payments to the factor capital, then total sales are equal to total factor payments:

$$C \cdot y = \sum_i p_i \cdot x_i \quad , \tag{2}$$

where C is the price of the output and p_i is the price of input x_i. If we divide both sides of Eq. (2) by output and use the I-O coefficients, then:

$$C = \sum_i p_i \cdot a_i \quad . \tag{3}$$

By including capital among the factors of production, we find that the price of the output is equal to the unit cost of production.

The basic concept in the economic theory of production is the production function. The production function is a summary of technology. The production function determines the maximum output that can be obtained from any vector of inputs:

$$y = F(x,t) \quad , \tag{4}$$

where t is time and reminds us that the maximum output from a given set of inputs can change with time. The production function summarizes all of the possible tradeoffs between factors of production.

For convenience, we shall assume that the production function is homogeneous of degree one. This assumption means that if we increase all inputs by a factor of two, the output will double:

$$F(\lambda \cdot x,t) = \lambda \cdot F(x,t) \quad , \tag{5}$$

where g is an arbitrary scale factor. At the factory level, this assumption means that there are no economies of scale. But at the sector level (where there are many factories), the assumption means that there is an optimum size for a factory and that expansion occurs by creating new factories of the optimum size. Since we are working at the sector level, this assumption is not very restrictive.

When there are sharp changes in energy prices, all of the managers of the factories in the sector must evaluate options for fuel switching and energy conservation. In the economic theory of production, each manager has a production function and a price vector and selects the combination of inputs that have the lowest unit cost. The result of this optimization process is a least cost set of I-O coefficients, that are determined by the price vector and the current state of technology.

Consider Eq. (3): the unit cost of production depends on the prices and the I-O coefficients. Since the optimum I-O coefficients depend on the prices, the minimum cost of production is determined by the prices:

$$C = G(p,t) \quad , \tag{6}$$

where G is called the cost function. The cost function is homogeneous of degree one with respect to price.

A central result of the economic theory of production is Shephard's lemma (Shephard, 1970). Shephard's lemma relates the optimum I-O coefficients to the partial derivatives of the cost function:

$$a_i = \frac{\partial G}{\partial p_i} \quad . \tag{7}$$

Shephard's lemma is an important theoretical result that has had a significant impact on the structure of energy demand models.

The cost share (s_i) is the fraction of the cost of production that is paid to each factor:

$$s_i = p_i \cdot a_i / C \quad . \tag{8}$$

The cost shares are positive and sum to unity.

The price elasticity of demand coefficient is the ratio of a percentage change in the demand of one factor to the percentage change in the price of another factor:

$$\sigma_{ij} = \frac{p_j}{a_i}\frac{\partial a_i}{\partial p_j} = \frac{\partial \ln a_i}{\partial \ln p_j} \quad . \tag{9}$$

If there are 10 factors of production, we can define 100 elasticity coefficients.

For an arbitrary cost function, the elasticity coefficients satisfy the following conditions:

$$s_i \sigma_{ij} = s_j \sigma_{ji} \text{ , and} \tag{10}$$

$$\sum_j \sigma_{ij} = 0 \quad . \tag{11}$$

Equations 10 and 11 are derived in the appendix to volume 2 of Ref. 2. If there are 10 factors of production, Eq. (10) determines 45 elasticity coefficients and Eq. (11) determines 10 coefficients. Thus, the constraints of economic theory reduce the number of unknown elasticity coefficients from 100 to 45.

To use Shephard's lemma to derive the structure of an energy demand model, we choose an arbitrary homogeneous of degree one cost function and use Eq. (7) to derive the demand equations. Consider a Cobb-Douglas cost function:

(12)

For the Cobb-Douglas cost function, the optimum I-O coefficients are given by:

$$a_i = \delta_i \cdot G / p_i \quad . \tag{13}$$

Thus, for the Cobb-Douglas cost function, the cost shares are constants:

$$s_i = \delta_i \quad . \tag{14}$$

Conversely, the base year cost shares determine all of the free parameters in a Cobb-Douglas cost function.

Contrast a conventional set of econometric demand equations with a set derived from a cost function. For conventional econometric equations, the independent variables could be prices, quantities, lagged variables, and other relevant variables, such as heating and cooling degree-days. For demand equations derived from a cost function, the only independent variables are price and time. Consider a log-log demand system with price as the independent variable:

$$\ln a_i - \ln \phi_i + \sum_j \psi_{ij} \ln p_j \quad . \tag{15}$$

Using conventional econometric methods, the parameters (ϕ_i and ψ_{ij}) in Eq. (15) are unrestricted. However, deriving a set of demand equations from a cost function places severe restrictions on the parameters. If Eq. (15) is derived from a cost function, then Eq. (15) must be identical to Eq. (13). Thus, $\phi_i = s_i$, $\psi_{ij} = s_i$ when i is not equal to j, and $\psi_{ij} = s_i - 1$. When the parameters in a log-log demand system are derived from a cost function, all of the parameters are determined by the base year cost shares.

To derive a system of demand equations with more free parameters, we will consider a Constant Elasticity of Substitution (CES) cost function:

$$G(p) = \{ \Sigma_i \alpha \cdot p_i^{\rho} \}^{1/\rho} \quad , \tag{16}$$

where the parameters α_i are positive and sum to unity. For the CES function, the shares satisfy:

$$s_i = \alpha_i (p_i/G)^{\rho} \quad . \tag{17}$$

For the CES cost function, a convenient assumption is that the factor prices are price indexes that are equal to 1.0 in a base year. With this assumption, G = 1.0 in the base year and the parameters (α_i) are equal to the base year shares.

For both the Cobb-Douglas and CES models, most of the parameters are determined by the base year shares. The CES model has one more parameter (p) than the Cobb-Douglas model. The parameter p controls the price response of the model. The range for p is $1 \geq p \geq -\infty$. When p = 1.0, the I-O coefficients are constant and have no price response. When p = 0.0, the shares are constant and the CES model is identical to the Cobb-Douglas model. When p is a large negative number, the I-O coefficients have a large price response.

If there are 10 factors of production, a model should have 45 parameters to determine the 45 free elasticity coefficients. Economists have developed several functions that have more parameters than the CES function; for example, the Generalized Leontief Function and the Translog Function. To gain more free parameters, we use nested CES functions:

$$G(p) = \{\Sigma_i \cdot g_i p\}^{1/p} \quad , \tag{18}$$

where:

$$g_i(p) = \left(\sum_j \beta_{ij} * p_i^{\eta_i} \right)^{1/\eta_i} \quad , \tag{19}$$

and the parameters α_i and β_{ij} are positive and sum to 1.0, while the parameters p and η_i are less than 1.0.

To identify the parameters in a nested CES structure, we assume that the multilevel cost function corresponds to a multistage production process. At the first stage (Eq. (19)), factors of production produce several intermediate goods. At the second level of production (Eq. (18)), output is produced from the intermediate goods. In Section 4, we will use a nested CES structure to specify an energy conservation model. We conclude this section by defining the logit function, that will be used Section 3 to specify a fuel share model.

Suppose that there are several options that can provide the same energy service. Let x_i be the amount of energy service from each option and let y be the total demand for energy service, then:

$$y = \sum_i x_i \quad . \tag{20}$$

If we consider Eq. (20) to be a production function, the I-O coefficients are the fuel shares, and the fuel shares sum to unity:

$$\sum_i a_i = 1 \quad . \tag{21}$$

If a cost function (such as the Cobb-Douglas function, the CES function, or the translog function) is used to specify a fuel shares model, the fuel shares will not sum to unity for any set of parameter values. The reason for the failure is fundamental. A cost function is derived from a production function by cost minimization. Given prices for each of the options, the least cost solution is to have a 100% market share for the least cost option.

To estimate the parameters for a fuel share model, a data base is developed for a set of economic sectors in a geographic region. Normally, several options have a partial share of the market. To develop a fuel share model in which several options have a partial share of the market, we assume that in the region each decision maker faces a unique set of costs for energy services. If each decision maker chooses the least cost option, the market share is the probability that the option has a lower cost than any of the competing options. If we assume that the

distribution function for the regional costs is a Weibull distribution, then we can derive the logit fuel share function (see Boyd, Philips and Regulinski, 1982, and Reister, 1983a):

$$a_i = \frac{p_i^{-\gamma}}{\sum_j p_j^{-\gamma}} , \tag{22}$$

where p_i is the cost of option i and γ is a parameter. Clearly, the sum of the fuel shares forecast by the logit function is equal to 1.0

3. FUEL SHARE MODELS

To illustrate how to use the logit function to specify a fuel share model, we will describe the fuel share module of the Oak Ridge Industrial Model (ORIM). Several versions of ORIM exist. Version 3 has the best documentation (Reister, Barnes and Edmonds, 1982). Version 5 is an annual model that forecasts energy demand for nine manufacturing sectors and for ten regions (Reister, 1983b). Version 6 is a one region model, that was used in the recent Energy Modeling Forum study of industrial energy demand (EMF-8). We will describe Version 6. ORIM has two fuel share models: one for the investment decision and the other for the operating decision (see Figure 1). We will describe the fuel share model for the operating decision.

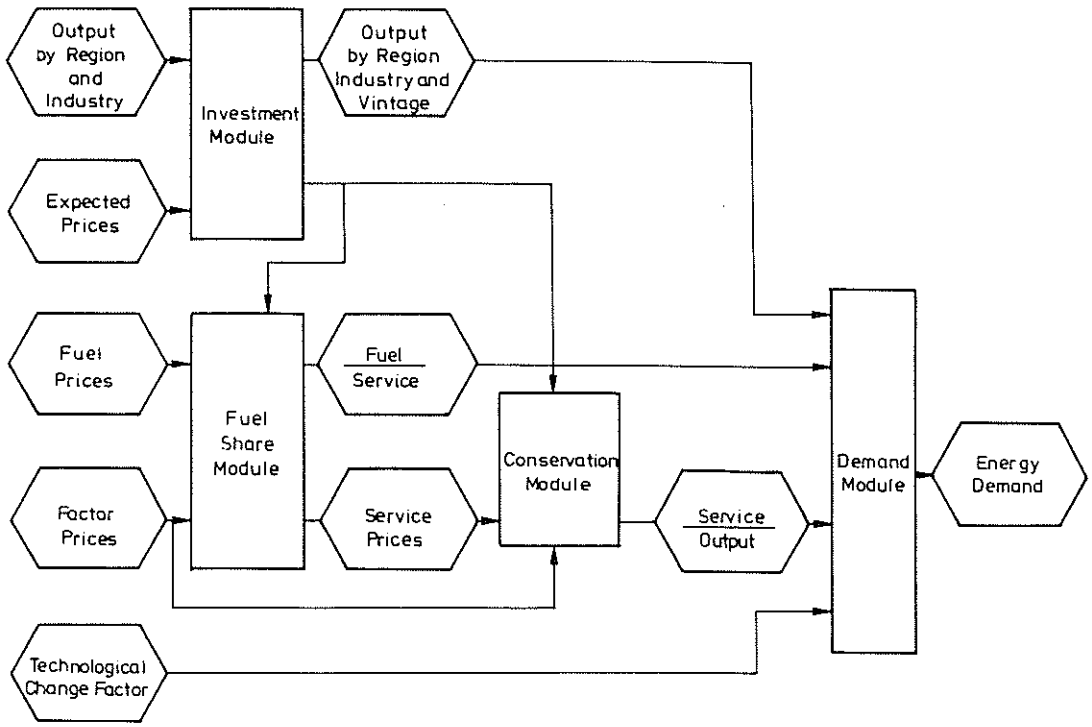


Figure 1. The Structure of the Oak Ridge Industrial Model.

At the time when ORIM was designed (1979), the existing industrial energy demand models were forecasting too much demand for coal. ORIM was designed to forecast the demand for the

six types of energy shown in Table 1. In the table, coal has the lowest price, but only 12% of the market. Natural gas has a higher price and 40% of the market. Electricity has the highest price and the second largest market share.

Table 1. Industrial Energy Demand in the United States in 1985.

	Type of Energy	Demand Trillion BTU	Market Share %	Price Dollars per Million BTU
1.	Coal	1677	12	1.59
2.	Natural Gas	5503	40	4.10
3.	LPG	1576	12	4.17
4.	Distillate Oil	1192	9	5.59
5.	Residual Oil	874	6	3.85
6.	Electricity	2804	21	15.15
	Total	13626	100	

To develop a fuel share model that can backcast the fuel shares displayed in Table 1, we must subdivide the market and include the other factors (capital, labor, and materials) required to produce energy service. For each industry in each region, energy demand is subdivided into four types of energy service: process steam, process heat, mechanical drive, and electrical service. The four types of energy service are further subdivided into market segments on the basis of characteristics such as size, capacity factor, and fuel option. Version 6 of ORIM has 11 types of steam, 12 types of heat, 6 types of drive, and 3 types of electrical service.

Coal has a small market share because it is a solid fuel that is expensive to handle and burn. If we include the total cost of capital, labor, and materials, the cost of energy service from coal is more than the total cost of energy service from natural gas. Furthermore, the cost of energy service provided by coal is higher for small plants with low capacity factors than for large plants with high capacity factors. The advantage of subdividing the market is that the fuel shares can be different in each market segment.

Electricity has a large market share because of its competitive advantage in certain market segments. Electrical services such as lighting, electronics, and electrochemical processes must be performed by electricity. Electric motors have a large share of the mechanical drive market. Electricity has a small fraction of the process heat and process steam markets.

In ORIM-6, there are nine factors of production: capital, labor, materials, and the six types of energy. The prices {P(M,T)} are exogenous inputs for each factor (M) and year (T). For each type of energy service (N) and each market segment (NN), there are several competing options (NI) that can provide energy service. The options are described by I-O coefficients (AI). The total cost of each option (PI) is the sum of the product of the prices and the I-O coefficients:

$$PI(N,NN,NI) = \sum_{M} AI(N,NN,NI,M) \cdot P(M,T) \quad . \tag{23}$$

To allow us to adjust the base year market shares, we add a set of control variables {CN(M)} and define PIC by:

$$PIC(N,NN,NI) = \sum_{M} AI(N,NN,NI,M) \cdot \{P(M,T) + CN(M)\} \quad . \tag{24}$$

When the I-O coefficients are correct for each option in each market segment, the control variables will be zero. However, in most cases the control variables will not be zero.

The market shares (SI) are determined by a logit function:

$$SI(N,NN,NI) = PIC^{RN} / \sum_{NI} PIC^{RN} \quad , \tag{25}$$

where the logit exponent depends on the type of energy service {RN = RN(N)}.

Let BN be the I-O coefficients for the market segment determined by the market shares:

$$BN(N,NN,M) = \sum_{NI} SI \cdot AI \quad . \tag{26}$$

We assume that the actual I-O coefficients (AN) track the BN with a lag (QA):

$$AN(N,NN,M,T) = QA \cdot BN(N,NN,M) + (1 - QA) \cdot AN(N,NN,M,T\text{-}1) \quad . \tag{27}$$

The average cost of energy service in each market segment (PN) is the sum of the product of the I-O coefficients and the prices:

$$PN(N,NN,T) = \sum_{M} AN(N,NN,M,T) \cdot P(M,T) \quad . \tag{28}$$

In ORIM-6 capital stocks have a lifetime of 30 years and 30 vintages of capital stocks are active each year. Each year all of the 30 year old equipment retires and the investment module determines the characteristics of the new equipment. For example, the investment module must choose between boilers that can use coal and boilers that can only burn fuel oil, LPG, and natural gas.

The results of the investment module are summarized in the array AT(N,NN,L). For each type of energy service (N) and for each vintage (L), the AT coefficients are the market shares for each of the characteristic types of energy service (NN). The sum of the AT coefficients over NN is equal to 1.0. For example, suppose that for a certain industry in a certain region, 30% of the process steam is produced in large boilers with high capacity factors. The investment module chooses between coal boilers and oil/gas boilers. If coal captures two thirds of the market in a certain year, the AT coefficient will be 0.20 for coal and 0.10 for oil/gas boilers.

To calculate the I-O coefficients for each type of energy service (AE), we sum over all of the market segments:

$$AE(N,L,M,T) = \sum_{NN} AT(N,NN,L) \cdot AN(N,NN,M,T) \quad . \tag{29}$$

The cost of each type of energy service (PS) includes both the operating costs and the sunk capital costs (ACSK):

$$PS(N,L,T) = \sum_{NN} AT(N,NN,L) \cdot PN(N,NN,T) + ACSK(N,L) \cdot P(K,T) \quad , \tag{30}$$

where P(K,T) is the price of capital.

The AE coefficients are the output of the fuel share module. They are the amount of fuel of type M required to produce one unit of energy service N in year T from vintage L. To calculate the fuel share coefficients, we used a logit function to forecast market shares for each option in each market segment.

4. ENERGY CONSERVATION MODELS

ORIM forecasts energy demand using the following basic equation:

$$\text{Energy Demand} = \frac{\frac{\text{Fuel}}{\text{Service}} * \frac{\text{Service}}{\text{Output}} * \text{Output}}{\text{Technological Change}}$$

The level of output and the technological change factor are exogenous inputs (see Figure 1). The fuel share module forecasts the ratio of fuel and service (AE) and the conservation module forecasts the ratio of service and output (AS). The technological change factor simulates changes in unit energy consumption that are driven by technology rather than by price. An example is a lower cost and/or better quality dye for fabric that works in cold water and replaces a hot water dye.

Energy conservation is easier for a new plant than for an old plant. The conservation module simulates retrofit energy conservation, while the investment module simulates energy conservation in a new plant. In the conservation module, energy conservation is modeled by a short-run cost function. In the investment module, energy conservation is modeled by a long-run cost function. We will describe the conservation module.

In ORIM there are 9 factors of production (capital, labor, materials {KLM}, and 6 types of energy). In the conservation module there are 7 factors of production (KLM and 4 types of energy service). In the conservation module, we want to substitute KLM for each type of energy service but we do not want to substitute one type of energy service for another.

To specify the conservation module, we use a nested CES structure (see Eqs. (18) and (19). To prevent substitution of one type of energy service for another, we use fixed coefficients for the top level (p=1). We define seven intermediate goods: one each for KLM and one for each of the energy services. The only input to the capital intermediate good is capital. The labor and materials intermediate goods are similar. For the energy service intermediate goods, the inputs are KLM and one of the types of energy service.

In the ORIM structure, KLM are subdivided into three groups: production of energy service, conservation of energy service, and balance of plant. There are practical problems but no conceptual problems with adopting accounting rules to allocate the KLM to one of the three groups.

For CES cost functions a convenient assumption is that the price is a price index, that is equal to 1.0 in the base year. We define the price index (PP) by:

$$PP(L,T,I) = P(I+6,T)/BPF(L,I) \text{ for } I = 1 \text{ to } 3, \text{ and} \tag{31}$$

$$PP(L,T,N+3) = PS(N,L,T)/BPS(L,N) \text{ for } N = 1 \text{ to } 4, \tag{32}$$

where the base prices for factors (BPF) and services (BPS) are calculated in the investment module.

We will call the two levels of the CES structure: the output level and the conservation level. At the output level, output is produced from inputs of the seven intermediate goods:

$$G(L,T) = \sum_{I} CS(I,L) \cdot GS(I,L,T) \ , \tag{33}$$

where the CS parameters are calculated in the investment module.

At the conservation level, the cost functions for the seven intermediate goods are given by:

$$GS(I,L,T) = PP(L,T,I) \quad \text{for } I = 1 \text{ to } 3, \text{ and} \tag{34}$$

$$GS(I,L,T) = \{\Sigma\ CT(I,J,L) \cdot PP(L,T,J)^{RT}\}^{1/RT} \quad \text{for } I = 4 \text{ to } 7, \tag{35}$$

where the CES exponent can have different values for each type of energy service {RT = RT(I)} and the CT parameters are calculated in the investment module. For the energy service intermediate goods, the inputs are KLM and one of the types of energy service. The CT parameters can be positive when J = 1 to 3 and must be positive when J = I. The CT parameters are zero when J = 4 to 7 unless J = I.

The input-output coefficients for each type of energy service (AS) are derived using Shephard's lemma (Eq. (7)). The partial derivative of G (Eq. (33)) with respect to the price of one of the types of energy service has a single term:

$$AS[N,L,T] = \frac{CS[I,L]*CT[I,I,L]}{BPS[L,N]} \{PP[L,T,I]/GS[I,L,T]\}^{RT-1} , \tag{36}$$

where N is the type of energy service and I = N + 3.

The fuel share module forecasts the ratio of fuel and service (AE(N,L,M,T)) and the conservation module forecasts the ratio of service and output (AS(N,L,T)); where N is the type of energy service, L is the vintage, M is the type of fuel, and T is the year. To calculate the demand for energy (E(M,T)), we sum over vintage and energy service:

$$E(M,T) = \sum_N \sum_L AS(N,L,T) \cdot AE(N,L,M,T) \cdot XO(L,T)/RC(N,L) , \tag{37}$$

where XO(L,T) is the output from vintage L and RC(N,L) is the technological change factor.

5. CONCLUSIONS

This chapter has discussed a hybrid approach for developing energy demand models that combines the economic theory of production with engineering process data. In Section 2, we reviewed the economic theory of production and introduced the Constant Elasticity of Substitution (CES) function and the Logit function. Using the Oak Ridge Industrial Model (ORIM) as a case study, we showed how the Logit function can be used to specify a fuel share model and how the CES function can be used to specify an energy conservation model.

Shephard's lemma is a key result of the economic theory of production. The lemma can be used to derive input-output coefficients from a cost function. We use CES cost functions because they have good global properties. To obtain more flexibility, we use nested CES cost functions. To identify the parameters in a nested CES structure, we assume that the multilevel cost function corresponds to a multistage production process.

When input-output coefficients are derived from a cost function, the tradeoffs between the factors satisfies the law of diminishing returns. Energy conservation occurs when capital, labor, or materials are substituted for energy service. In most cases, the substitution obeys the law of diminishing returns. Thus, a CES cost function can be used to specify an energy conservation model.

Normally, fuel switching does not obey the law of diminishing returns. For example, two brands of regular gasoline can be substituted without diminishing returns. Since we include all of the costs for each option in ORIM, the options are perfect substitutes and do not satisfy the law of diminishing returns. Since we cannot use a CES cost function to specify a fuel share model, we use the Logit function.

ORIM can be viewed as a five level production process. Starting at the most detailed level, the levels are: options, choice, markets, conservation, and output. The outputs of each level are the inputs to the next level. Three of the five levels have fixed input-output coefficients: options, markets, and output. Fuel switching occurs at the choice level, while energy conservation occurs at the conservation level.

The data base for ORIM uses engineering process data at the option level, end-use survey data at the market level, and aggregate statistical data at the output level. In the base year, all of the price indexes are equal to 1.0 and many of the parameters in the model can be determined from the base year data.

REFERENCES

Boyd, D. W., R. L. Philips and S. G. Regulinski (1982). "A Model of Technology Selection by Cost Minimizing Producers," Management Science 28: 418-424.

Reister, D. B., R. W. Barnes and J. A. Edmonds (1982). "The Oak Ridge Industrial Model," ORNL/CON-56, Oak Ridge: Oak Ridge National Laboratory.

Reister, D. B. (1983a). "An Analysis of Industrial Demand for Natural Gas," Energy 8: 749-755.

Reister, D. B. (1983b). "Results of the Oak Ridge Industrial Model," ORAU/IEA-83-10(M), Oak Ridge: Oak Ridge Associated Universities.

Shephard, R. W. (1970). Theory of Cost and Production Functions, Princeton: Princeton University Press.

Energy Vol. 15, No. 3/4, pp. 261–278, 1990
Printed in Great Britain

0360-5442/90 $3.00 + 0.00
Pergamon Press plc

2.4. ACCOUNTING AND END-USE MODELS

BERTRAND CHATEAU

AFME
Paris, FRANCE

and

BRUNO LAPILLONNE

Asian Institute of Technology
Bangkok
THAILAND

Abstract - Accounting and end-use models were developed in industrialized countries in the early seventies. Since then they have become rather popular and are now widely used all over the world. These models are directly derived from the so-called technico-economic analysis of energy demand.

Based on the concepts of applied system analysis, this approach aims at identifying all economic, social, technical and political factors which are related to energy demand and also at understanding how these factors explain changes in energy demand in the medium and long term. Focusing on the long term, this methodology is above all, a prospective approach which enables an analysis of all possible energy demand patterns in the long range instead of trying to predict just one in particular.

1. INTRODUCTION

Accounting and end-use models were developed in industrialized countries in the early seventies. Since then they have become rather popular and are now widely used all over the world. These models are directly derived from the so-called technico-economic analysis of energy demand.

Based on the concepts of applied system analysis, this approach aims at identifying all economic, social, technical and political factors which are related to energy demand and also at understanding how these factors explain changes in energy demand in the medium and long term. Focusing on the long term, this methodology is above all, a prospective approach which enables an analysis of all possible energy demand patterns in the long range instead of trying to predict just one in particular.

As a consequence, end-use models are not conceived as predictive tools (as an econometric model may be) but as tools which help the analysis of long term possible evolutions of energy demand. In particular, they make it possible to rapidly and accurately assess the consequences on the energy demand evolution of all types of uncertainties that affect our appraisal of the long term development of the society.

In a first part, the main characteristics of end-use models will be pinpointed. We will show that behind this general modeling approach, three different types of such models exist. Then as

an illustration, the MEDEE models and their applications will be presented. From the experience achieved with these models, particular considerations will be made upon the field of application of end-use models. Lastly, the main issues one confronts when adapting or using end-use models will be discussed.

2. END-USE MODELS: GENERAL CHARACTERISTICS AND HISTORY

General Characteristics

These models were initially developed to analyze the long-term effects on demand of higher energy prices and energy savings policies in some industrialized countries (e.g. France with MEDEE, (Chateau and Lapillonne, 1977, 1978), UK with Leach (Leach et al., 1979), USA with CONAES (1978)).

Although there may exist certain differences from one study to another, these models nevertheless have the same general basis:

- a detailed representation of energy end-uses (e.g. cooking and lighting for households or thermal energy for industrial processes);
- calculation of the demand for useful and final energy for each end-use, on the basis of technical, economic and social indicators usually measured in physical units (such an approach deals explicitly with the equipment used by energy consumers);
- use of a scenario approach to account for the development of all factors which depend on policy options both in the energy field and elsewhere.

Total demand is thus obtained from the aggregation of the demand of the various end-uses considered. For this reason, this approach to energy demand is often referred to as the "bottom-up approach".

For a given end-use, the final energy demand (e.g. oil products, coal, electricity) is considered a result of the satisfaction of a physical energy requirement (e.g. thermal or mechanical energy) in an appliance or equipment (e.g. boiler, furnace, motor), which can be characterized by the fuel required and its end-use efficiency (Figure 1). This physical energy requirement represents energy in the form that is really "useful" for the consumers (the "useful energy") either to satisfy a given need (e.g. comfort in dwellings, mobility) or to carry out an economic activity (e.g. production of steel, of a service) in a given physical and technological context (e.g. climate, dwelling, production process). The distinction between useful and final energy makes it possible to distinguish two components in the energy demand. On one hand, the "useful energy" is directly related to the level of development of the country; on the other hand, the transformation of useful energy into final energy, is related to the type of energy and equipment used (end-use efficiency). With this distinction made, it is possible to dissociate the influence of life styles, from the growth of energy demand and more generally the pattern of development from the influence of changes more related to energy (in particular, the effect of energy policies). In addition, this distinction makes it possible to measure the effect of energy substitutions on the level of total demand, especially for fuels with very different end-use efficiencies. For instance, the replacement of fossil fuels by electricity in some industrialized countries, or the change-over from non-commercial to commercial fuels in the developing countries.

Thus, this type of model requires a very thorough statement of the assumptions concerning future changes in the strictly technological factors as well as in the style of economic and social development: structure of transport (modal distribution) and housing; the patterns of land use, which will determine the mobility needs; GDP structure (relative share of the various sectors) as

well as industrial structure, which will depend on the domestic consumption pattern and the nature of foreign trade.

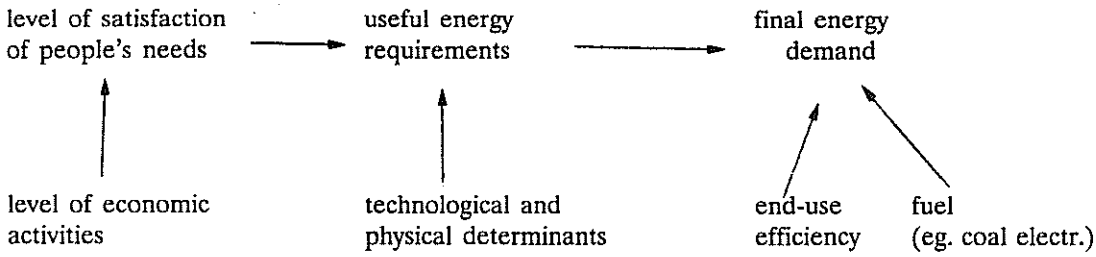


Figure 1. General Scheme for Energy Demand Analysis in End-Use Models.

Accounting Models

The accounting models represent the first generation of end-use models. In these models, the progression of most variables is exogenous and defined within the framework of a scenario. The main function of the model is to rapidly perform a whole series of multiplications and additions (thus, they are accounting tools rather than real models). The changes in the social and economic structures, the technology (especially energy savings), and the fuel mix (i.e. energy substitution) are exogenous and specified in the scenario, which thus plays a key role. Users of these models must therefore have a very accurate knowledge of the social, economic and technical problems linked to energy. The MEDEE 2 (Lapillonne IAEA, 1986) as well as the models used in the Leach et al. (1979) or CONAES (1978) studies, belong to this category. These type of models have been recently adapted to deal with energy demand issues of developing countries (e.g. MEDEE S (Lapillonne, 1985; Dumort and Lapillonne, 1987), CIRED (Dartois, Hurcade, Lagandre, 1980), OLADE (1984), European Commission (1984) and IDEE-BARILOCHE (1986) within an EC network). Accounting models have a number of advantages which, to a great extent, explain their popularity:

- Firstly, they provide simple and transparent answers to a number of questions concerning the long-term development of energy issues (e.g. the impact of contrasting energy policies, especially with regards to energy savings and the development of substitutes for hydrocarbons, and also the effect of the style of development on energy demand.)

- In addition, these models enable a better analysis of energy supply issues by investigating the balance of supply and demand at a more detailed level (e.g. taking account of temperatures, and spatial dimension).

- Finally, with these models the major demand determinants are made explicit, particularly the variables of action to control the long-term development in energy demand (with a distinction between technical and socio-economic variables of actions).

Prices are not considered explicitly in accounting models. The role of prices is supposed to be taken into account in the scenario assumptions, particularly in those related to energy savings and substitutions. Prices are not included because of a belief that price elasticities are not relevant for long term forecasting and that prices are not the only factors to explain energy savings and substitutions. However, this absence of prices quickly revealed its limitations. Indeed on the basis of exogenous objectives of energy savings and substitutions, the many studies carried out with accounting models have well demonstrated the range of variation of energy demand over the long term, but have failed to indicate the conditions to reach these objectives, especially the price level and policy measures required. As will be seen, recent developments of these models

have attempted to overcome these limitations, either by introducing econometric relationships to account for energy savings and substitutions, (this time at a disaggregated level), or by relying on models simulating consumer behavior ("simulation models").

Econometric End-use Models

In these models, the energy substitutions for space heating, are based on an econometric sub-model of market share (e.g. "logit" model). The heating requirements may be related to energy prices. As examples of these models, we can mention the ENDIM model developed by J.K. Parikh for India (Parikh,1981), the model of the Canadian Department of Energy (ISFD, Seeto and Erdma, 1980), and the MIDAS model of the Commission of European Communities (1985). This type of approach is especially popular for the modeling of road transport fuel demand. In this case gasoline or diesel demand is considered as the resultant of three components-a stock of vehicles, a mileage (annual distance driven by vehicle) and a unit consumption (miles per gallon, liters per 100 km or km per liter), each modelled through separate econometric models taking into account prices, incomes and other variables. A good example of such approach are the models developed by G. Kouris for the US (Kouris, 1980) or the OECD and Less Developed Countries (LDC's) (Kouris, 1985).

Some of these econometric end-use models are actually combined to a macro-economic model and have been developed to represent energy demand determinants better. End-use modelling is dealt with in "satellites models" (see for instance, the model of the Department of the Environment and Transportation of U.K. (1981), and the EPM model (Ohman, 1980) in Sweden).

Such econometric end-use models have the advantage of making energy demand projections more consistent at the macro-economic level, yet have the drawback of being over-dependent on the past and the market mechanisms. The fact that they rely on price elasticities for long term forecasting is also highly questionable. Their great advantage with respect to the traditional econometric models is to be able to cope with the long term structural changes (because of their high level of disaggregation) and the role of prices at the end-use level.

Simulation End-Use Models

Another development of end-use models has been to simulate investments in energy savings and substitutions on the basis of relationships reflecting consumer behavior, taking into account energy prices, costs of energy savings actions or of fuel substitution, as well as other factors (subsidies, interest rates, financing constraints, etc.). In general, investment behavior of the various types of consumers is characterized by a discount rate and variables taking into consideration the fact that consumer do not necessarily carry out cost effective investments (lack of information, consideration of non economic factors etc.). Among this type of "simulation model" we can mention the MEDEE 3 model (Chateau and Lapillonne, 1979; Camos et al., 1986) (especially in its version "energy demand management", MEDEE 3 ME (Chateau and Lapillonne, 1985)), as well as other models in Belgium (SPPS, 1983), the United Kingdom (SPRU, Skea, 1987) and the United States (e.g. Lawrence Berkeley Laboratory, (Meier, 1982); E. Hirst et al., (1976)).

The treatment of demand in energy supply optimization models like EFOM or MARKAL is somehow similar to that of simulation models. Energy savings or substitutions are considered on the basis of energy, prices and investments costs of actions. However, the representation of consumer behavior is different because in those optimization models it is assumed that there exist a unique decision maker behaving according to collective interest (national discount rate). On the opposite simulation, models differentiates among homogeneous groups of consumers and aim at representing their real behavior for each of them.

3. THE MEDEE MODELS

As compared to the general characteristics of the end-use models, the MEDEE models present three main particularities.

First of all, the disaggregation of energy demand within the model is neither arbitrary nor directly issued from the usual energy consumption statistics. It results from a systematic process aiming at identifying homogeneous energy demand categories (or so-called modules), with regard to:

- end-uses;
- economic function of the consumers;
- consumers behavior;
- physical and technological context in which the energy requirements are satisfied (climate, dwelling, process).

This approach aims at two major objectives:

- to make use of conclusions and results of micro-analysis of consumer behavior for macro-modelling purposes, assuming that the more homogeneous the module is the more accurate the use of the "mean behavior" is to depict the behavior of the module;
- to strengthen the long term stability of the model's relations by minimizing the bias introduced by all kinds of structural changes (e.g. changes in the structure of consumers, of production, of end-uses, of equipment, of energy supply).

Secondly, the use of the concept of "useful energy" is restricted to those modules where a competition does exist among fuels with significantly different end-use efficiencies. In this case, the useful energy requirements are measured either in terms of relative useful energy or in terms of final energy for a reference fuel:

- The concept of relative useful energy is used if the social needs or economic activities are not affected by the energy substitutions. It is measured as the final energy consumption of the most efficient fuel/equipment.
- The reference energy is measured as the final energy consumption of a reference fuel/equipment to meet social needs or economic activities if those are to be affected by the fuel/equipment in competition. As an example, energy requirement for cooking in rural areas of developing countries are measured in terms of a reference energy, since the transition from fuel wood to modern fuels is usually accompanied by a transformation of the cooking habits. The purpose of restricting the use of the "useful energy" concept and measuring it in terms of relative useful energy or reference energy is to cope with both the necessity of calculating useful energy and the difficulty of measuring the physical reality of useful energy.

Finally, the exogenous variables of the models, whose long term evolution is affected by strong uncertainties, are organized in a hierarchical scenario structure in which the following are specified:

- those variables whose evolution is independent on decisions that may be taken by national decision-makers (e.g. variables related to the international environment);

- those variables dependent on the former level, which describe the main pattern of the socio-economic development on the country or region under study and which are affected by the general policy of the government;

- those variables dependent or linked with the variables of the two former types, which may be used to specify sectoral policies (either public or private), in particular the various aspects of the energy policies (promotion of such energy form, energy conservation objectives, etc.);

- those variables whose evolution is largely determined by the variables of the three former types.

The purpose of establishing such a structure is:

- to make the design of the scenario itself easier, more systematic and more consistent;

- to make it possible to better identify, classify and specify the various types of uncertainties affecting the long term evolution;

- to allow one to clearly specify the various political options and directly simulate their consequences on energy demand evolution.

We shall review now the main versions of the MEDEE models that have been developed so far.

MEDEE 2, MAED and MEDEE S

MEDEE 2 (Lapillonne, 1986) is a simple accounting model, initially designed for the Energy Program of IIASA (International Institute for Applied System Analysis, Laxenburg (Austria)). The energy demand sub-model of MAED (IAEA, 1986) (called "module 1", the other sub-models dealing with electricity load curve projections) is almost entirely made up of the MEDEE 2 model (only slight changes have been introduced). The objective of these models is to rapidly calculate prospective final energy balances.

Basically, the activities are either projected exogenously or produced by a macro-economic submodel. The energy intensities, unit consumption and end-use efficiencies are projected on the basis of technological evolution, taking into account the combined effect of energy price evolution and energy conservation policies. On other hand, the fuel mix is projected by end-use on the basis of the historical evolution, taking into account relative energy price evolution and energy substitution policies.

So far, MEDEE 2 has known a very large diffusion all over the world, either under IEJE responsibility (e.g. State of Sao Paulo in Brazil, Ireland, Italy, Sweden, Quebec), or IIASA responsibility, or IAEA responsibility (e.g. Algeria, Indonesia, Pakistan) under the name of MAED.

MEDEE S (Lapillonne, 1985; Dumort and Lapillonne, 1987) has been developed as a substitute to MEDEE 2 for developing countries. It is also an accounting end-use model. But the major difference stands in the fact that, being initially conceived to be applied for developing countries energy demand assessment, it has been designed in such a way that it can cope with most of the very different situations that may be encountered in such countries. This means that:

- the structure of the model is highly flexible in order to be easily adapted to various economic and energy situations and also various other situations in which data and statistics are available;

- most of the situations that can be encountered, in terms of energy end-uses, fuels, economic activities, types of consumers, energy consuming equipment, may be grasped by the model;

- the user of the model is directly involved in the design of the structure of the model, through a didactic procedure; this nevertheless requires from him a basic knowledge of the energy demand mechanisms, of the economic situation and of the available data and statistics.

This great flexibility made it even possible to apply this model in the Southern part of Europe where heating requirements are not too important (e.g. Greece, Spain).

The overall structure of MEDEE is described in Table 1. Most of the flexibility is obtained by the possibility of adding supplementary modules to the standard model that concentrate on specific sectors or end-uses.

Basically, the type of equations or the nature of the input data and results are similar to those of MEDEE 2. However the number of equations or of input data, and the format of the results are not fixed, but depend on what the model user decides while designing the structure of the model. The data base format is automatically produced by the software.

So far, MEDEE S has been transferred to many countries and international organizations, which, usually, further adapted some aspects of the model to their own purposes (in particular some have developed an IBM-PC micro-computer version of MEDEE S).

MEDEE 3

As compared to MEDEE 2 and MEDEE S, MEDEE 3 (Chateau and Lapillonne, 1979; Camos et al. 1986) presents four major differences:

- It is not an accounting model but rather a simulation model, in the sense that the behavior of the energy consumers and of the micro decision-maker in the energy demand field is modelled ("simulated");

- To reach that purpose, it has been necessary to disaggregate MEDEE 3 much further as compared to the previous models to be able to deal with homogeneous classes of behaviors;

- The scenario design process is supported by a systematic and computerized procedure, called "tabular functions", which translate the hierarchical structure of the scenario in a set of mathematical relations;

- Due to the high number of equations (between 1000 and 1300 depending on the version) and to keep flexibility in the design of the structure (for trans-country adaptation within Europe), the model is supported by a PASCAL software called SAMSS which is specially designed for this purpose. MEDEE 2 or MEDEE S are written in Fortran without the use of a specific software. By checking the overall consistency within the model and with the data base, compiling the model, resolving the mathematical equations and editing results at the request of the model user, SAMSS creates the logical sequence of the equations after making any changes in the variables specification (endogenous versus exogenous) or equations

- Control-relations have been introduced within the model, either to indicate major inconsistencies among input variables or to prevent the model from running if major contradictions, physical impossibilities or, economic inconsistencies appear during the simulation.

Thus, one of the difference between MEDEE 3 and the other MEDEE models is that in MEDEE 3 the projection of a part of the socio-economic or technical determinants is no longer exogenous, but results from endogenous simulations:

- simulation of the evolution of the physical output of the large energy consuming industries. (e.g., the stock of dwelling, the stock of cars, the personal and freight mobility, etc.)
- simulation of the choices of households among heating appliances, among modes of transportation;
- simulation of the breakdown of physical output of LECI's among industrial processes, etc.

The general structure of MEDEE is shown on Table 1.

MEDEE 3 has been developed for the European Commission (EC) as part of the energy modelling system of the General Directorate "Research and Development" (DG 12). It was then used for national studies within EC's projects. It has also been adapted by the Ministry of Energy of the Quebec Province in Canada, where it is currently used for long term energy demand planning (MER, 1984).

Table 1
Structure of the MEDEE Model

	STANDARD SUB-MODELS	SUPPLEMENTARY MODELS
INDUSTRY	analysis by branch and end-uses: • motor fuel uses • thermal uses • specific uses of electricity	• steel • energy intensive products (by type of product)
HOUSEHOLD	distinction urban/rural and according to classes: • cooking (and other thermal uses) • lighting (and other electrical uses	• heating • water heating • household electrical appliances • cooling
TERTIARY SECTOR	analysis by sub-sector • thermal uses • specific uses of electricity	• heating • cooling • public lighting
TRANSPORT	distribution between: • passenger transport • goods transport • international transport	• urban transport
AGRICULTURE	analysis by end use or by equipment	

MEDEE 3 ME (Chateau and Lapillonne, 1985) (ME stands for "Maitrise de l'Energie" i.e. Energy Demand Management) has been adapted from MEDEE 3 for the AFME (Agence Francaise pour la Maitrise d l'Energie), mainly for assessing medium and long term energy demand policies issues. In that respect most emphasis has been put on the following issues:

- representation of the behavior of micro decision-maker as regards investments in energy savings or energy substitutions in existing facilities (e.g. equipment, buildings);
- consideration of new equipment and technologies;
- introduction of the key variables, characterizing any energy demand management policy (e.g., prices to the final consumer, subsidies, tax credit, mandatory measures, standard, non financial incentives, etc.)

While MEDEE 3 was a rather large model structure, MEDEE 3 ME has been progressively transformed into a flexible structure, from which the model-user can derive its own version (the various versions of the MEDEE models may be produced from this general model). Thus the model is made of:

- a core-model;
- optional supplementary sub-models that provide an additional disaggregation of some end-use or sectors;
- optional satellite models, in which simulations are carried out.

The core model can be used separately as an accounting model (as MEDEE S). Each satellite model can be run independently to provide in depth analysis of certain issues. The model may also produce projections of input variables of the core model, which would otherwise be considered as exogenous. These models can be coupled and run together.

4. APPLICATIONS OF THE MEDEE MODELS: CONSIDERATIONS ABOUT THE FIELD OF APPLICATION OF END-USE MODELS

The MEDEE models have been applied in very different contexts: OECD countries, centrally planned economies or developing countries in Latin America, Asia or Africa; at a regional, national or international level; and for various purposes including long term forecasting, analysis of specific energy policies, sectoral studies etc. In appendix 1 we indicate the main studies carried out with the MEDEE models, making a distinction between overall long term forecasting studies and specific studies.

Most of these studies aimed at providing indications about the possible level and structure of long-term energy demand to various types of institutions (international organizations such as Commission of European Communities, United Nations-ECE), research institutes (IIASA, Greece, Morocco, IEPE), energy companies (Brazil, Algeria, Indonesia, France) and finally, energy administrations (France, Portugal, Canada-Quebec, Ecuador, Colombia, Mexico, Indonesia, Spain)). In this case, the models were mainly considered as tools of exploration for the future.

In some applications, however, the models were used not only as forecasting tools but also as aids to decision making by providing useful information about the future for decision makers--construction of a refinery (Ecuador) or a large power plant (Indonesia); evolution of potentials of export of electricity (Quebec); identification of energy savings objectives (France).

Looking back to all these studies one may draw some conclusions about the fields of applications of the MEDEE models, and more generally of end-use models.

First of all, these models are well suited to evaluate the possible range of long term energy demand. In particular, and many studies have concentrated on this aspect, they are quite appropriate to indicate the possibility of decoupling energy demand and GDP growth (e.g. the various levels of demand that can be associated to a given scenario of economic growth). They can also provide relevant answers to "what if" questions (e.g. what will happen to energy demand if there are such changes in the industrial structure or if such and such standards are decided for buildings, vehicles or equipment). In addition, they are also very helpful to decision makers each time the decision to be made requires a good knowledge of energy demand over the long term (e.g. construction of refineries, power plants).

In both cases the relevance of the studies will strongly depend on the scenarios considered. For instance, contrasted scenarios, with a low probability will lead to very open energy demand projections, that will not be at all useful. Thus, although the model may be relevant to deal with a wide range of issues, the way they are used is of utmost importance. We will come to this aspect in the next section when discussing the problem of scenario construction.

However, these models also have their limits related to the quantity of data required, to the way price effects are accounted for, to the time-horizon on which they are relevant and also to the overall consistency between all exogenous variables.

The implementation of an end-use model requires a large quantity of data and information which is not systematically available (e.g. economic data, technical data concerning end-uses). This may be considered as an obstacle for the use of such models. One should actually distinguish between the socio-economic data, which usually is regularly collected for non energy purposes (e.g. household appliances, cars), and the technical information which can be estimated or adapted from similar countries or regions (e.g. end-use efficiencies) and checked against the overall statistical information of the energy balance (the model being used to juggle with the values of the different parameters until the balance consumption by sector is returned). In our opinion, difficulties in the access of information should not be used either as an argument against using end-use models or an excuse for using aggregated models. Instead, these difficulties should be the starting-point for data research and surveys, which will enable us to progressively improve the understanding of energy demand, make the use of accounting model more accurate and identify the key energy demand issues.

End-use models cannot directly evaluate price effects and cannot give direct answers to such questions as---What will be the influence on demand of a change in the oil price? (such as in 1973, 1979/80); What will be the influence of a different pricing policies, in many developing countries? However, end use models provide a good framework to contemplate such issues.

Although these models do not have limits in terms of context of application (country, region or even cities), they do have limits on the time horizon. Apart for the econometric models, the end use models cannot deal with the short term and even medium term since in general they do not include determinants or relations that are important to explain variations of energy demand on such horizons (e.g. capacity utilization factor, climate, stocks, etc.).

Another limit is the absence in most of these models of a macro-economic framework assuring a consistency between all economic variables and even between technical variables (e.g. rate of energy savings) and economic variables (e.g. investments). This raises the following contradiction---on the one hand these models are simple, transparent and easy to understand, thus they may be used by a wide range of people. On the other hand, because of the absence of a macro-economic framework, they need a good knowledge of the relations existing between the main energy demand determinants as well as a certain expertise with respect to energy

economics and energy demand analysis. Thus, although they look simple, these models (especially the accounting ones), should in principle only be used by people who have a certain expertise. Improvements of end use models are nevertheless possible on that aspect that will be discussed in the next section.

5. MAIN ISSUES TO DEAL WITH WHILE USING OR ADAPTING AN END-USE MODEL.

As a conclusion drawn from 10 years experience in the application of the MEDEE models, we will draw some general empirical conclusions about how to accurately and properly use end-use models and also how to improve them. Five major issues will be addressed:

- choice of the level of disaggregation,
- accounting for price effects and policy issues,
- accounting for technological changes,
- problems of consistencies in the scenarios,
- the scenario writing process.

Choice of the level of disaggregation.

As already stressed, the choice of the relevant disaggregation level (i.e. energy demand categories or modules) is a central part in the building or adaptation of end-use models.

Two basic ideas underlie the search for disaggregation:

- the more energy demand is disagreggated, the more the consequences of structural changes can be assessed---which is a necessity for long term forecasting since significant structural changes may occur (e.g. change in the industrial structure);
- because of the differences in nature of energy demand determinants and energy demand evolution mechanisms, all heterogenous situations should be grasped separately.

To some extent, these two basic ideas may lead to the wrong statement that "the more disaggregated the model is, the more accurate the forecasts are". However, the search for disaggregation usually meets three main obstacles which temperate this statement:

- the availability of data and information needed to properly run the model and to validate it;
- the more the model is disaggregated, the more complex, time consuming and costly it is to set up a data base to make projections and to finally check the relevancy and the consistency of the results;
- the growing difficulty to ensure the consistency among scenario variables when their number is increasing.

Thus, the choice of the right level of disaggregation will in any case depend on a trade-off between a need for further disaggregation and the possibility and accuracy to do it properly.

As a general rule, the main idea is to first identify which energy demand categories should be considered with respect to the issues that the model should address. Then, the categories

which appear to be the most important, in terms of their weight in the total energy demand (present and future), should be considered separately and the others aggregated. It is much better to correctly assess 90% of the total energy demand and to be rough on the remaining 10% than to try to correctly assess the whole demand at the price of a much higher complexity and difficulty. In this approach, information is not restricted to usual statistics, but rather involves the results of technical studies and punctual surveys and, sometimes, if necessary, expert's guess.

Accounting for price effects and energy policies issues.

It is a matter of evidence that any model designed to assess medium/long term evolution of energy demand should be able to account for the impacts of energy policy changes on demand (e.g. energy price changes, introduction of additional measures to save or to substitute energy).

In end-use models, energy demand is considered a physical resultant of various policy, social, economic and technical factors, including energy prices and policy measures. Thus it is not possible in such models to assess the relation of price quantities by merely examining aggregate price elasticities (as may be done in economic models where energy demand is considered as an aggregate function between quantities and prices). Instead, price effects may be taken into account in end use models according to their particular influence on the various energy demand determinants---short term micro-economic effects on the level of satisfaction of social needs (mobility, comfort, for example); medium/long term effects on technology through technical choices; macro-economic effects through competitiveness of industries, budget constraints of fuel consumers, RD technological programs, etc.

Assessing the influence of energy prices in end-use models raises two types of problems:

- Firstly separating short term from medium/long term price effects can be a problem if short term price effects can be easily identified and measured right after a sharp price increase (like in 74/75, 80-81). It is much more difficult to understand and measure how these short term price effects combine with the medium/long term ones (for example how has the decrease in the unit consumption of cars (medium/long term gasoline price effect) slowed down the decrease in the annual car mileage (short term consequences of price increases in 1974/75 and 1980/81)?)

- More theoretically, the second problem lies in the degree of confidence that anyone may have in the free market (e.g. will the same price increases have the same effect on consumption if they are, or not, accompanied by other policy measures aiming at saving or substituting energy?) Clearly, the implementation of energy conservation policies in most of the countries which experience free market shows that almost no public authorities really believe that only market mechanisms can solve energy problems.

As a matter of fact, accounting for price effects and energy policies issues is not easy in end-use modelling. An optimal combination of expert guess, for technical scenario assumptions, simulation functions and short term price elasticities, should be used. But the main and extremely useful advantage of such modelling is to give the user a very clear and exhaustive understanding of all implications of price variations and non-price policy measures on energy demand evolution.

Accounting for technological change.

Technological change plays an important role in end-use modelling. As opposed to economic modelling where change is taken into account through time related parameters or in terms of changes in production factors combination, which are extremely difficult to assess in the long term, end-use modelling considers explicitly this change at two levels:

- the end-use technologies---the various industrial processes (at least for large energy consuming industries), the various types of building and insulation standards (for the households and tertiary sector), the various modes of transportation, etc.

- the energy appliances---boilers, motors, furnaces, ovens, vehicles.

The technologies are characterized by their unit consumption, while the appliances are specified in terms of energy efficiencies and fuel structure. Such distinction is essential for the following reasons:

- the lead times of technological changes are quite different at both levels. The end-use technology usually changes slowly, with very long lead times, while changes in energy appliances may occur much more rapidly, with lead times not exceeding the life time of such equipment(i.e. between 7 and 15 years);

- moreover, due to their long lead times and the high investment usually involved, changes in end-uses technologies can not be assessed only as responses to energy price changes and must be related to strategies expressed in terms of standard setting, research and development, and industrial development. Changes in appliances may be assessed within a much more classical economic approach, and modelled in that way;

- due to the differences in end-use efficiencies of the various energy appliances, the substitution among fuels may result in a general improvement of the overall efficiency that may be confused with an improvement in the technology itself, if not analyzed separately.

Usually, technological changes are taken into account by combining:

- assumptions of the future technical characteristics of the technologies and appliances that will be put on the market in the future (based on technical studies of the technologies that may be made available in the future);

- assumptions or simulations of the technical improvements of technologies and appliances that are already installed and used, whether these improvements may result from price mechanisms, energy saving policies, or both;

- simulation of the diffusion of new technologies and appliances and of their impact on the average technical characteristics within each energy demand module;

- simulation of the retirement of old technologies and appliances and of their impact on the average technical characteristics within each module.

It should be noticed that this way of accounting for technological changes indicates the time horizon limits for an accurate application of end-use models. Indeed, this limit can be defined as the time delay necessary for a new physical discovery to turn into a wide-spread technology affecting energy demand in a significant way. According to various estimations provided in studies on the technological innovation process, it seems that this horizon lies between 25 and 30 years, but not beyond.

Problems of consistencies in the scenario.

In any type of modelling, the results are always the processed image of the input assumptions and usually have no specific meaning independently of these assumptions. In that respect, the end-use models do not differ from the others. They do present two characteristics which make the scenario writing more difficult:

- as already outlined, they have quite a great number of exogenous variables on which it is necessary to make assumptions. Thus the relation assumption-result is more diffuse and difficult to interpret;

- most scenario variables being linked to each other (non independent), assumptions should be made consistently according to these relations.

If the first problem may be easily solved through sensitivity analysis, the second one remains crucial although it is possible to rely on various additional techniques to improve this consistency (e.g. macro-economical models). Two major questions should be addressed as far as this problem of consistency is concerned:

- what is really meant by consistency?

- what are the implications - or results - of inconsistencies?

In principle, two assumptions are inconsistent if the realization of one prevents the realization of the other to happen. Obviously, this principle can be applied only to variables which are linked through strong causal relations. In most cases, scenario variables are linked with much weaker causal relations (sometimes no causality at all) and these relations can only be described in very qualitative and general terms. Three major fields of consistency may be found:

- the macro-economic consistency of scenario assumptions and of results in respect of macro-economic equilibria in the long term (e.g. external trade, investment saving);

- the consistency of assumptions on scenario variables depicting sectoral policies (e.g. regulation versus economic incentives);

- the social consistency of assumptions on scenario variables depicting the life styles and the income distribution.

What are the dangers of inconsistencies? In principle, an inconsistent scenario is a scenario that cannot take place in reality. The results associated with such a scenario do not correspond to any possible future. But, since the overall consistency can never be guaranteed and end-use models do not intend to predict any future, the question is more to know how far an inconsistency can bring the results away from a possible future than to know precisely how to reach this unforeseeable possible future. Depending on the results depicted by scenario variables, this question may be set in different ways.

In the case of sectoral policies, inconsistencies may be interpreted either as a mis-view of political consistency or as a failure of the policy to reach its goals or to correctly implement its measures. In both cases, it is of the most importance to have a clear idea of the consequences of these inconsistencies since they are part of the process of highlighting the future for decision making;

The lack of a long term macro-economic model clearly indicates the limits of any attempt to check the macro-economic consistency on the long term. But, since end-use models rely heavily on physical quantities and technical factors, the impact of possible macro-economic inconsistencies may not be exaggerated. Let's take industry as an example. Less than ten energy intensive products (e.g. steel, cement, aluminum, paper pulp, glass, ethylene, chlore, ammonia, bricks and tiles) account for more than 70% of the total energy consumption of industry, yet these products account for less than 15% of its total value added. 50% error in the value added associated to the projected physical output of these products will lead to a 7.5% error in the value added of the rest of the industry and a little more than a 2% error in the total energy consumption of industry (7.5% x 30%). In other words, energy demand is more determined by the physical and technical realities of the society than by the way these realities are converted into economic

values and aggregates. If these realities are accurately accounted for, their economic translation may be of minor importance as far as energy demand is concerned.

In case of life styles, it should be kept in mind that the changes are very slow and take place mainly with the changes in generations. Thus it can be assumed that for the 20 to 30 years to come, almost all types of life styles that will play a significant role, already exist and can be specified. Nevertheless, two major types of inconsistencies may remain in life styles assumptions:

- inconsistencies on the possible weight of each class of households with different and antagonist life style;

- inconsistencies between the social and economic behavior considered for certain classes of households and those implicitly assumed in the macro-economic scenario.

If, to some extent, this second type of inconsistencies may be of minor importance due to its negligible macro-economic implications, the first one is more serious. Indeed, the consumption per household may vary by a factor 2 to 3 from one life style to another. Inconsistencies at this level are difficult to check, since the related variables may always be considered from a subjective point of view.

The scenario writing process.

Apart from its role as a consistent framework to provide the necessary input variables of the model, the scenario in the MEDEE approach proved to fulfill two further main functions:

- Being a systematic and didactic method for long term evaluation, the scenario approach forces the user to think about the main events and decisions that may affect the socio-economic development when selecting those having a particular influence on energy demand. This function proved to be essential since one of the main difficulties in assessing long term evolution is to take some distance from the weight of the present and immediate past, which is the necessary condition to shift away from a deterministic view (incompatible with the long term) to a conditional and constructive view;

- The scenario proved also to be an efficient method to identify, classify and specify all types of uncertainties which dominate our appraisal of the long term future and any assessment of the economic risks associated with energy decisions. Apart from the uncertainties due to the insufficient knowledge of the present and historical evolution of energy and socio-economic systems (statistical uncertainties, theoretical uncertainties), two main types of uncertainties dominate our appraisal of the long term. They have different implications as far as decisions are concerned:

  - uncertainties on events that may affect the whole socio-economic systems, but which are not under the control of national decision-makers ("external uncertainties");
  - uncertainties on the important decisions that may be taken by national decision-makers, particularly in response to the external events ("internal uncertainties").

If the external uncertainties cannot be diminished by the scenario writing process, the internal ones, on the opposite, may be reduced, since the scenario may contribute to highlight some contradictions or problems to which the decision may lead.

The selection of scenarios appeared to be always, to some extent, a political issue, at least in those applications which were not purely academic. It always "favors" certain views on the development of the society among various possible, and leads to results which may always be utilized for their policy implications.

Depending on the context of application of the model and the purposes of the study, this policy characteristic of the choice of the scenario may more or less have important consequences. As a general observation, the closer to the decision-making process the model is, the more politically sensitive the choice of the scenario is. As a consequence, the capacity of the model application to influence any decision-making process proved to be highly determined by the political sense attached to the choice of the scenario and also by the personality of the decision maker. As a matter of fact, each time this choice was made by high rank people in energy administrations (Minister Cabinet or Energy Direction), there has not been any problem of policy acceptance of the results of the study.

The choice of the scenario usually resulted from a trade-off between three main objectives which have sometimes been antagonistic:

- selecting the scenarios which seemed the more likely to happen;
- giving priority to scenarios which seemed to have the best political acceptance and which are directly consistent with other political assessment of the future (socio-economic plan for example);
- choosing the scenarios which are supposed to provide more and better information for the assessment of the economic risks of the macro-decisions (in that case, the fact that the macro-decision is or is not taken becomes the basic guideline of the scenarios).

Most of the time, scenarios designed only on probabilities or political criteria proved to be of little help and was also often misleading for energy policy and for energy decision analysis. Probabilistic or political criteria have been short or medium term criteria, dominated by both a deterministic view of the future and an over-reliance on the immediate past. This criteria has not provided possibility for analyzing the behavior of the whole system in response to shocks either external or internal.

In that respect it is a clear that all attempts to caricature over pessimistic or over optimistic scenarios in order to force anyone to "believe" in the mean scenario constitute a perversion of the method by hiding a deterministic view under a false notion of probability or of social acceptance.

REFERENCES

Camos, M. et al. (1986). MEDEE 3, Modele de demande en energie pour l'Europe, Paris: Lavoisier, Collection Tec-Doc.

Chateau, B. and B. Lapillonne (1977). La prevision a long terme de la demande d'energie: propositions methodologiques, Paris: Editions du CNRS.

Chateau, B. and B. Lapillonne (1978). "Long term Energy Demand Forecasting, a new approach," Energy policy 6(2): 140-157.

Chateau, B. and B. Lapillonne (1979). "Long-term energy demand simulation," in Energy models for the European Community, p. 120-128, Surrey: Guildford (MEDEE 3).

Chateau, B. and B. Lapillonne (1985). La simulation des actions et investissements de maitrise de l'energie pour l'industrie et le chauffage des logements dans le modele MEDEE 3-ME, Grenoble: IEPE.

Commission of European Communities (1984). Energy and Development, Paris, Boston: Lavoisier, Tec-Doc.

Commission of European Communities (1985). The MIDAS Energy Model, Bruxelles: E.C.

Committee on Nuclear and Alternative Energy Systems (CONAES) (1978). "US Energy demand: Some Low Energy Futures," Science 200: 142-152.

Dartois, Hourcade, Lagandre (1980). Strategies energetiques alternatives pour le Tiers-Monde et besoin d'investissement, Paris: CIRED.

Department of the Environment and Transport (1981). A simulation model for UK energy demand, London: HMSO.

Dumort, A. and B. Lapillonne (1987). Le modele MEDEE Sud: un modele de demande en energie pour les pays du Sud, Luxembourg: Office des publications europeennes.

Hirst, E., W. Lin and J. Cope (1976). Engineering Economic Model of Residential Energy, Oak Ridge: Oak Ridge National Laboratory.

Instituto de Economia Energetica (IDEE) (1986). Requirements Model, Buenos Aires: IDEE.

International Atomic Energy Agency (IAEA) (1986). Model for Analysis of the Energy Demand (MAED), Vienna: IAEA.

Kouris, G. (1980). "Fuel Consumption for Road Transport in the USA," Energy Economics 2: 31-37.

Kouris, G. (1985). Transportation Oil Demand to 2000, London: The Economist Intelligence Unit.

Lapillone, B. (1986). MAEP Model, International Atomic Energy Agency, Vienna.

Lapillonne, B. (1986). MEDEE 2: A Model for long-term Demand Evaluation, Laxenburg: International Institute for Applied System Analysis, Laxenburg.

Lapillonne, B. (1985). The MEDEE S Approach for Energy Demand Evaluation in Developing Countries, Grenoble: IEPE.

Leach, G. et al. (1979). A Low Energy Strategy for the United Kingdom, London: IIED and Science Review.

Meier, A. K. (1982). Supply curves of conserved energy, Berkeley: Lawrence Berkeley Laboratory.

Ministere de l'Energie et des Ressources (MER) (1984). L'implantation de MEDEE 3 au Quebec: la methodologie et les resultats, Quebec: MER.

Ohman, C. (1980). "Studies of optimal energy conservation programs," ECE seminar modelling studies and their conclusions on energy conservation and its impact on the economy, Washington DC (ECE/SEM3) (EPM model).

OLADE (1984). Metodologia OLADE para el analysis de la demanda, Quito: OLADE.

Parikh, J. K. (1981). Modelling Approach to long-term energy demand and energy policy. Implications for India, Laxenburg: IIASA.

Seeto, C. J. and R.W. Erdma (1980). "Assessment of potential natural gas penetration using a Canadian interfuel substitution model," in Energy System Analysis, p. 201-213, Dordrecht: D. Reidel Publishing Company.

Service de Programmation de la Politique Scientifique (SPPS) (1983). Modelisation technico-economique des consommations energetiques dans le secteur domestique, Bruxelles: SPPS, (Programme national RD Energie).

Skea, J. (1987). "A Simulation model of Interfuel Substitution in the industrial Boiler Market," Energy Economics 9: 17-30.

APPENDIX 1: EXAMPLES OF APPLICATIONS OF THE MEDEE MODELS

1. LONG TERM FORECASTING

FRANCE*
PORTUGAL*
UNITED STATES*
SWEDEN*
BRAZIL*
ALGERIA++
WORLD* (by regions)
GREECE+
CANADA (QUEBEC)**
EUROPEAN COMMUNITY**
SPAIN (BASQUE REGION)+

2. SPECIFIC STUDIES

Assessment of Energy Demand Management Policies

FRANCE**
BRASIL*
UNITED STATES, USSR, EEC COUNTRIES, EASTERN EUROPEAN COUNTRIES*

Evaluation of Long-term Electricity Demand
INDONESIA++

Evaluation of Long-term Refining Requirement
ECUADOR+
COLOMBIA+
MEXICO+

Perspectives of Transportation Energy Demand

MOROCCO+

Evaluation of the possibilities of development of local and regional energy sources in the residential and tertiary sector

FRANCE (Region Provence Alpes Cote d'Azur)**

*MEDEE 2 model
**MEDEE 3 model
+MEDEE S model
++MAED model

Energy Vol. 15, No. 3/4, pp. 279–284, 1990
Printed in Great Britain

0360-5442/90 $3.00 + 0.00
Pergamon Press plc

2.5. PROBABILISTIC FORECASTING

ANDREW J. LIPINSKI

Institute for the Future (ret)
Menlo Park, California
U.S.A.

1. INTRODUCTION

At a risk of repetition, let us restate that the primary purpose of forecasting is to present the decision makers with information about the future. All interesting things about the future are uncertain and unpredictable. When a forecaster attempts to predict the essentially unpredictable future, what he or she really means is that they have selected a particular set of outcomes from a much larger probabilistic domain of outcomes. Their hope is that this domain has a central tendency (the outcomes near the center of the distribution have a higher probability than those near the edges) and that they have selected either the most likely or the expected value of the distribution (the two coincide if the joint distribution is symmetrical).

Do the forecasters know the shape of the distribution? Sadly, the answer is that they have no idea because they have never checked. If they did check, say the probabilistic forecasters, they would find out that the distributions are most often very "untidy", with several peaks of almost the same probability. Small changes in probability assessment are likely to make the most likely value jump from one peak to another. No wonder then that deterministic forecasting is often such an act of faith. Worse, it can seriously mislead the decision makers who interpret the narrow forecast as evidence of conviction about the values contained in the forecast.

Probabilistic forecasting attempts to communicate to the decision makers the information about the characteristics of the domain of outcomes, such as its spread, shape and other measures. In doing so, probabilistic forecasting encounters two major obstacles: the difficulty of eliciting probability assessments and the difficulty of communicating the results of the forecast to the decision makers.

Rarely can probabilities be assessed from historical data alone. In most cases the assessment comes from human, subjective judgment by individuals or groups.

The following example of probabilistic forecasting illustrates the use of group judgment in assessing probabilities of future events and functional relationships between future uncertain quantities.

Several models were used, the first one probabilistic, a static probability tree, modified to incorporate binary information about three time periods.

As in most probabilistic modeling, building a probabilistic model was merely the first step: the probabilistic model had to be followed by another: complex, deterministic and dynamic. The "timing" information included in the probabilistic tree was used to select appropriate "time shots" from the dynamic forecasts and the outcomes of the tree were used to generate the values of the exogenous variables driving the deterministic, dynamic model.

We will first describe the methodology of probabilistic forecasting and then show how the probabilistic model was used in conjunction with other models. Finally, we will comment on the scope off the effort and suggest ways of communicating the results to the decision makers.

2. METHODOLOGY

The most commonly used probabilistic model is the probability tree shown in Figure 1.

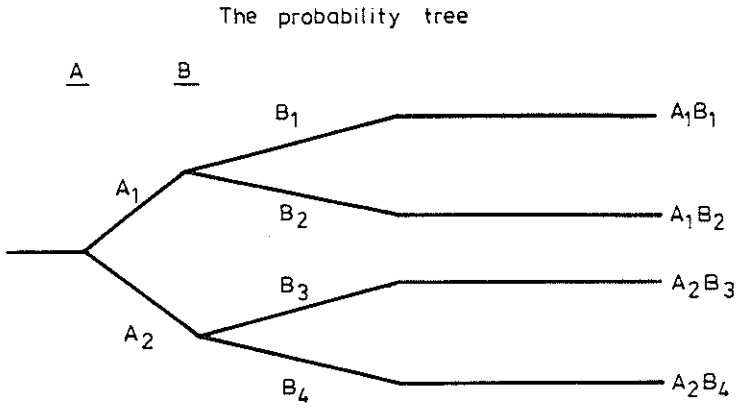


Figure 1. The Probability Tree

Using a probabilistic tree model presents a special set of challenges. Variables are arranged in sequence, the first one being the "truly" unconditional one (in reality all probability assessments are conditional) -- that is, the probability distribution of the first variable is unaffected by the values of the remaining variables. The probability distribution of the second variable in the sequence is conditional on every discrete value of the first variable. For example, the probability distribution of inflation might depend on the value of GNP growth. High GNP growth might raise the value of inflation but the actual value of inflation would still remain uncertain. Its general level would rise, and, depending upon the value of GNP growth, the shape of the distribution would also change. For example, very high growth might result in a higher probability of runaway inflation, a peak at the right hand side of the probability density distribution.

It is immediately apparent that a lot will depend on the sequence. Deriving a satisfactory sequence is a matter of judgment of what depends on what. Not only that we can't be quite sure of what the sequence ought to be, but many situations are circular, that is the GNP could also be dependent on inflation rather than simply driving inflation. One can see that there could be many sequences, with probabilities assigned depending on one's belief of how economics or some other mechanism actually works, such as the relation of our own sales to the prices charged by the competitor.

Then there are computational and elicitation consequences. How many values of the first distribution should be used to assess the second distribution? Let us assume only three, the equally likely values (Amara and Lipinski, 1983). Even then, we must obtain three distributions of the second variable and, if we continue the division by three equally likely values per distribution, we need nine distributions of the third variable, twenty seven of the fourth and so on.

One can see that probabilistic modeling requires careful thinking about the variables that are really essential lest the elicitation effort and expense overwhelm the analyst. An example of a "frugal" tree, that was part of the modeling effort described next, is shown in Figure 2.

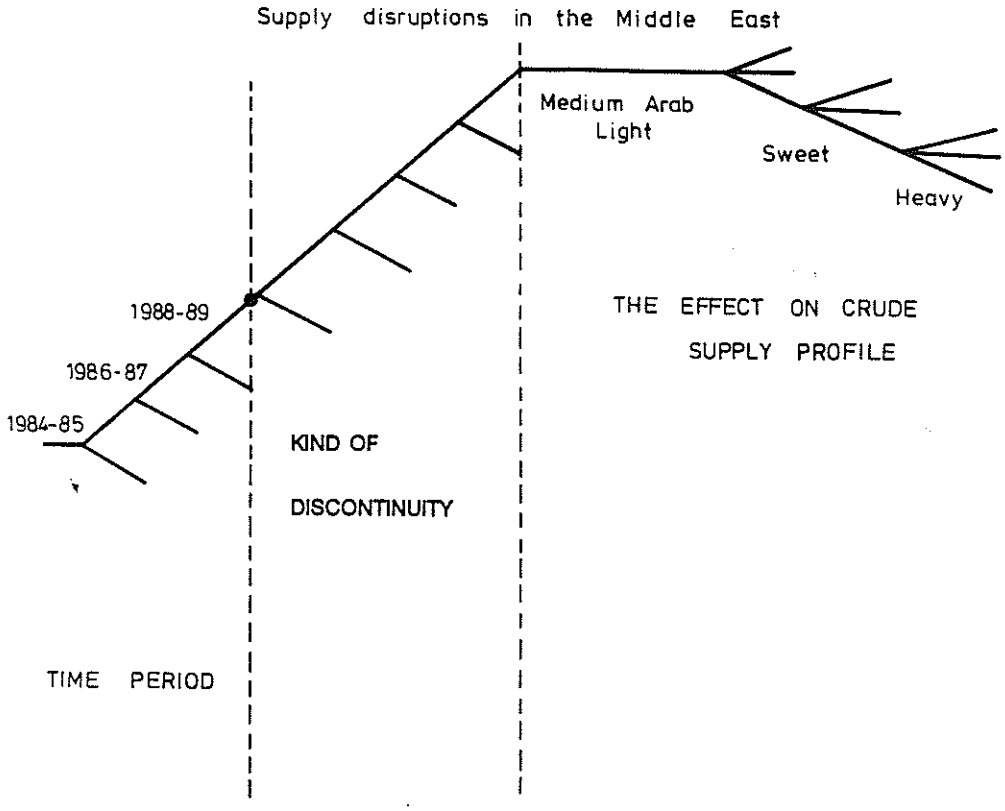


Figure 2. Supply Disruptions in the Middle East

3. USE OF A PROBABILISTIC MODEL IN FORECASTING

The objective of the forecasting effort was to provide decision makers with information about future price differentials between heavy and light crude and between spot and contract prices for a "reference" crude (Brent Light).

In the consensus of political and oil industry experts participating in the construction of this model, five possible events would adequately account for changes in the supply profile of Medium Arab and Mexican Light, the Light and the Heavy crude oils. These were:

-- Cessation of the Iran-Iraq war

-- Damage to Gulf facilities (including the pipelines radiating from the Gulf)

-- Breakdown of OPEC discipline

-- Revolution in Nigeria or Libya

-- Restriction of output by Saudi Arabia.

The timing of these events was considered important because of the changes in the supply demand situation within the 1984-89 time frame. Consequently, we divided the 1984-89 time frame into three intervals: 1984-85, 1986-87, and 1988-89. In each time period, a disturbance (from any source) in the oil production profile could happen, or not. To qualify as a disturbance, a change in one of the three crude production levels had to exceed 300,000 barrels per day for a period of nine months or more.

The task of the probabilistic model was to project the supply levels of heavy, medium and light crude, taking account of these developments in the Middle East. The inputs to the probabilistic model were the assessments of political and oil industry experts. The former assessed the probabilities of changes in the political situation and the latter translated these changes into changes in the supply profile. The experts were interrogated individually but the judgments of the two panels, political and oil industry, were averaged, using weighing based on self-calibration (Amara and Lipinski, 1983), Table 1 and 2 in 3.6.

The reason for relying on group judgment was the realization that no single individual has adequate knowledge to quantify the entire complex, fluid and uncertain future of oil prices or, indeed, of any comparable situation in energy or any other field. In such a case the principal task of the modeler becomes the building of an explicit framework in which such expert judgment can be embedded. In this case, the framework was the probability tree shown in Figure 2.

The advantage of having an explicit framework (contrasted to another possible approach such as a set of interviews followed by a summary) is that it allows for subsequent revisions. Each expert contribution "fits" in a precise place in the structure. If the contribution is made by a group of experts, rather than by a single individual, the method of arriving at the joint judgment is also explicit so that it can be updated as new experts are brought in. After updating, the final result can be recomputed. For example, the new result would be a fresh set of probability distributions of the supply profile.

The user of the forecast often wishes to know how sensitive are the results to the judgments of the participating experts (in our case the results are probability distributions). In our experience, having a panel of a dozen experts makes the results insensitive to a single assessment. We can also add that the averaged judgments of another panel, elicited with an uneventful time interval, would be quite similar. However, many situations, and certainly the situation in the Middle East, are extremely fluid. A second expert panel may possess information that the first panel did not have (for example, that the Iran-Iraq war has ended). We should not be surprised when their assessment is different. Forecasting results are the best information available at the time and when the situation changes, the information should be updated.

Figure 3 shows the block diagram of the entire forecasting effort.

Three types of models were involved in the effort described here: 1) a probabilistic representation of the Western World oil supply affected by the developments in the Middle East, 2) a deterministic portrayal of the interrelations between major Western World energy-related variables, such as demand for crude, crude supply or prices of petroleum products, and 3) a simple linear regression model to represent the results of model 2).

The deterministic model, DYCIM, belongs to the "cross-impact" family of structural models (Lendaris, 1980 and section 3.6).

The probabilistic model of the Middle East provided thousands of different sets of production levels of the three crudes, depending on what has happened in the Gulf. The DYCIM model, on the other hand, provided a forecast of prices for a set of production levels over time. How were the two models combined?

Ideally, the DYCIM model should be run once for every end branch of the probabilistic tree. But that would be too expensive and too time consuming. Instead, the output of the tree was sampled and the DYCIM model was run once for every thirty or so sampled production levels. From the DYCIM outputs we extracted three time snapshots of price differentials corresponding to production levels in the three time periods specified by the probabilistic tree.

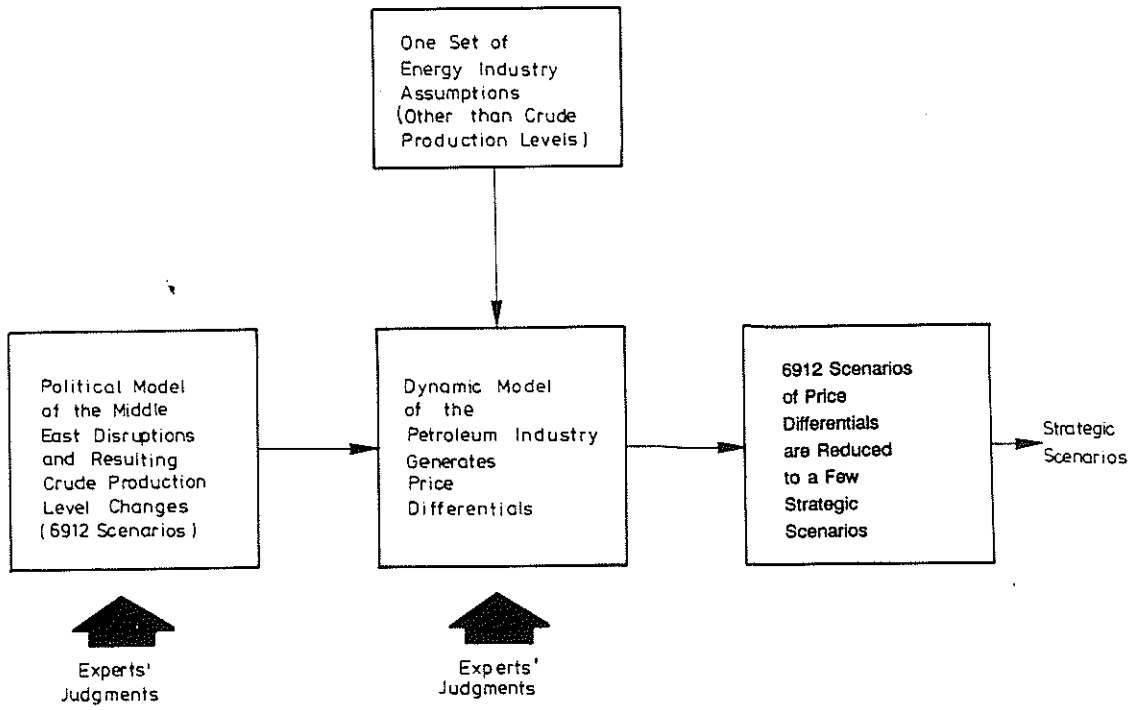


Figure 3. Block Diagram of the Forecasting Effort

For each time snapshot a simple regression model was run with the appropriate price differentials (light to heavy or spot price to contract price) as the dependent variable and the three production levels as independent variables. Once this was done, the DYCIM model could be replaced by the regression equation which became the "surface model" of the DYCIM.

There was no problem in running these six equations (two price differentials times three time snapshots) for all end branches of the tree. The outcome corresponding to each branch contained: 1) events sequence, 2) probability of that events sequence, and 3) the corresponding price differential. Sorting the outcomes by the value of price differentials allowed construction of the probability distribution of price differentials at any of the three time frames.

Three groups of 12 experts each participated in the effort, the last one being involved in the DYCIM workshop. One computer program dealt with the computation of conditional probabilities, the events sequences and the output level of crude. Other programs, using algorithms provided by our expert consultants, computed the interactions between crude oil producers of the three grades of crude, and processed the probability distributions. In all, 50 man days of expert time were used (a considerable expense) in addition to the analyst's time of six months.

4. COMMUNICATING RESULTS TO THE DECISION MAKERS

In the final step, that of communication the results, the probability distributions were translated into "metascenarios" (Amara and Lipinski, 1983).

Metascenarios are bundles of outcomes or scenarios, and are therefore fewer in number than the number of scenarios. For that reason, they are easier to communicate. Our preference is for the "equally likely" metascenarios -- typically, two or three, that manage to capture most of the uncertainty and can serve as input to strategy determination or operational decision making.

An example of the "Christmas tree" of equally likely values of price differentials is shown in Figure 4.

Figure 4. Equally likely Price Differentials Time Frame: 1990; 1984 Dollars

REFERENCES

Amara, R., and A. J. Lipinski (1983). Business planning for an uncertain future: Scenarios and Strategies, New York: Pergamon.

Lendaris, G.G. (1980). "Structural modeling - A tutorial guide," IEEE Transactions on Systems, Man, and Cybernetics SMC-10: 807-40.

Energy Vol. 15, No. 3/4, pp. 285–296, 1990
Printed in Great Britain

0360-5442/90 $3.00 + 0.00
Pergamon Press plc

2.6. DEMAND FORECASTING METHODOLOGIES: AN OVERVIEW FOR ELECTRIC UTILITIES+

AHMAD FARUQUI

Barakat, Howard & Chamberlin, Inc.
Oakland, California
U.S.A.

THOMAS KUCZMOWSKI

Energy Modeling Forum, Department of Engineering-Economic Systems, Stanford University
Stanford, California
U.S.A.

PETER LILIENTHAL

Energy Modeling Forum, Department of Engineering-Economic Systems, Stanford University
Stanford, California
U.S.A.

Abstract - Prior to 1973, most industries--including electric utilities--forecasted growth using rather straightforward time-trend approaches. While these approaches served the industry well during periods of steady and rapid growth, they failed to capture the underlying causal factors of growth. Thus, whey were unable to predict or explain the sudden changes in growth rates that have occurred since the 1973-74 oil embargo. Therefore, more sophisticated econometric and end-use models forecasting techniques were introduced to the utility industry. The growing awareness of the implications of a finite fossil fuel supply, as well as the increases in electricity rates, led forecasters to consider electricity prices and conservation issues explicitly in their models. In the industrial sector, changes in technology, structural changes in the economy, and fuel switching issues have led to an interest in process models.

The current trend in utility forecasting is toward strategic forecasting models. A key characteristic of strategic models is their ability to examine explicitly the factors and issues affecting future growth. This implies combining elements of the econometric approach with the technology detail found in the end-use/process models. Strategic models must be capable of doing more than merely forecasting future requirements. They must be able to provide planners with additional information on which policies to pursue to shape future loads.

1. ISSUES IN FORECASTING

Today's energy planning environment, particularly as it relates to industrial customers, is characterized by a number of important issues. First, estimating the future economic strength of the manufacturing industries is a key challenge to the industrial forecaster (1). Second, the purchasing patterns of industries' clients have changed. The advent of light weight, high performance materials is changing structurally the traditional intermediate materials sectors away

+An earlier version of this paper was presented at International Atomic Energy Agency course on Demand Forecasting at Argonne National Laboratory on September 21, 1988.

from firms producing basic metals to those producing composites, engineering ceramics and plastics, and specialty metals. Similarly, the advancement of electronics is rapidly changing the competitive advantage of traditional consumer product manufacturers in favor of those firms who have capitalized on new found consumer desires, and on firms which can effectively meet the expanded information needs of a dynamic economy. Finally, developments in new technologies are creating new markets. Thus, utilities must be able to assess the competitive position of industrial firms in their service area as well as that of the products those firms are producing. Only by examining both of these factors can a reliable, long-term forecast of electricity consumption be made.

The complexity of forecasting is further increased through the emergence of new technologies, particularly electricity-intensive technologies. As new technologies are introduced, the demand for inputs changes due to the requirements of the new technology and changes in relative competitive position of that industry as a whole. The key questions associated with new technologies are which ones will be adopted by industry and how will their adoption affect the input mix (2).

Since the early 1970s, economic, political, social, technological, and resource supply factors have combined to change the operating environment. As far as the electricity markets are concerned, many utilities are implementing demand-side management (DSM) programs. DSM consists of planning and implementing activities designed to influence customer use of electricity in ways that will produce desired changes in the utility's load shape, i.e., changes in the time pattern and magnitude of a utility's load. The approach is in contrast to traditional supply-side planning where utility capacity was built to satisfy a given level of demand. Many utilities are actively pursuing DSM programs (3). Figure 1 provides a framework for analyzing these issues. The diagram presents the scope of the INDEPTH modeling system which is described elsewhere in this chapter. Thus, an approach is required which can capture the influence of numerous, closely related issues: international competitiveness, product competition, consumer demand, technological change, and DSM programs. See (2) for additional perspectives.

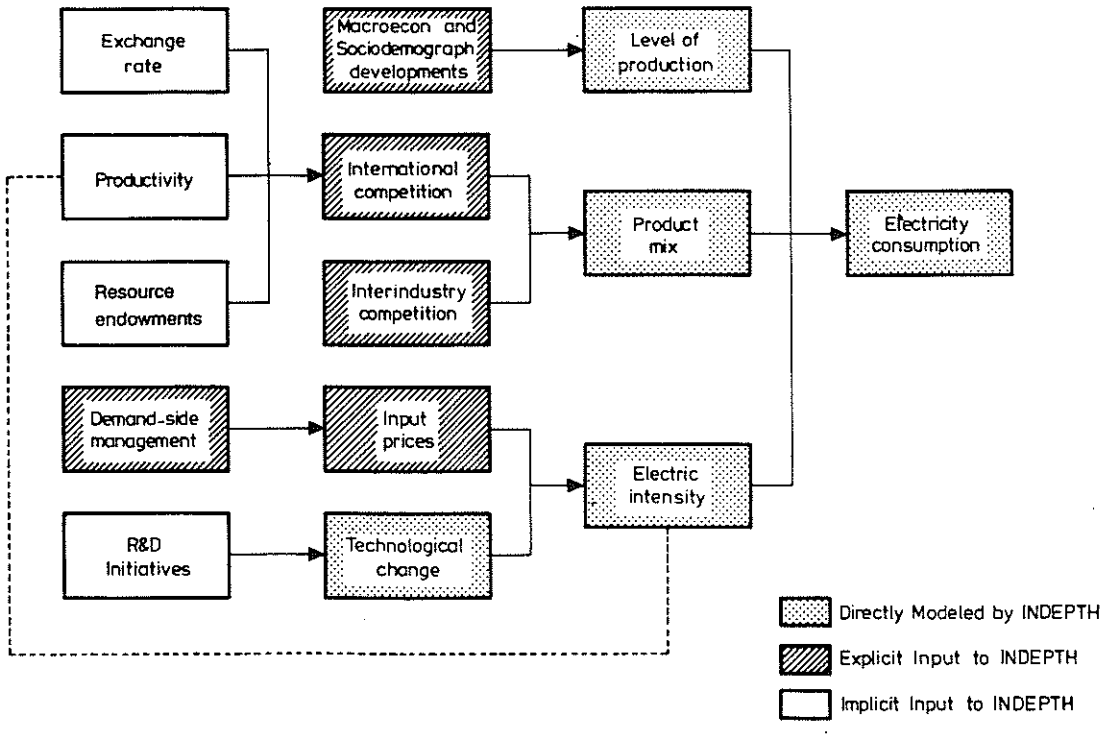


Figure 1. Structure of Industrial Electricity Demand

2. EVOLUTION OF FORECASTING METHODOLOGIES

Prior to 1973, most industries, including electric utilities, forecasted growth using rather straightforward time-trend approaches. While these approaches served the industry well during periods of steady and rapid growth, they failed to capture the underlying causal factors of growth. Thus, they were unable to predict or explain the sudden changes in growth rates that have occurred since the 1973-74 oil embargo. Therefore, more sophisticated econometric and end-use models forecasting techniques were introduced to the utility industry. The growing awareness of the implications of a finite fossil fuel supply, as well as the increases in electricity rates, led forecasters to consider electricity prices and conservation issues explicitly in their models. In the industrial sector, changes in technology, structural changes in the economy, and fuel switching issues have led to an interest in process models. The evolution of forecasting approaches over time is illustrated in Figure 2.

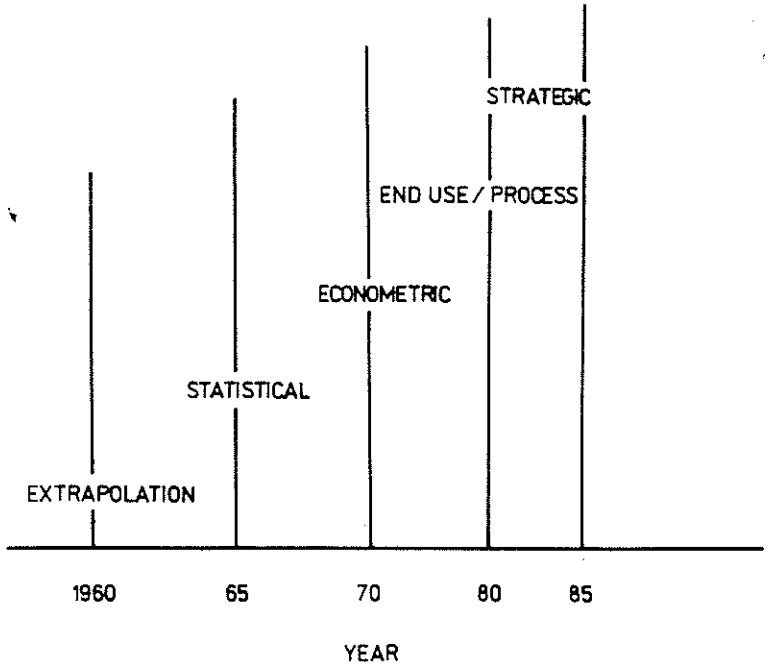


Figure 2. The Evolution of Forecasting Approaches

The current trend in utility forecasting is toward strategic forecasting models. A key characteristic of strategic models is their ability to examine explicitly the factors and issues affecting future growth. Strategic models recognize the impact that policy decisions can have on future loads. This requires detail--detail on the customers' operations, their current and potential demand for electricity, their competitiveness in the market place and their options with respect to production processes, switching alternatives, etc.

In the industrial sector, this implies combining elements of the econometric approach with the technology detail found in the end-use/process models. Strategic models must be capable of doing more than merely forecasting future requirements. They must be able to provide planners with additional information to help shape future demand. None of the modeling methods overcomes the uncertainty associated with forecasting.

3. CURRENT STATE OF PRACTICE

As noted, the methods used in forecasting have been changing rapidly since the early 1970s. Key findings from surveys of utility forecasters indicate:

1. Utilities spend significant resources on energy forecasting.
2. There is a growing desire to incorporate more detail on electricity use into the forecast.
3. Information on electricity consumption by industry type, customer equipment, and operations are generally available, although the quality of the information and the time frame over which it is available vary widely.

The overwhelming majority of utilities currently use an econometric approach to forecast industrial energy consumption. Survey results indicate that a combination of approaches is often used. In the sample of 14 utilities:

- Three use an aggregate econometric approach; that is, they produce forecasts for the industrial class as a whole.
- Ten use a detailed econometric approach; that is, the industrial class is disaggregated according to such factors as SIC, customer size, and growing industries and declining industries.
- One uses a customer survey as the sole basis for the forecast.
- One uses a customer survey to supplement other econometric forecasting methods.
- Two use end-use/process models for forecasting electric consumption of the major industries.

Because some utilities use more than one technique, the sum of the above responses exceeds the number of utilities interviewed.

The use of formal questionnaires to obtain information relevant to generating the industrial forecast is the exception rather than the rule. Only three of 14 respondents collect and incorporate customer information into the forecast through use of a formal questionnaire. The far more common practice is to collect customer-related information informally through customer representatives. Because industrial forecasts are typically generated at a more aggregate level of detail than the specific customer information, such information is incorporated into the forecast in an ad hoc manner. That is, the information is used judgmentally to adjust the statistically derived electricity consumption level up or down as appropriate. The practice of interviewing industrial customers, rather than relying on surveys, is quite a contrast to the residential sector where appliance surveys are conducted regularly by the industry.

The following information was collected regarding the overall familiarity of the utility industry with process models that estimate energy consumption by focusing on the underlying industrial processes and technologies. The familiarity of the respondents with the concept of process models is as follows:

- Twelve have gained familiarity with the models through seminars, conferences and/or the literature.
- Two used the process models in the preparation of their industrial forecast.
- Four expressed the opinion that process models are too data-intensive and cumbersome for use as forecasting tools at the service area level.
- Two have no familiarity with process models.

In addition to technology, industrial energy consumption is influenced by the level of economic activity. Economic drivers used by the respondents include:

- Final Reserve Board production index.
- Value of shipments.
- Service area employment.
- National/regional production levels for important industries in the service area (number of automobiles, tons of paper, etc.).

- Gross state product.

All but one of the respondents indicated they used outside economic forecasting services to provide them with at least the starting point for generating the drivers for their models. Four indicated that they had in-house economic services that regionalized or adjusted national forecasts to reflect service area conditions.

A common concern in the development of an industrial forecasting model is the availability of data, particularly with the increased tendency of the industry to incorporate more detail into the forecasting models. Of the 14 utilities interviewed, 12 classified their customers, and consequently, industrial electricity consumption, by at least the two-digit SIC level. Some of the customer classifications go back as far as the 1960s, others only to 1982. There is some doubt about the validity of the data from the earlier years due to misclassification, changes in customer product mix, and changes in the SIC categories.

In addition to consumption data, some models, such as process models, require information on industrial customer processes and equipment. One-third of the respondents indicated that their customer representative had good knowledge of their customer's electricity consuming equipment. Another third indicated that they obtained similar information through recently conducted surveys. The remaining one-third indicated that they had little knowledge of their customer's equipment and processes.

All but one respondent indicated a general level of satisfaction with their present forecasting procedures. This, however, does not imply that all of the forecasters' needs are currently being met. Specific needs, as expressed by the forecasters, include:

- More detail on the customer's operations to assess his business competitiveness.
- More detail on the pattern of consumption.
- More detail on consumption and end-uses to be used for marketing purposes.
- Better understanding of the impact of conservation on customer's consumption.
- Better understanding of the effect of technology changes on industrial consumption.
- Increased accuracy of economic variables driving the forecast.

Regarding DSM, it is interesting to note that for most respondents DSM in the industrial sector was limited exclusively to rate-related programs. Specific demand-side programs mentioned include:

- Time-of-use rates.
- Off-peak energy rate.
- Interruptible rate.
- Special metal melting rate.
- Economic development rate.
- Industrial customer audit program.
- Financial incentives for the purchase of energy efficient equipment.
- Educational activities on the efficient use of electricity.
- Marketing of conversions from fossil fuels to electricity.

Utilities devote considerable resources to the system-wide forecasting activity. The total level of effort devoted to forecasting residential, commercial, and industrial energy consumption, as well as the economic drivers, ranges from 2 to 40 person-years per year for the utilities interviewed. The average level of effort is 11 person-years per year.

4. SYSTEMS OF MODELS

A detailed review of the strengths and weaknesses of the alternative modeling approaches is described in (4). Based on this review, a hierarchical, menu-based approach appears well suited to meeting utility needs.

Such an approach is adopted in the INDEPTH modeling system, described elsewhere in this chapter. The concept is however perfectly general. The system should have some general properties such as the following. It should form an internally consistent, nested modeling system that would enable each utility to select an appropriate level of detail based on the availability of data and the specific characteristics of its service area. System of models consist of several levels of analysis. The broadest level will normally consist of a series of econometric models that produce electricity forecasts for the 20 two-digit major industry groups such as food products, lumber products, and pulp and paper products. These econometric models will provide complete coverage of the manufacturing sectors. The second level of models will focus on a few key industrial sectors through the use of process models. These models provide greater detail about the industry being modeled, but cannot, by themselves, be used to generate forecasts for the total manufacturing sector because only a few sectors are modeled. The third level of models is the equipment models. These models are designed to forecast change in market penetration of key end-use equipment, such as motors.

In the integrated modeling system, all the component models need to be supported by a common set of assumptions, external drivers, and a consistent logic. These common elements are necessary because the forecasts obtained from the second level--process models--are implicitly contained in the forecasts produced by the first level--econometric models. Similarly, the results of level three--the equipment models--are implicitly contained in the econometric and process models. Therefore, all levels of the system utilize the same macro-level economic forecasts to assure that the electricity forecasts are based on common perceptions of the underlying economic growth of the service area, of the structural effects on industry sectors, and of technological change. Therefore, a critical component of the modeling system is the data base and the methods to regionalize economic forecasts.

5. SELECTION OF MODEL SPECIFICATION

Selection of the particular model specification is based on satisfying several objectives. It is necessary to strike a balance between model specifications that are abstract, complex, theoretical, and difficult to implement with those that are overly simplistic, possess little statistical and theoretical justification, and are being used because of the expediency of implementation. The following objectives need to be considered:

- Validity--Is the model conceptually and theoretically sound? Are the relationships, specifications, and data accepted by the professional community? Do models of a similar nature possess a track record of accurate forecasting?
- Transferability--Is the model capable of being specified, estimated, and implemented in a way that makes it applicable to a variety of areas?
- Data availability--Are the data available for model estimation? Is the quality of the data sufficient to make key planning decisions?
- Output capabilities--Is the model capable of being formulated in such a way as to satisfy the need?
- Ease of understanding--Can the model structure and implementation be understood by critical decision makers who must have confidence in the model and its results to make decisions?
- Issue handling--Does the model allow examination of critical issues facing forecasters? Can the model address any or all of the following:
  - Structural change in product demands
  - Regulation
  - Technological change
  - Demand-side management issues such as rates, incentives, and education
  - Alternative input prices

A classification scheme useful for describing the trade-offs between the level of detail contained in models versus their ability to explain is shown in Figure 3.

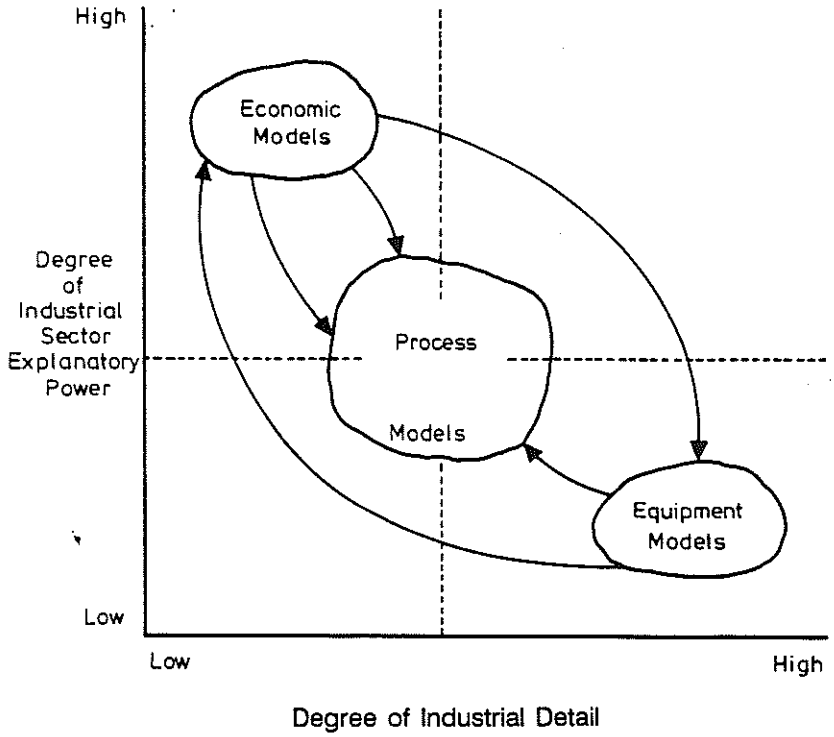


Figure 3. Positions of the Different Forecasting Models in Given Taxonomy

6. APPLICATION OF MODELS

Equipment Model Applications

The equipment models are used to examine the impact of technological change in key equipment common to most manufacturing establishments. Examples include electric motors, process heating equipment, and lighting fixtures. These models are run to analyze the potential impacts of DSM programs designed to promote the use of certain types of end-use equipment.

Equipment models when used for electric utilities will produce results in the form of electricity consumption by motors, electric lights, etc. The consumption of energy by this equipment will, however, already be contained within the forecasts made by the econometric and process models. Therefore, sector totals for each industry cannot be obtained by merely adding up the consumption levels forecasted by equipment models.

The implementation of the equipment models, therefore, requires two steps: (1) identification of equipment usage in the industrial sectors of other model levels, and (2) utilization of the equipment models in a simulation mode.

A bridge is built that describes the baseline level of equipment usage in each industrial sector in the econometric and process models. The bridge allows identification of the implicit baseline equipment consumption. By utilizing the equipment models in the simulation mode--with the baseline being the level of usage described in the bridge--the net change in equipment usage can be added to any other model result.

Econometric Model Applications

The econometric models are the most general and comprehensive of the models. Built at the two-digit SIC level, these models form the basic forecasting capability of most modeling systems. Through simple summation, the aggregate manufacturing consumption can be determined. All other models within the system modify and supplement the results obtained from the econometric model.

The macro-driven, service area economic forecasts, and decision-making assumptions used in the econometric models must be consistent with those used at the other levels. Only if the other levels are used in a simulation mode can different data or assumptions be used--and then, the results from the different data must be compared to baseline results from the data used in the econometric models.

Thus, the econometric models are critical tools for forecasting. Utility-specific applications will use the elasticity estimates derived from the system of cost-minimizing input demand models in single-equation models after regional calibration. The process and equipment models are used to analyze in greater details results which are implicitly contained in the econometric models.

Process Models Applications

The process models are for use by utilities interested in better understanding how a particular four-digit industry may evolve over time. Thus, the process choice models are used only if that industry segment is important in a local service area.

The process models yield a forecast of electricity consumption at a more detailed level than is produced by the econometric models. Therefore, the results of the process models cannot be added to the results of econometric models without careful consideration of possible double counting. The process models can be integrated with the econometric models as follows. Two sets of econometric models are specified. One for the set of entire two-digit SICs and another based on two-digit SICs except for the four-digit sector being modeled by the process models. For example, if a process model is built for SIC 3311--basic steel--then a special econometric model containing all activity in SIC 33 except 3311 is also built.

Econometric Versus Process Models

Application of econometric and process models is outlined on Figure 4. The driver model operating at the bottom level generates input (production mix) in monetary units. Therefore, calibration is required for the process model into the physical units.

Both econometric and process model have their specific data bases as listed in the figure. Behavioral lags representing an estimate of time required to achieve the equilibrium market share are often included in process models.

Model outputs can be described briefly as future demand trends from the econometric model and technology choice from the process model (as an optimization algorithm is applied).

Mixed Models Application

In the middle the mixed models are placed as described before. This type of modeling can use a combination of orthodox generic approaches described before for econometric and process models.

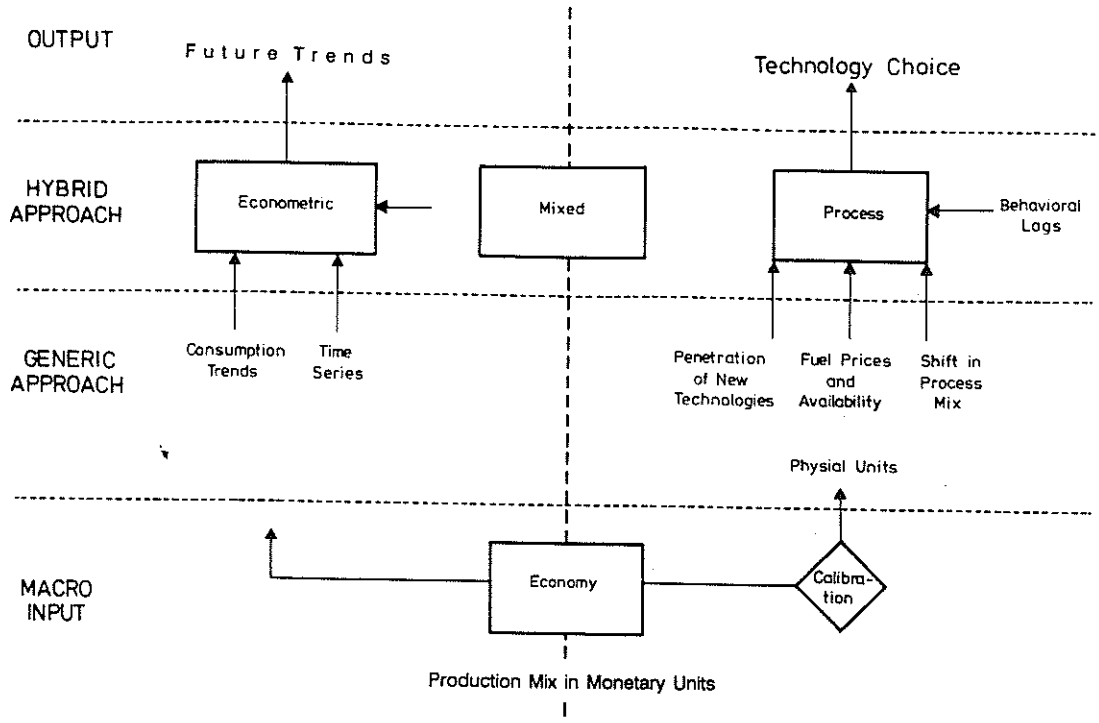


Figure 4. Econometric Versus Process Models

The mixed approach might be just one model, called hybrid in our taxonomy. It can also be a system of models, which can use separate econometric, process, and other types of models. INDEPTH is an example of such an integrated modeling system. In the most general case, it can be a set of different type of models employed to generate the demand.

Such a model set was used in an Energy Modeling Forum study to evaluate the impact of structural shift in the US economy on industrial energy demand. On the macroeconomic level, three economy-wide models were employed. To predict the major trends in the US economy, the PILOT model was used (see Chapter "Multisectoral Planning Models"). This is a macroeconomic model designed for long-term projections of both energy demand and economic output. The PILOT model is a large-scale dynamic input-output model of the US economy. The main goal of the model is to evaluate technological change, which is assumed to be the driving force of economic growth. The INFORUM model, an input-output macroeconomic model, was used for short-term projections only (i.e., through 1995).

Finally, the WHARTON model was used as a driver for industrial models. WHARTON generated sectoral outputs, which were used to forecast industrial energy demand. The PILOT model was calibrated to WHARTON's base-case scenario. It was impossible to use the PILOT model to generate energy demand by industrial sectors, as the aggregated energy services submodel is employed.

All the types of models represented in this study can be used to forecast the levels of demand for different energy carriers. However, the three models described above were designed generally to describe the macroeconomy. Therefore, to conduct the detailed research, the sectoral output levels obtained from these macroeconomic models were introduced as drivers to three different types of industrial sector models:

- Process (ISTUM)
- Hybrid (ORIM)
- Econometric (PURHAPS)

Observation on Models Behavior or Response

As a result of this comparative study, some observation on models behavior can be made. Preliminary observations were that the choice of method applied can bias the results. As a rule, the higher growth in demand level, especially for electricity, was obtained from the econometric model. The nature of the bias in the process-type models is to be overly optimistic on the rate of efficiency improvement in the input-output coefficients. The rationale of the econometric models is that our best guess of the future is provided by a study of past trends. Therefore econometric models can easily capture long-term regularities such as electrification and automation if they are contained in historical data.

During the process of detailed comparisons, it was found that differences in model outputs were not very large, after proper calibration. There are, however, some aspects of model types which might be of some interest for users.

The penetration of new technologies can be an extremely important factor in the future. This issue can be examined only by engineering process models.

Econometric models tested in the EMF 8 study show surprisingly low sensitivity to changes in the input prices, especially oil prices. This might be explained as a result of the relatively long time periods used by the models.

All models have built-in tendencies that favor or discourage use of certain inputs. All models show a trend toward electricity with virtually no substitution away from electricity toward fossil fuels. Process models display a tendency toward coal, possibly due to lower nonfuel costs of coal technologies.

In general, econometric models are highly dependent on time series data and are best suited for short-term projection. As a consequence, econometric models may not be appropriate in forecasting the effect of new technologies that depart from the historical pattern.

An excellent illustration of the difference between econometric and process models is afforded by the forecasted role of cogeneration by the models participating in the EMF 8 study. In Figure 5 both of the process models show declining intensity of purchased electricity in manufacturing while the other models show it to be increasing. This discrepancy was primarily due to self-generation.

The long term trend for cogeneration and self-generation had been declining for almost a century and the recent surge in activity had not yet been reflected in the data available to the modelers at the time of the study. Consequently, the forecasts of the econometric models reflected very little cogeneration activity in the future.

Because the process models base their forecasts more on future conditions than past trends, the favorable economics of cogeneration resulted in a very substantial penetration of this technology. This was the greatest single difference in the results of the two modeling methodologies.

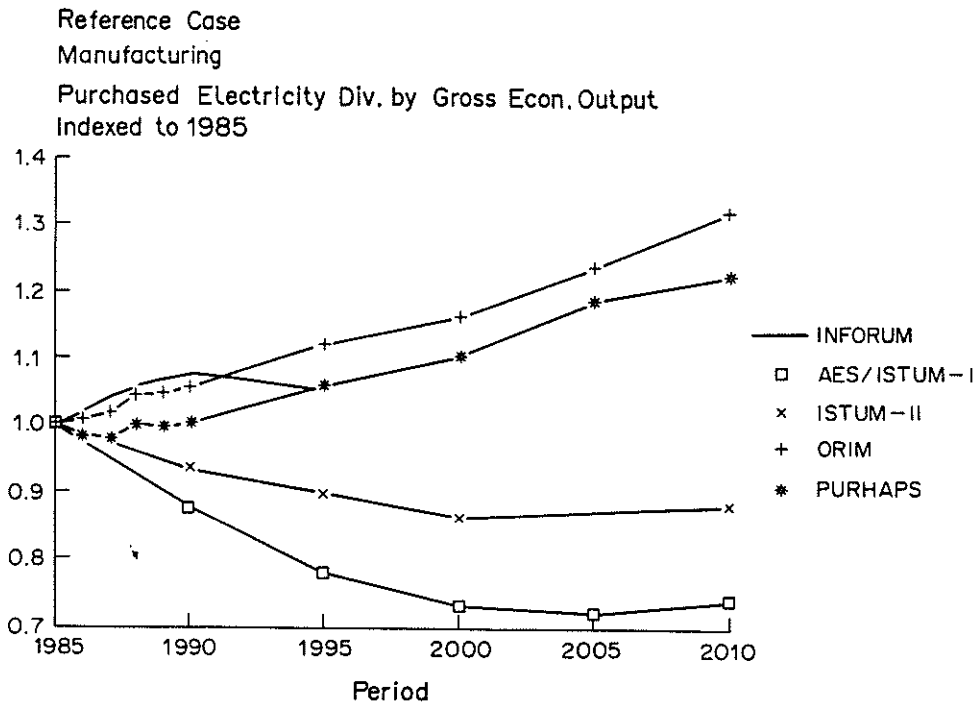


Figure 5

The inherent optimism concerning market penetration rates exhibited by process models stems from the difficulty of modeling institutional obstables to new technologies. In the case of cogeneration, the contractual and financial complexity of many of these projects has hindered projects that might otherwise have been economically attractive. Furthermore, the aggregate nature of national models obscured regional variations that limited cogeneration activity to those utility districts that needed capacity. These issues are typically handled by a logistic market penetration function but there is little theoretical basis for estimating lag parameters other than those associated with capital vintaging.

As was stressed before, the promising direction is to take advantage of the chief features of both methods. The models should not rely entirely on historical trends. Even recent history cannot point out future trends. Future trends are driven by many other factors such as penetration of new technologies, input prices and availability, shifts in processes, and product mix, etc. Econometric models are also plagued by the fact that supply constraints have been a major trend factor affecting consumption in the past in many countries. Thus, demand relationships cannot be easily identified through normal econometric techniques.

VIII. CONCLUSIONS

No one type of model can be recommended as the best one for a particular forecasting task. The choice is with the analyst, who can make it based on the experience and availability of techniques and resources. Often a system of models approach will be preferable over that based on a single model.

REFERENCES

Boyd, G. and T. Kuczmowski. (1989). "Impact of Structural Shift in the Economy on Energy Demand and Conservation" EMF Working Paper 8.14, J. Weyant (ed), Industrial Energy Demand, North Holland.

Battelle-Columbus Division. (1984-1987). Demand-Side Management, Vols I-V, EPRI EA/EM 2597, Palo Alto, CA.

Faruqui, A. and J. Broehl. (1987) (editors). The Changing Structure of American Industry and Energy Use Patterns, Columbus, OH: Battelle Press.

Lilienthal, P. (1986). "Overview of Model Structures." EMF Working Paper 8.13.

Resource Dynamics Corporation (1986). Electrotechnology Reference Guide, EPRI EM-4527, April.

Battelle-Columbus Division (1985). Industrial End-Use Planning Methodology (INDEPTH): Recommended Design, EPRI EA-4019, Palo Alto, CA: EPRI, Section 3.

Weyant, J.P. (1988) (editor). Industrial Energy Demand, North Holland.

SECTION 3

MODELING ENERGY–ECONOMY INTERACTIONS

Energy Vol. 15, No. 3/4, pp. 297–300, 1990
Printed in Great Britain

0360-5442/90 $3.00 + 0.00
Pergamon Press plc

3.0. MODELING INTERACTIONS+

WILLIAM W. HOGAN

J.F. Kennedy School of Government, Harvard University
Cambridge, Massachusetts
U.S.A.

and

SHAILENDRA C. PARIKH

Energy Division, Oak Ridge National Laboratory,
Oak Ridge, Tennessee
U.S.A.

1. INTRODUCTION

The models represented in this chapter address many issues. Further, the conceptual orientations of the models differ and different components of the problem are emphasized. This diversity is valuable for our present discussion and is essential for the extension of the models to wider studies. When concentrating on the energy-economic feedback, however, a common structure of the models emerges. At cost of ignoring some important individual model details, this common framework provides a background for designing and interpreting the comparative model tests. The purpose here is to develop this common framework and sketch the participating models in this context.

2. GENERAL STRUCTURE

The general framework derives from the familiar economic view of the circular flow of products and resources (Figure 1). The many choices in this economy are divided into the decisions made by two representative groups, the producers and the consumers.

The producers utilize the resources of the economy (e.g., capital, labor, and energy) to make products (e.g., gasoline and food). These products can be classified as energy and nonenergy products, each of which is provided directly for consumers or utilized in other production processes. The consumers in turn provide the economic resources to the producers and demand their products. The demands of consumers may be for energy or nonenergy products and there is an interaction between these components.

These dual relationships, the supply and the demand for resources and products, must be appropriately balanced. Each model has some implicit or explicit mechanisms for achieving this balance, and some of the most fundamental model differences can be found in the alternate descriptions of these interfaces. Once obtained, the balanced flow of resources and products is the primary measure of economic activity that can be compared to the level of energy utilization.

+The entire materials for this section was taken from EMF 1 Report, Energy and the Economy, Appendix 1, "Comparison of Models of Energy and the Economy", 1977

For the present discussion, certain features of the producers, consumers, and balancing mechanisms become prominent. These are the modeling characteristics which seem most relevant in explaining the different approaches to the modeling of the energy-economic feedback.

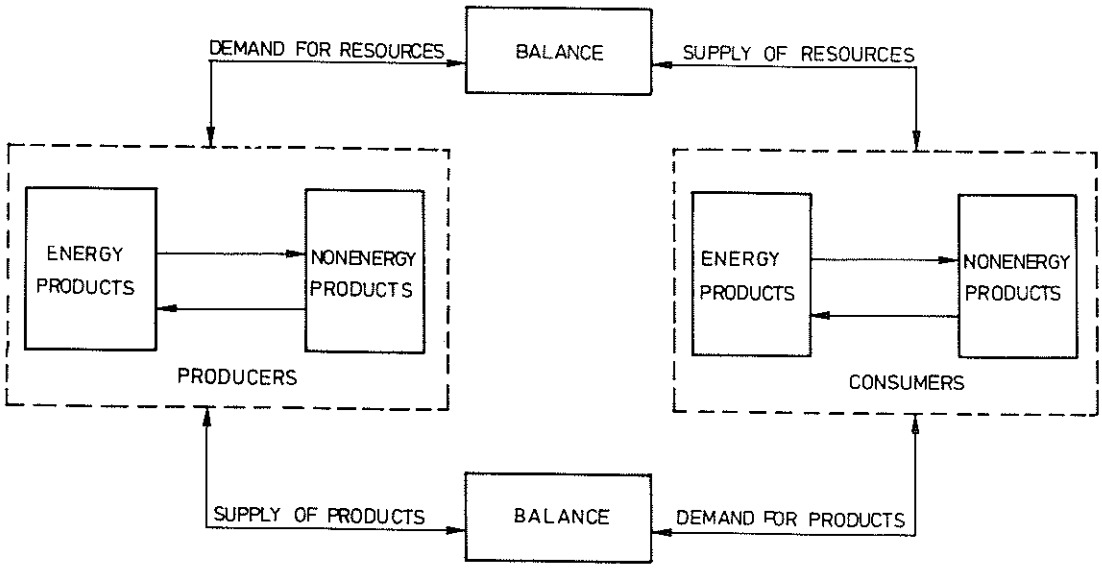


Figure 1. Flow of Products and Resources

Producers

The models uniformly organize the production sector accounting through the use of an interindustry input-output structure. A major source of the demand for energy is the link in the production process where energy is an intermediate good used in the manufacture of other products. But the models differ in several features in the design of the input-output structure and the treatment of energy. These features are the level of aggregation, the degree of substitution, the representation od dynamics, and the incorporation of trends.

Aggregation. All the models distinguish between energy and nonenergy products, but the level of further disaggregation varies from none to a separate representation and accounting for more than 50 industrial groupings. Designed for different purposes, these varying levels of aggregation may provide some insight regarding the aggregate effects of possible changes in the composition of industrial output.

Substitution. A key element in the measurement of the feedback from the energy sector to the economy is the assessment of the flexibility of energy utilization in the production system of the economy. This flexibility can be broadly classified into two components: interfuel substitution and factor substitution. Inclusion of the interfuel substitution permits examination of those interactions where a conversion from one form of energy to another is important. On the other hand, the inclusion of factor substitution permits construction of scenarios depicting switches from energy to other factors of production (labor, capital, and intermediate goods and services).

The manner in which any of these two broad categories of substitutions is incorporated in the model also provides yet another source of model differences. Some substitution characteristics are implicit in the choice of the level of aggregation. Across the components of an aggregate,

these implicit substitutions may assume either perfect substitutability or perfect complementarity. For example, in a model containing the aggregate "oil and gas", it is possible to assume either that "oil and gas" can meet known fixed proportions of demands for oil and gas (perfect complementarity). The actual result may vary with different uses of the same model and may depend on how aggregate quantities are treated when exogenous changes are made. In complex model where there are different levels of aggregation in different parts of the model, it is important to define the rules of disaggregation. This is seldom done.

The area where great care usually is taken is in deserving the explicit substitutions across fuels and factors in the input-output structure. There are three prominent approaches: use of fixed coefficients (assumption of perfect complementarity), use of behavioral relations based on econometric analysis of historical data (econometric representation), and use of engineering descriptions of alternative process technologies (engineering representation). The possibilities for flexibility in the input-output coefficients are an important component of any flexibility in the energy-economy feedback and the score of some of the major differences in the model conceptions and implementations.

Dynamics. All the system deal with evolving economy in a dynamic framework, but the underlying dynamic structures differ amount the models. Two attributes may be used to describe the differences across models: interaction of variables and speed of adjustment. The interaction of variables may be myopic, with current decisions determined entirely by current parameter values, or clairvoyant, with current decisions determined simultaneously by all parameter values. The dynamics of a model might permit instantaneous adjustment in variables resulting from exogenous shocks, or the adjustment may be gradual. In all the models, the key variable that forces some gradual adjustment is the stock of capital goods carried from one period to the next.

Trends. All of the models deal with some important exogenous parameters and structural changes through the application of simple trends. In this context, it is possible to define a standard set of variables that determines the long run growth of output in absence of any bottlenecks. This standard set consists of time profiles of population, labor force, and labor productivity.

Consumers

The models are less uniform in the accounting structure for final demand, but a similar set of characteristics provides a framework for describing the consuming sectors.

Aggregation. The level of product aggregation in each model consuming sector generally matches the corresponding level of aggregation in the production sector. However, these components of final demand are separated further into different types of consumption activities, investment groupings, classes of government expenditures, and types of net exports. Further, in some models the demands are estimated directly and in others they are derived from more fundamental end use requirements. As in the production sector, the level of detail hare may permit examination of the effects of changes in the composition of final demand.

Substitution. A set of substitution characteristics parallel to those defined for producers is found in the structure of final demand. Thus, it is possible to broadly classify the total substitution into two components: interfuel substitution and factor substitution. Also, these substitutions are implicit in the level of aggregation of final demand as well as explicit and modeled through assumption of perfect complementarity, an econometric representation, or an engineering representation.

Dynamics. Classification of the dynamic behavior in the models parallels the classification for the producing sector. Thus, one looks for either myopic or clairvoyant interaction of variables, and instantaneous or gradual adjustment in the profiles of demands for goods and services brought about by changes in their relative scarcities of prices.

Trends. Whether due to government regulation, evolving tastes, or other unspecified factors, all the models recognize the existence of important demand changes that are essentially exogenous to the model structure. For example, the basic MRG assumptions include an eventual 10% reduction in energy demand through non-price-induced conservation, possibly as the result of standards being imposed by the government. Thus, it is possible to define a standard set of exogenous trends in specification of time profiles of exogenous government expenditures and exports, and non-price-induced conservation of energy demand.

Balancing Mechanism

The structure of the model's view of the balance between supply and demand is a source of major difference in the systems. All of the models preserve a physical balance in terms of the real flows of products and resources and, for models with explicit markets, a balance of monetary flows is maintained. However, the behavioral balance of the systems is a source of model diversity, of which the most important features can be described in terms of the model objective and treatment of dynamics.

Objective. The models are either positive or normative.

- Positive models postulate the existence of certain behavior on the part of consumers (e.g., maximizing utility) and the producers (e.g., maximizing profits) and assume that these sectors communicate through competitive markets. The primary focus of information exchange in the markets is the distribution of relative prices of the products and resources. For given prices, the behavioral assumptions plus a balancing of incomes yield the corresponding supplies and demands for resources and products. For the positive models, balance is achieved when these quantities and prices are equal. The system is then said to be in a market equilibrium.

- Normative models here start from the same technological description of production and consumption possibilities, but assume that the producing and consuming sectors operate cooperatively to maximize some joint criterion function. The system is in balance when the physical flows match and the production-consumption activities are at levels which maximize the criterion function over the feasible values.

Dynamics. The dynamics of the balancing mechanism cannot be viewed independently. Rather, its nature depends upon the treatment of the dynamics of production and consumption in a particular model. If the implementation of the dynamics of production and consumption is myopic, then the determination of the market equilibrium also is myopic. The positive models have followed this approach. However, such myopic characterization is not essential. It is possible to define a market equilibrium in the presence of perfect foresight, and thus have a clairvoyant implementation of the market equilibrium.

Energy Vol. 15, No. 3/4, pp. 301–324, 1990
Printed in Great Britain

0360-5442/90 $3.00 + 0.00
Pergamon Press plc

3.1. SHORT AND MEDIUM-TERM MODELING AND PROBLEMS OF MODELS LINKAGE

P. CAPROS[+], P. KARADELOGLOU, G. MENTZAS, and J.-E. SAMOUILIDIS

Department of Electrical Engineering, National Technical University of Athens
42, 28is Octovriou Street
106 82 Athens
GREECE

Abstract - In this paper we present the main theoretical and empirical issues encountered in the construction of a short/medium-term energy-economy linked system of models, namely the Hermes-Midas system. The paper focuses on the problems of model linkage and proposes a methodology. The Hermes-Midas Linked System covers all aspects of energy-economy simulation and policy analysis for European countries.

1. INTRODUCTION[++]

The role of energy in the economy in the short-to-medium-term is an important subject of energy economics, with multiple policy implications. External energy price shocks, domestic energy pricing schemes, patterns of energy taxation, and problems of financing and scheduling large energy investments are some of the crucial policy analysis subjects over the medium and short term. These subjects should not be disjointed from classic macroeconomic regulation themes, such as economic growth, inflation, employment, deficits and monetary policy.

In economics, there is generally a dichotomy between short/medium-term and long-term representations of economic adjustment mechanisms. The distinction is based on the philosophy underlying the representation of the economy, rather than strictly on the time frame. For the long-term, it is assumed that there is enough time for economic agents to adjust their behavior to optimum levels and for markets to clear. Thus, general equilibrium and structural changes are inherent features of the long-term representation of economy. In the short/medium-term, although the economy tends towards general equilibrium, the analysis needs to include a representation of adjustment processes, pressures and temporary disequilibria. Another common assumption is that important structural changes can be ignored in a short-term representation.

After the 1973 crisis, numerous analysts became conscious of the fact that energy has important short/medium-term consequences on the economy and vice versa. However, the design of structural changes in a nation's energy policy requires a consistent long-term perspective. This explains the priority given to modeling long-term energy economy interactions.

Policy questions concerning structural changes in energy supply patterns, changes in consumer life styles and energy consumption processes, as well as technology shifts, have been studied in

---

[+]To whom all correspondence should be addressed.

[++]We are indebted to Mr. Pierre Valette, Mr. Eric Vanhalewyn and Dr E. Romberg of the DG/XII of the Commission of the European Communities for their support.

depth in most countries. The modeling tools used are all based on a neoclassical representation of economic mechanisms: each economic agent maximizes his profit or utility while commodity and factor prices are determined from the general (Walrasian) economic equilibrium under competitive markets (see e.g. G. Ackley (1961)). This representation is either explicitly or implicitly retained in long-term energy/economy models independently of their form, i.e. optimization or simulation; for the uniformity of long range energy economy models see Weyant (1985).

This approach fails to represent adequately the short/medium-term adjustment of the economy. Within that time frame the existence of markets in disequilibrium, adjustment processes and expectations need to be explicitly represented.

As in most short/medium-term macroeconomic models, we retain the Keynesian representation of the economy (see L.R. Klein (1947)) for integrating energy. Temporary disequilibria are permitted, mainly in the labor, goods and money markets. Thus, unemployment, under-utilization of production capacities and deficits can be simulated; see Figure 1 for a schematic comparison of the Keynesian and the neoclassical approaches.

Since a short/medium-term energy economy model evaluates the immediate economic consequences of energy pricing, taxation and investment policies, it should be considered as being a complement (and not a rival) of long-term equilibrium models. In fact, it provides insight into the characteristics and the feasibility of paths driving to long-term restructuring objectives.

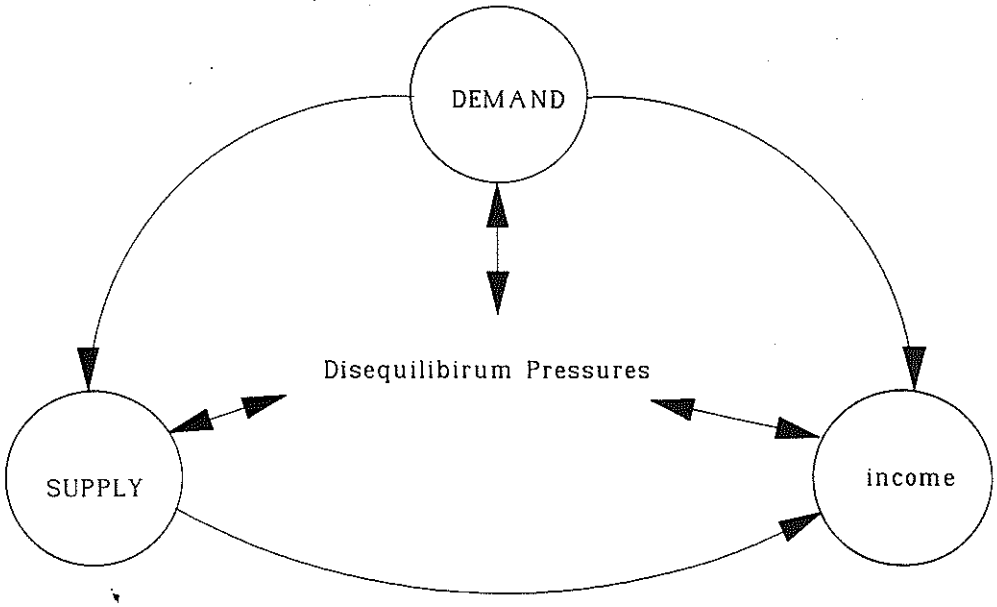
In this paper we present the main theoretical and empirical aspects of the construction of a short/medium-term energy/economy system of linked models based on neoKeynesian theory. In Section 2 of the paper, we present the main methodological issues related to the integration of energy in a macroeconomic model. Section 3 analyses the theoretical and empirical issues inherent in model linkage. A large-scale modeling application is described in Section 4 based on the Hermes-Midas modeling system, which is available for the European countries. Conclusions and perspectives for applications, as well as topics for further research are provided in Section 5.

2. METHODOLOGY

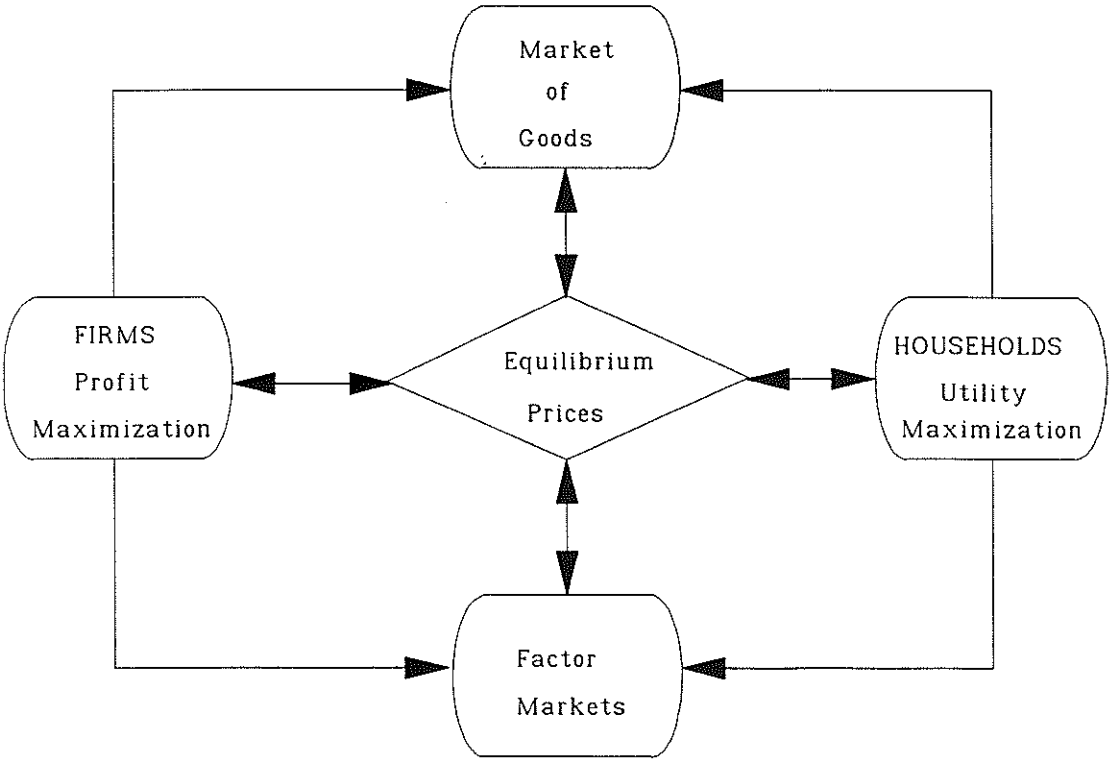
The Keynesian Approach to Macroeconomic Modeling

The basic elements of Keynesian economic modeling include the rejection of Walrasian competitive equilibrium and reflect the view expressed by Keynes in the General Theory that equilibrium is obtained via quantity rather than price adjustments. The Keynesian economy is demand driven: demand (i.e. consumption, investment, public expenditures etc.) is determined in the first place. If we denote this demand by Q_d, the behavior of producers is constrained by $Q_s = Q_d$, where Q_s denotes quantities supplied. The demand for labor (L_d) is determined afterwards, from a relation evaluating the quantity of labor required for producing Q_s (e.g. $L_d = f_L(Q_s,K)$ where K stands for capital). Labor supply (L_s) is determined by households and equilibrium employment is evaluated by $L = \min(L_s, L_d)$. Thus unemployment is equal to $L_s - L_d$ if $L_s \geq L_d$. The labor market is not cleared by prices as it would be in neoclassical models. In addition, since a demand-driven economy is the case, under-utilization of production capacities may occur. In a similar way, the interest rate is restricted to the money market disequilibrium and it does not clear the financial market (investment-savings) as would happen in neoclassical models.

This basic Keynesian framework has been considerably extended in recent models giving rise to the so-called neoKeynesian approach to empirical economic modelling.



The (neo)Keynesian Scheme



The Neoclassical Scheme

Figure 1. The Keynesian and the Neoclassical Schemes

Extensions include mainly the integration of price, wage and profit equations, the use of production functions, and the use of foreign trade relations. Production functions evaluate potential output, which is used in the determination of the rate of capacity utilization, which in turn is used in the price determination equations. Thus, prices and demand are affected by supply side factors, interactions that are mainly found in neoclassical models. For such reasons the term "neoKeynesian" should be interpreted as neoclassical-Keynesian rather than "new" Keynesian. Anticipation formulations and adjustment processes (either backward or forward looking) are also introduced in most models in order to improve their dynamic properties.

The pioneering work of applying these ideas to the empirical modeling of macroeconomic mechanisms is due to Klein and Goldberger (1955). Since then it has been extended and applied in many circumstances and many countries. Large-scale neoKeynesian macroeconomic models known at an international level include the following: Wharton, see Preston (1972); Brookings, see Fromm and Klein (1975); Chase, see Evans (1970); DRI, see Eckstein et al (1974); DMS, see Fouquet et al (1978); METRIC, see Artus, Bournay et al. (1979), etc.

A Conceptual Aggregate NeoKeynesian Model

In order to explain how energy can be integrated in a neoKeynesian macroeconomic model, we will first present the equations of a conceptual aggregate model without any explicit representation of energy. The aggregate model reproduces the structure of a typical neoKeynesian model, as well as its essential features in functioning.

Consider an economy with a single composite good and three economic agents, namely a representative household, a representative firm and government. The composite good is produced by means of two production factors, namely labor and capital. The monetary sector is represented in reduced form by the interest rate which is assumed to be arbitrarily predetermined, an assumption that may be plausible in short term analysis for a small economy. Following the neoKeynesian methodology we assume excess supply in all markets (goods, labor, current account) and thus quantity adjusts supply to demand. We assume also that prices are determined by mark-ups on supply costs which reflect imperfect competition.

Demand is the sum of private consumption (C), public expenditures (G, considered exogenous), investment (I) and net exports (i.e. exports, X, minus imports, M). National production (Q), i.e. the supply of goods, is adjusted to demand through the following relation:

$$Q = C + G + I + X - M \tag{1}$$

In a neoKeynesian model, it is usually admitted that there is a feedback of prices and real interest rate on private consumption, in addition to the traditional effect of income, via the marginal propensity to consume. Thus, we have[3][4][5]

$$C = \Phi_c(wL/p, \hat{p}, r-\hat{p}) \tag{2}$$

where w is the wage rate, L is employment, p is the price index and r the interest rate. It is clear that wL/p is a measure of real income.

---

[3] $\hat{x}$ denotes the variation of x.

[4] Φ is a functional form which may include distributed lags.

[5] x_{-1} denotes x lagged by one period.

Imports depend on the level of demand and on competitiveness, measured by the ratio of national to foreign prices in foreign currency. In neoKeynesian models the effect of pressures on production capacities is often included. Thus:

$$M = \Phi_M(Q, p/(p^*e), 1-U_c) \tag{3}$$

where p^* denotes foreign prices, e the exchange rate and U_c the rate of capacities utilization.

$$X = \Phi_X(Q^*, p/(p^*e), 1-U_c) \tag{4}$$

where Q^* is an exogenous index of foreign demand.

The producer is constrained to produce Q by means of capital (K) and labor (L). His behavior can be viewed as a problem of minimization of cost cK + wL, subject to a constraint representing technically feasible production mixes, whose efficient frontier is the production function Q = f(K,L). Since Q is restricted, we will include only the derived factor demand equations in the model.

For short/medium-term models it is preferable to assume also that producer's decisions concern only variations (and not the levels) of the production factors. One of the approaches used is the one based on a putty-clay technology, according to which only ex ante substitutability between the production factors is allowed, while the ex post capital-labor ratio remains fixed. Hence the producer decides only about the proportions of capital and labor in the latest vintage of production inputs, i.e. about investment (I) and the change in labor input, while stocks remain unmalleable. Under cost minimization this assumption yields[6]:

$$I/\Delta Q_c = \Phi_I(t, w/c) \tag{5}$$

$$\Delta L/\Delta Q_c = \Phi_L(t, w/c) \tag{6}$$

$$K = (1-\delta)K_{-1} + I \tag{7}$$

$$L = \Delta L + \sum_{i=1}^{T} \Delta L_{-i} \tag{8}$$

where c denotes the cost of capital, t is time representing technical progress and δ the replacement rate of capital. Q_c denotes potential output which may be evaluated by a production function:

$$Q_c = f(K, L, t) \tag{9}$$

Equations (5), (6) and (9) should be related by the use of common coefficients. It is possible now to evaluate the variables that transmit disequilibrium pressures in a neoKeynesian framework, namely the rate of capacity utilization (U_c) and the rate of unemployment (U_N):

$$U_N = (L_o - L)/L_o \tag{10}$$

---

[6]In equations (5) and (6) additional explanatory variables are often included representing short-term profitability, financing conditions etc.

$$U_c = (Q_c - Q)/Q_c \tag{11}$$

where the supply of labor (L_o) may be considered exogenous in the short-term. Disequilibrium pressures affect the determination of prices, wages and the foreign trade.

The price of the composite commodity (p) depends on the supply cost, which is decomposed in factor cost and cost of imported materials:

$$\hat{p} = \xi(\mu(L/Q)\hat{w} + (1-\mu)(K/Q)\hat{c}) + (1-\xi)((\hat{p}^* + \hat{e}) - \tau U_c \tag{12}$$

A Philips curve formulation is generally used for evaluating the evolution of the wage rate:

$$\hat{w} = \theta\hat{p} - \eta U_N \tag{13}$$

Among the various formulations for the cost of capital we retain a form proposed by Ando, Modigliani et al. (1974):

$$c = p(1 + \delta + r - \hat{w}) \tag{14}$$

The neoKeynesian model is, thus, a system of 14 simultaneous equations that are solved dynamically in a time-forward manner. The interactions are presented schematically in Figure 2. Notice that, unlike the traditional Keynesian scheme, endogenous demand depends also on the feedback effects coming from pressure variables and from prices. Prices and wages are simultaneously determined.

Introduction of Energy in the Aggregate Model

Energy is introduced by considering its triple role in the economy, namely as a production factor, as a consumable and tradeable commodity, and as a producing industry.

The representation retained is illustrated by an Input/Output scheme presented in Figure 3, which consists of two sectors/products, namely energy and non energy. This scheme is general enough to allow the representation of economies that produce and/or import energy.

Concerning private consumption, households should determine not only total consumption (and thus savings) but also the relative shares of energy and non energy products. Thus, there must be an expenditure allocation system in addition to equation (2):

$$C = \Phi_C(wl/p, \hat{p}, r - \hat{p}) \tag{1}'$$

$$C_N = \Phi_{CN}(C, p/p_E) \tag{2}'$$

$$C_E = \Phi_{CE}(C, p/p_E) \tag{3}'$$

where p_E is the price of energy and p the price of the non energy commodity. Functions Φ_{CN} and Φ_{CE} are related through the constraint $C = C_N + C_E$.

Imports and exports of non energy products are defined as in equations (3) and (4):

$$M_N = \Phi_{MN}(Q, p/(p^*e), 1 - U_c) \tag{4}'$$

$$X = \Phi_X(Q^*, p/(p^*e), 1 - U_c) \tag{5}'$$

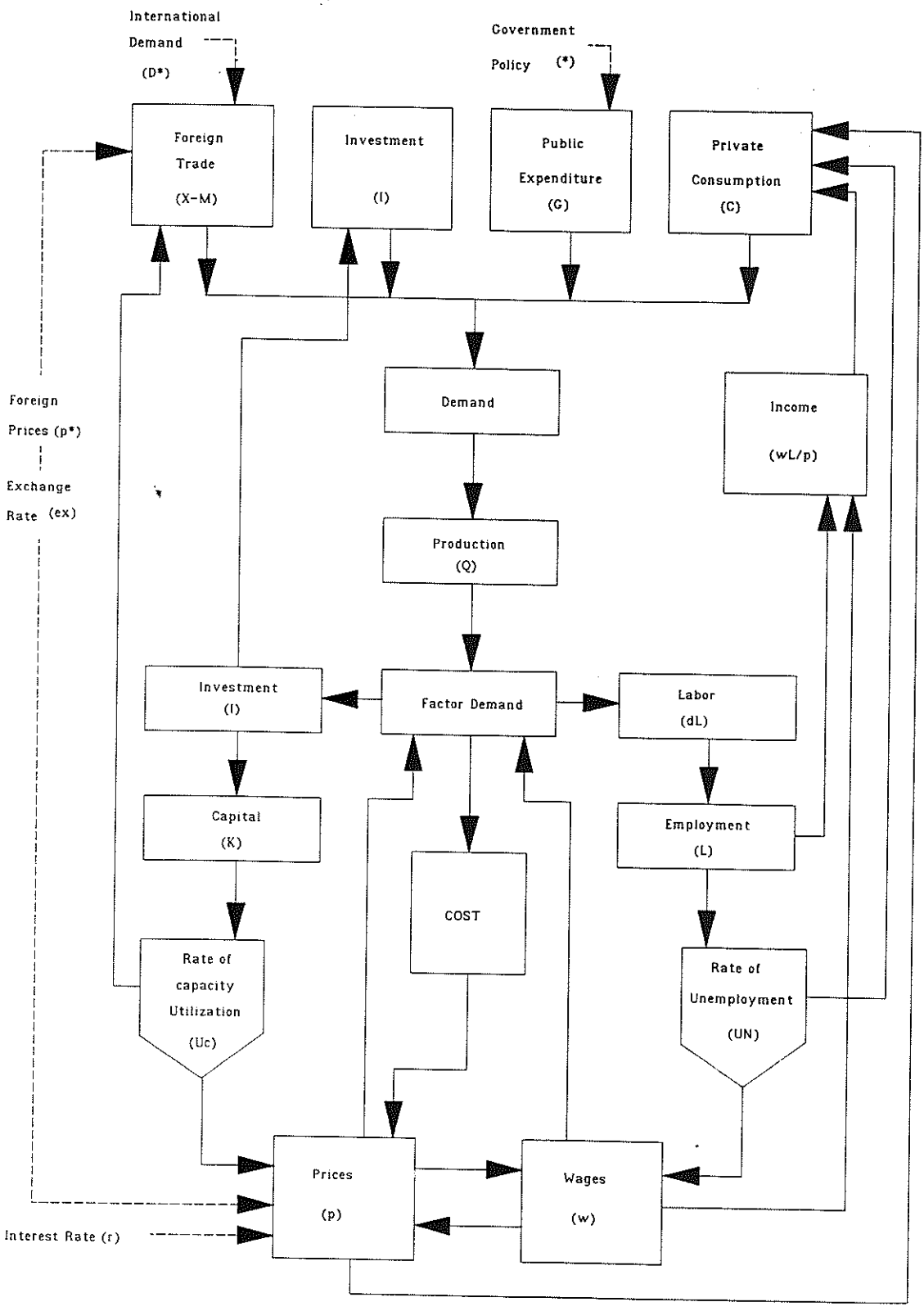
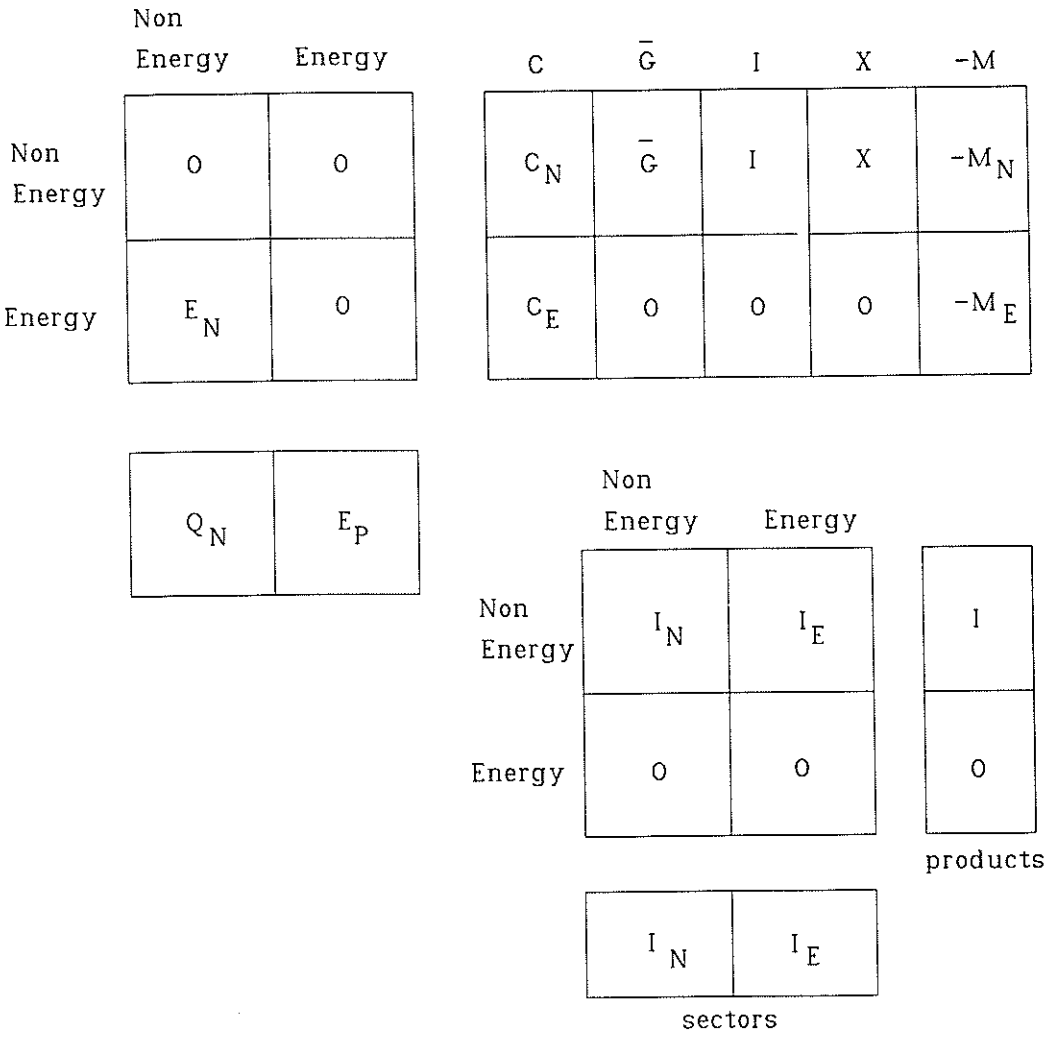


Figure 2. The Complete Neokeynesian Model



Q : production
C : consumption
M : imports
I : investment
E_P : energy production
index N : non energy
index E : energy

Identities

$Q_N = C_N + \bar{G}_N + I + X - M_N$

$E_P = E_N + C_E - M_E$

$I_N + I_E = I$

$C_N + C_E = C$

$M_N + M_E = M$

Figure 3. Input/Output and Investment Matrix

The introduction of energy as a production factor in the production system is performed by extending the production function. The so called K.L.E. (K: capital, L: labor, E: energy) production function is an important contribution of energy economics and its pertinence has been proven. Since the pioneer work of Hudson and Jorgenson (1974) and Berndt and Wood (1977), KLE or KLEM (with M standing for input materials) production functions have been estimated for most countries. Most authors adopted a putty-putty technology which allows for immediate adjustment at the optimum factor demand levels. This could be interpreted only within the context of general economic equilibrium. In the short/medium-term, it is preferable that substitutions between production factors are allowed only for the last vintage of capital. This leads to the adoption of a putty-clay technology. This was first introduced (with energy) by Artus and Peyroux (1981). Factor demand equations under cost minimization and putty-clay technology can be written as follows:

$$I_N/\Delta Q_c = \Phi_{IN}(t, w/c, p_E/c) \tag{6}'$$

$$\Delta L_N/\Delta Q_c = \Phi_{LN}(t, w/c, p_E/c) \tag{7}'$$

$$\Delta E_N/\Delta Q_c = \Phi_{EN}(t, w/c, p_E/c) \tag{8}'$$

$$K_N = (1 - \delta_N)(K_N)_{-1} + I_N \tag{9}'$$

$$L_N = \Delta L_N + \sum_{i=1}^{T} (\Delta L_N)_{-i} \tag{10}'$$

$$E_N = \Delta E_N + \sum_{i=1}^{T} (\Delta E_N)_{-i} \tag{11}'$$

These equations concern the non energy sector, denoted by the subscript N.

The consideration of energy as a producing industry includes all production, conversion and distribution of energy. The energy sector is usually considered a capital dependant industry. It is also common to consider that energy investments (I_E) are determined within long range planning so that they are given for short/medium-term analyses. Thus, the energy sector's capital stock (K_E) is determined by:

$$K_E = (1 - \delta_E)(K_E)_{-1} + I_E \tag{12}'$$

where δ_E is the replacement rate of energy capital. Domestic energy production (E_p) depends, generally, on the stock of capital and the availability of natural resources (R_E) which of course is exogenous in the short/medium-term:

$$E_p = \Phi_E(K_E, R_E) \tag{13}'$$

By assuming imperfectly competitive energy markets (oligopolistic markets and/or energy price regulations), which is a quite plausible hypothesis in most European countries, we represent the domestic production price of energy as being determined by mark-ups on production, conversion and/or distribution total costs. World energy product prices, however, have an important role in price determination, especially for coal and oil products, in the sense that world prices are driving the domestic ones. This corresponds to the small open economy assumption, which again seems valid for most European countries. Energy supply activity uses capital and labor services the cost of which should affect the production cost. The general form of the equation which determines domestic production energy prices is as follows:

$$p_{EP} = \Phi_{EP}(p_{EM}e, w, c, t) \tag{14}'$$

The energy industry is also a source of employment and income which is represented by:

$$L_E = l_E E_p \quad (15)'$$

Equations (12)' to (15)' should not be used if energy supply activity is assumed to be exogenous, or if energy is considered, for simplicity, as being only imported. Such assumptions are often made in conceptual or reduced form energy-economy models.

Net imports of energy are evaluated from Input/Output identities and provide a quantity adjusted demand/supply equilibrium of energy:

$$M_E = E + C_E - E_p \quad (16)'$$

The consumer price of energy, which influences the non energy production block, i.e. (6)' to (8)', and the expenditure allocation system, i.e. (2)' and (3)', can now be evaluated as a weighted sum of domestic and imported energy prices:

$$p_E = (1 + t_{vat})(t_E + (p_{EP} E_p + p_{EM} eM_E)/(E_p + M_E)) \quad (17)'$$

where t_{vat} and t_E are tax rates on energy. The rest of the model is similar to that of the previous section:

$$Q_c = f(K_N, L_N, E_N, t) \quad (18)'$$

$$Q_N = C_N + G + I + X - M_N \quad (19)'$$

$$L = L_N + L_E \quad (20)'$$

$$U_N = (L_o - L)/L_o \quad (21)'$$

$$U_c = (Q_c - Q_N)/Q_c) \quad (22)'$$

$$\hat{p} = \xi\left(\mu\frac{(L_N)}{Q_N}\hat{w} + v\frac{(K_N)}{Q_N}\hat{c} + (1 - \mu - v)\frac{(E_N)}{Q_N}\hat{p}_E\right) + (1 - \xi)(\hat{p}^* + \hat{e}) - \tau U_c \quad (23)'$$

$$\hat{w} = \theta\hat{p} - \eta U_N \quad (24)'$$

$$c = p(1 + \delta + r - \hat{w}) \quad (25')$$

Notice that (19)' guarantees demand/supply equilibrium of non energy goods and that additional cost elements due to the energy input are added to the price equation (23)'. Notice also that the introduction of energy, in its tripartite nature, implies the incorporation of important new loops in the model.

Problems of Representing Energy in the Model

There are no a priori restrictions for the form of the energy equations of the model (i.e. equations (3)', (8)', (12)' to (17)'). Various empirical formulations may be followed: a single functional form may be used for each one; they may be disaggregated by energy products; or, even, they may be replaced by systems of equations.

In energy modeling works a general distinction can be made between two types of formulation:

ex-post in which results are directly related to exogenous factors without any insight into the process of energy production, consumption or pricing; and

ex-ante in which this process is represented with some engineering, economic or decision details.

The empirical formulation of the ex-post approach relates to the application of econometric techniques to functional forms that link the output of the model to policy and/or exogenously defined variables, while the ex-ante approach often refers to several equations or relations with a higher level of techno-economic insight.

In energy modeling the ex-ante or mixed formulations are often preferred, since they are capable of representing energy technologies in a concrete manner. Research related to this approach includes: process analysis demand models; mathematical programming supply models; or even econometric models extended with process or engineering details. Such models are using energy balances as their accounting framework and consider energy in physical terms (e.g. kWh).

All modeling works for neoKeynesian energy-economy analysis begin with the construction of an econometric model similar to that of equations (1)' to (25)'in which ex-post direct econometric formulations are used for the energy equations. Let us term this type of energy-economy model a "compact" one, since it is consistent with a single statistical and definition scheme, the national accounting framework; see Askin (1978) for pioneer applications with the U.S. Wharton and DRI models, d'Alkantara and Italianer (1982) for the European HERMES model and Capros et al (1984) for the French MIEL model. Such models are integrated because they consider all aspects of energy and economy, but they are not satisfactory, with respect to energy.

Several attempts have been made to integrate ex-ante formulations of energy into a compact neoKeynesian energy-economy model. Two approaches should be distinguished:

- Single model in which energy equations are replaced by blocks of equations representing energy supply, energy demand and energy pricing, with some engineering and process detail; note that the mathematical form of the new blocks of equations is compatible with the whole model; see Brillet et al (1982) for the French Mini-DMS-Energie model which is the only example given in the literature.

- Linked systems of models in which the energy equations are eliminated from the compact model and the resulting one is linked with specific energy models which may be of different mathematical form; in addition there might be a global convergence procedure for solving the whole system. This approach has been extensively used mainly in neoclassical energy-economy models; example of neoclassical economic growth models linked to engineering energy ones include the LITM-TESOM-Illinois model, see Lucachinski et al (1979); the ETA-Macro model, see Manne (1978); the PILOT model, see Dantzig et al (1978); the IFFS-DAS-DGEM model, see Kydes (1983); the MEDEE-MESSAGE-IMPACT-MACRO model of IIASA, see Basile (1980); the German L.E.S.S. model, see Schmitz (1981); the Swedish MSG model, see Oettinger (1982); the Norwegian ENOR model, see Ek et al (1979); and the Japanese system of models, see Ogawa and Uchida (1979). Examples of linked model sets based on the neoKeynesian framework include the Wharton-Energy model, see Wharton EFA (1979); and the Belgian model system, see Osterrieth (1979).

Notice that there exists only one "single" neoKeynesian energy-economy model with a detailed engineering representation of energy. This is, in part, due to the significant difficulty in modeling the short/medium term fluctuations; a brief account of this difficulty is given below.

An important issue in neoKeynesian models is that they adequately simulate the dynamic evolution of alternative economic growth paths. Hence, they contain many static, as well as dynamic, feedback mechanisms; see also the aggregate model of the previous sections. The introduction of an ex-ante formulation for energy within such models generates important problems of inconsistency, which deal with:

- the system of definition of energy variables (energy balances instead of national accounts);
- the different disaggregation schemes for energy products; and
- the technical aspects of energy.

3. PROBLEMS OF MODEL LINKAGE

In order to understand the problems generated in the linked-model-set approach let us define some concepts related to model linkage. They deal with: (i) model inter-dependencies; (ii) convergence procedures; and (iii) dynamic resolution of the linkage.

If we assume that the two models represent two systems having some common parts, then there may be some overlap between them. Thus, common endogenous variables or closely defined endogenous variables should be eliminated from one of the models. To establish the linkage it is necessary to construct interface modules (i.e. sub-models) which perform the adequate transformations of variables, such as: the output of one model as input to the other; and the common or related exogenous variables.

When linking an energy with a macroeconomic model, the energy-related variables should be eliminated from the macroeconomic model and the macroeconomic evaluations should be deleted from the energy model. The interface modules guarantee consistency between the different definitions, representations and aggregation schemes employed in the two models.

A link variable is defined as being exogenous to one model and endogenous to the other. If link variables exist both from model 1 to model 2 and vice versa, the modeling system will be called a formal linkage. If this is not the case, one of the two models is satellite to the other. Satellite models are often used for disaggregating energy, once it has been evaluated at an aggregate level. The formal linkage needs a global convergence procedure for obtaining a solution. We must distinguish between two practices:

- empirical convergence: established by the user who runs the two models a few number of times; and
- numerical convergence: established automatically via a numerical algorithm (e.g. Gauss-Seidel) and by employing a severe convergence criterion.

Let x_t, $t = 1,2,...,T$ be a link variable evaluated by one of the models for all simulation periods (1,2,...T). Concerning the dynamic nature of the linkage, two cases should be distinguished:

- intertemporal linkage: the whole vector x_t, $t = 1,...,T$ is transmitted from model 1 to model 2 and vice-versa when a global iteration is performed; hence, global convergence is obtained once for all;
- dynamic linkage: the value of x_t for period t is transmitted from one model to the other during a global iteration; when global convergence is achieved for period t, the results feed period t+1 and the whole convergence procedure restarts for the next period.

When linking two models with different mathematical forms (e.g. a simulation with a linear programming model) it is quite impossible to obtain both numerical and dynamic convergence. This is the case of most neoclassical linked model systems, so the empirical and intertemporal convergence techniques are applied.

In neoclassical linked model sets the linking practices adopted are closely related to the theoretical framework of general economic equilibrium. Following the Arrow-Debreu theorem, each component model can be considered as expressing the optimizing behavior of an economic agent and the resolution algorithm implements market clearing mechanisms; that is, it finds a system of prices for which in all markets supply equals demand. This allows the consistent linkage of component models with different forms (such as linear programs, systems of nonlinear equations etc.). For a presentation of the underlying theory in general equilibrium model linkage see Hogan and Weyant (1982 and 1983).

The above approach, however, is meaningless in the case of linked neoKeynesian energy-economy model sets. First, the underlying theory is not that of general equilibrium. Second, the macroeconomic component has often the form of first-difference dynamic time-forward equations. Thus, one is more or less obliged to retain a convergence algorithm which simply finds the solution of systems of simultaneous equations. The energy model-components should follow similar mathematical forms, i.e. systems of dynamic equations. Nevertheless, there could be an exception to that if the energy model-component acts as satellite to the linked system (this was the case of an earlier version of the Wharton Energy model).

In brief, for short/medium term modeling both a numerical and a dynamic convergence are required. Notice that if we link two dynamic systems of equations (e.g. econometric models) and we apply a numerical and dynamic convergence procedure with severe convergence criteria, the resulting system is mathematically equivalent to a single system of equations, i.e. a single model.

In summarizing, we can conclude that within the context of a neoKeynesian energy-economy framework the introduction of ex-ante formulations for energy should be made by replacing the energy equations (3)', (8)', and (12)' to (17)' with systems of dynamic equations, which guarantee a sufficient representation of energy issues. Any definition or aggregation inconsistencies of the latter systems with the rest of the model should be suppressed by adding transformation equations. The resulting modeling system is either a "single" model or a "numerical-dynamic linkage"; the difference between the two is only a computational and not a mathematical distinction.

4. THE HERMES[7]-MIDAS[8] LINKAGE

Brief Description of the Models

The conceptual analysis presented in section 2 has been applied to the linkage of the macroeconomic model Hermes and the energy model Midas.

Hermes is a neoKeynesian multisectoral macroeconomic model with special emphasis on the energy sector. The structure of the model follows exactly the formulation of equations (1)' to (25)', but is extended to 9 sectors, including aggregate energy (which is further disaggregated

[7]HERMES stands for Harmonized European Research for a Multinational Economic and Energy System. It was initiated by d'Alcantara and Italianer (1982). The model was constructed by several national European teams under the coordination of DG/XII of the Commission of the European Communities. For a detailed presentation of the model see Italianer (1986) and Valette and Zagame (1989).

[8]MIDAS stands for Multinational Integrated Demand And Supply. The Midas-Demand model was initiated by ECOSIM Sarl (1986) and was constructed within the services of the Commission of the European Communities (DG/XII and DG/XVII). Details, as well as the TROLL implementation of the model are presented in Capros et al (1987). For a presentation of the Midas-Supply model see Capros, Karadeloglou and Mentzas (1988). The demand module of the model has been extensively revised by Detemmerman, Guillaume, Lecloux (1988).

in 8 energy products). Moreover, income distribution mechanisms are added in order to take into account taxes, transfers, subsidies and social funding. Aggregate energy demand in industry is evaluated as a derived factor demand, together with capital and labor. In private consumption, energy demand is included within 14 more general consumption categories and especially within "fuels for domestic use", "power for domestic use" and "fuels for personal transportation".

In the version of HERMES we worked with, a nested CES/Cobb-Douglas production function is used:

$$Q_t = Q_o\, e^{-gt}[\delta I_t^{-\sigma} + (1 - \delta)E_t^{-\sigma}]^{\frac{-a}{\sigma}} L_t^{1-a}$$

This function is formulated in terms of the putty-clay technology and the optimum levels of the marginal technical coefficients of the production factors are derived by cost minimization. This leads to the evaluation of $I^*/\Delta Q^c$, $\Delta E^*/\Delta Q^c$ and $\Delta L^*/\Delta Q^c$ as functions of the relative factor prices, where the asterisk denotes optimal (desired) factor demand.

Total optimal factor demand, for instance energy demand E^*, is evaluated by the use of a capital replacement process, and, finally, effective energy demand is obtained by assuming an error correction adjustment process to its optimal level. This is a method of transforming the traditional KLE production function for the short/medium term.

In some national Hermes models, energy demand is disaggregated by type of fuel following a translog formulation involving relative fuel prices as the only explanatory variables. Hence, interfuel substitutions are allowed. Separability assumptions guarantee that substitutions between energy products do not directly modify the aggregate energy demand; nevertheless, they affect the mean energy price, since consumer prices of aggregate energy are evaluated as a weighted sum of fuel prices. Thus, interfuel substitutions influence aggregate energy demand via prices. However, in many instances, this formulation did not prove accurate in the estimation or simulation phases and has been abandoned in several national Hermes models. In our case, we have not used the interfuel substitution option.

Midas is a medium-term annual energy model, combining econometric with process analysis formulations. It simulates the evolution of both energy demand (Midas-Demand) and energy supply and prices (Midas-Supply). Energy is measured in physical units following the energy balance system. A number of important concepts and capabilities found in Midas include the following:

- a complete module for the evaluation of energy prices;
- a fuel substitution mechanism that takes into account both prices and technical limitations;
- the use of load duration curves for the estimation of electricity supply and prices;
- the considerable emphasis on the role of the interest rate, the production and financing costs; and
- the evaluation of normal and abnormal scrapping of energy-using equipment.

Energy demand is evaluated separately for each sector. Useful demand (UE) for sector j is evaluated by the following equation which corresponds to the reduced form of a complex system of equations in Midas:

$$UE_j = G(Q_j)^a \Phi\left(\frac{PE_j}{P}\right)^{-b} \Phi\left(\frac{PE_j}{PK}\right)^{-c} \qquad (26)'$$

where G stands for technical progress, Q represents activity or income, P is a deflator of economic activity and PK is an approximation of the capital cost. The functions of relative prices include distributed lags.

A similar algebraic form is used for the evaluation of the desired (optimal) demand for energy products:

$$C\hat{E}_{ij} = G_{ij}(\bar{Q}_j)^a \Phi\left(\frac{P\hat{E}_{ij}}{P}\right)^{-b} \Phi\left(\frac{P\hat{E}_{ij}}{PK_{ij}}\right)^{-c} \qquad (27)'$$

The results of (27)' are considered the optimal shares of energy products. In a given time period these shares are applicable only to the non-allocated part of energy demand, that is to energy demand of new consumers and of those who replace their energy-related equipment. Two types of replacement processes are applied: normal scrapping, which corresponds to the life time of equipment; and abnormal scrapping, which corresponds to the replacement of equipment due to economic inefficiency. Thus, substitutions between energy products are applied only to the latest vintage of capital, which is analogous to a putty-clay dynamic process. Notice that this replacement and product substitution process does not affect total energy demand evaluated by (26)'; rather it concerns only its allocation in energy products. However, this allocation affects the mean energy price (PE) used in (26)'. This mechanism corresponds to the separability assumption in energy demand formation.

The MIDAS-Supply model includes the following modules: electricity production; refining sector; coal production; natural gas production; coke and synthetic gas production; and energy prices.

The electricity production module operates as follows : (i) a linearized load duration curve is constructed by using the energy demand results of MIDAS-Demand, (ii) electricity plants are allocated within the load curve following cost minimization, (iii) the marginal cost is approximated for each consumption category, and is introduced in the energy prices module.

The refining sector module uses an aggregate representation of a typical average refinery, including distillation, cracking and reforming. Desired production of each refining unit is compared to existing capacities in order to obtain estimations of the rate of capacity utilization and of production cost. Refinery throughput and the flows within the refinery are evaluated by means of engineering relations. Cost influences the competitiveness of national refineries in exports and the formation of product prices.

The production of fossil energy is formulated by means of supply curves (involving reserves), which also serve in the evaluation of cost elements.

Midas evaluates energy prices by using the dollar price of crude oil as the main exogenous variable. Most of the equations follow a rate of variation formulation, connecting an energy product price (PE_i) with: the price of a product (LP) which is considered as a leading price; and national and international labor (WG) and capital (PK) factors. Oligopolistic market structure is assumed and prices are defined by a mark-up process on costs (i.e. they are not competitive equilibrium prices):

$$\hat{PE}_i = A(\hat{LP})^a (\hat{WG})^b (\hat{PK})^{1-a-b} \qquad (28)'$$

Concerning electricity pricing the model adopts a special algorithm which corresponds to marginal cost tarification for each consumer type. Taxes are then added to these prices, in order to obtain consumer prices.

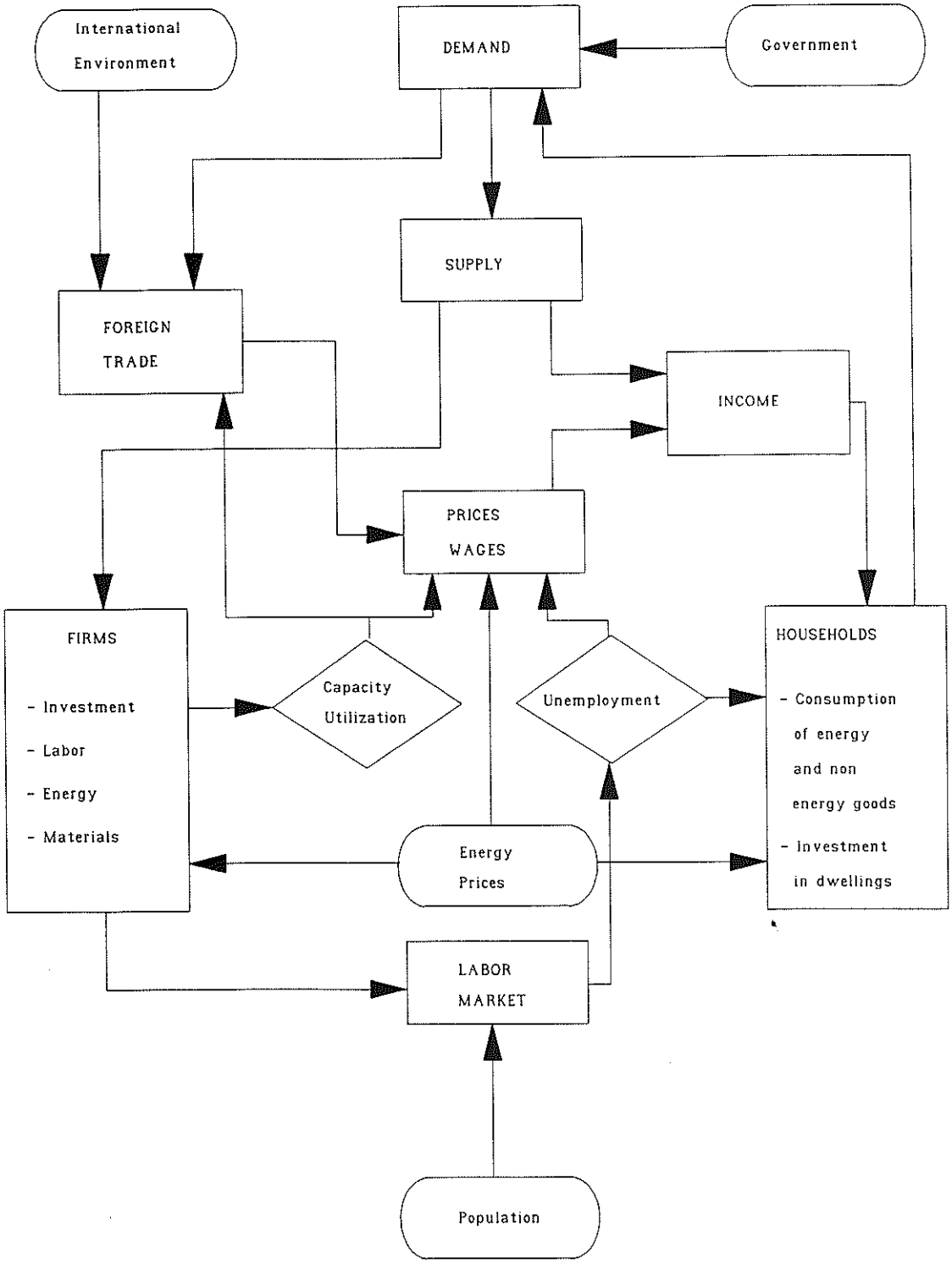


Figure 4. Simplified Representation of Hermes.

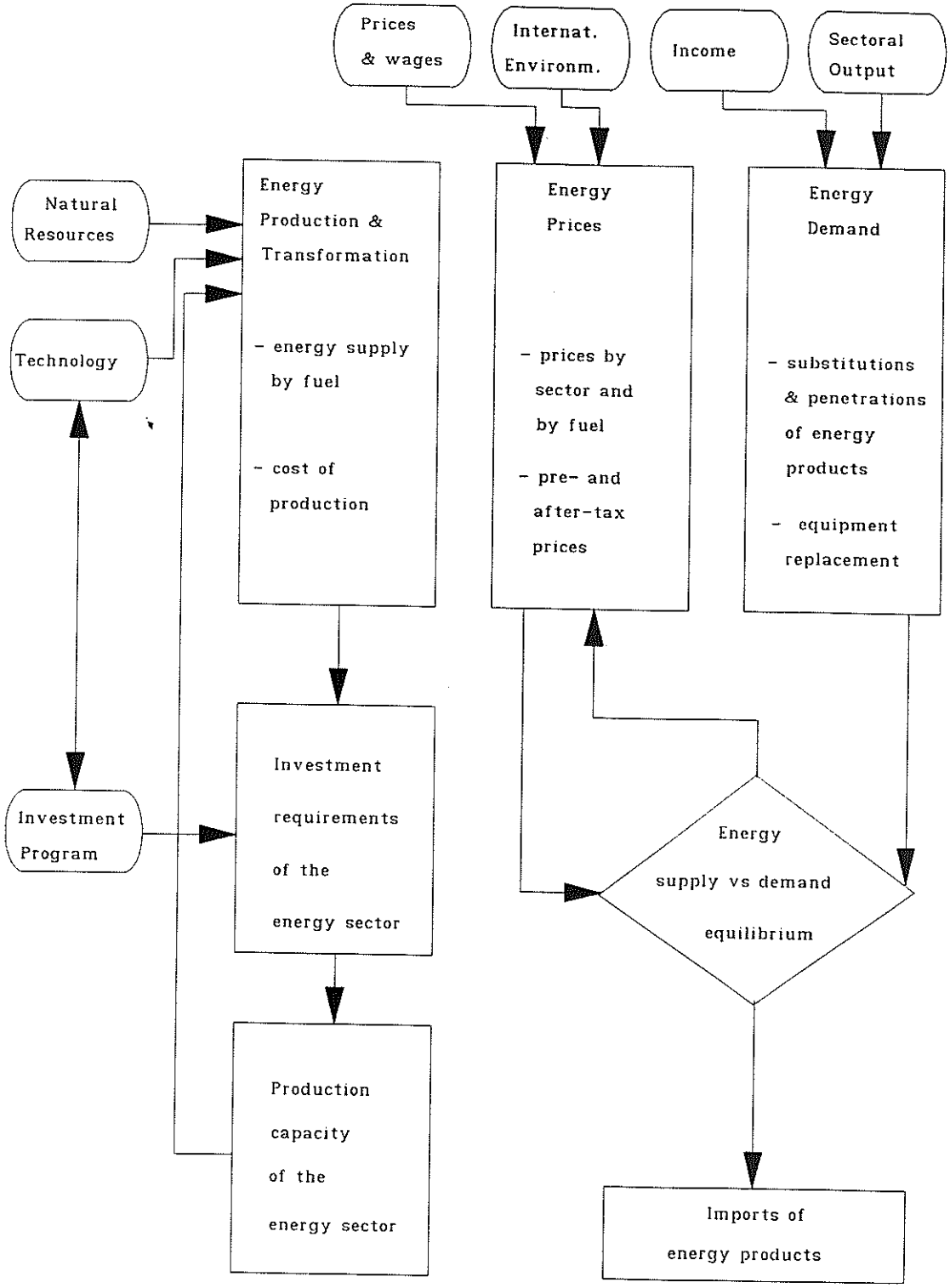


Figure 5. Simplified Representation of Midas

Consistency of energy supply, transformation and demand flows is guaranteed in the MIDAS model, so that complete balance sheets constitute its output.

Description of the Linkage

HERMES is a typical example of a compact energy-economy model, as defined in section 2.4. With respect to energy, however, it has several limitations. The model uses:

- ex post translog energy equations for evaluating industrial energy demand by energy product;
- ex post econometric demand equations for the evaluation of energy demand in the residential sector;
- time-series data for energy, measured with the national accounts principles; and
- it does not treat the tax policy, pricing, investment and technological issues which are tied to energy in a detailed manner.

The above constitute limitations, especially when the model is to be used by energy engineers and planners.

On the contrary, the Midas model is much more accurate concerning energy issues. In fact the model uses:

- mixed ex-ante and ex-post formulations for evaluating energy demand and the allocation in energy products;
- several engineering and policy details that are very useful in energy-related scenario construction;
- time-series data for energy, measured with the energy balances principles; and
- a sectoral classification scheme adapted to the energy analysis requirements.

For these reasons, it was decided to link Hermes with Midas, and, by considering their medium-term simulation nature, to design and implement a numerical/dynamic convergence procedure, as defined in section 3. The strategy is to modify both models by eliminating energy equations from HERMES and macroeconomic equations from MIDAS and then link them by means of interface modules performing transformations of link variables.

Concerning HERMES the following variables were made exogenous and the respective equations were eliminated:

- energy supply (production, imports, energy to energy intermediate consumptions, investment of the energy sector);
- industrial and residential energy demand equations of type (3)' and (8)';
- all energy prices equations of type (14)' and (16)'; and
- the government receipts from energy related taxes.

Modifications of MIDAS, in view of the linkage, include the elimination of a small block which evaluated macroeconomic activity, income and price variables by means of aggregate multipliers.

The two-way interface between the modified HERMES model and the modified MIDAS model is organized in two modules which perform transition of three types of variables: common exogenous (e.g. oil price, population, interest rate, exchange rate) which should be the same in both models; endogenous in HERMES and exogenous in MIDAS (e.g. gross domestic product, wage rate, rate of inflation, consumer's disposable income, etc.); and endogenous in MIDAS but exogenous in HERMES (i.e. energy demand, supply and prices).

The problems of inconsistencies between the energy and the economic models are due to both their different sectoral aggregation schemes and their different statistical systems. This required extensive efforts for the transformation of the endogenous variables of MIDAS into adequate input to HERMES and vice-versa. In order to overcome discrepancies it was assumed that the rate of growth of an energy variable in value units at constant prices is equal to the rate of growth of the same variable measured in physical units. Incompatibilities concerning the sectoral aggregation schemes were treated in the following manner: the rates of growth between comparable energy demand figures were assumed equal (e.g. transport demand for energy in Midas was assumed comparable to the sum of energy demand of the transportation sector and the demand for transportation fuels in the residential sector in Hermes); in a second phase, the exact shares of sectoral mix were evaluated in Hermes and used directly into the interface modules without having any effect on the rest of Hermes. A similar approach was followed for the transformation of prices, while tax receipts from energy were directly fed into Hermes.

The Main Issue of the Linkage

Whe we want to link a macroeconomic model based on a KLE or KLEM production function with an energy demand model, the following dilemma arises:

- should we conserve the aggregate energy demand equation originally used in the macroeconomic model and use the energy demand model only for disaggregating energy in various products? or

- should we eliminate the aggregate energy demand equation from the macroeconomic model and retain the energy demand model for all energy related evaluations?

The first approach is consistent with respect to microeconomic theory, but it does not take advantage of the accuracy of the energy model in evaluating energy demand. This approach implies that the energy demand model becomes a pure submodel, i.e. a satellite of the macroeconomic model. The second approach corresponds to a formal linkage as defined in section 3 and is clearly more attractive in energy-economy policy analysis since it overcomes the limitations mentioned above. The problem is, then, to justify the adoption of the second approach, or in other terms, to define the implications of replacing arbitrarily the energy demand equation of a macroeconomic KLEM-based model with an energy demand model which has been separately estimated. A similar problem arises also for the expenditure allocation system of consumers, for which we want to replace the energy demand equations with an energy demand model which has also been separately estimated.

In terms of the description of section 2.3, this problem concerns the replacement of equations (3)' and (8)' with separate energy demand models. When faced with this dilemma, we decided to eliminate the aggregate energy demand equations from the macroeconomic model (i.e. Hermes) if we could demonstrate that the qualitative simulation properties of Hermes would not be modified by this replacement.

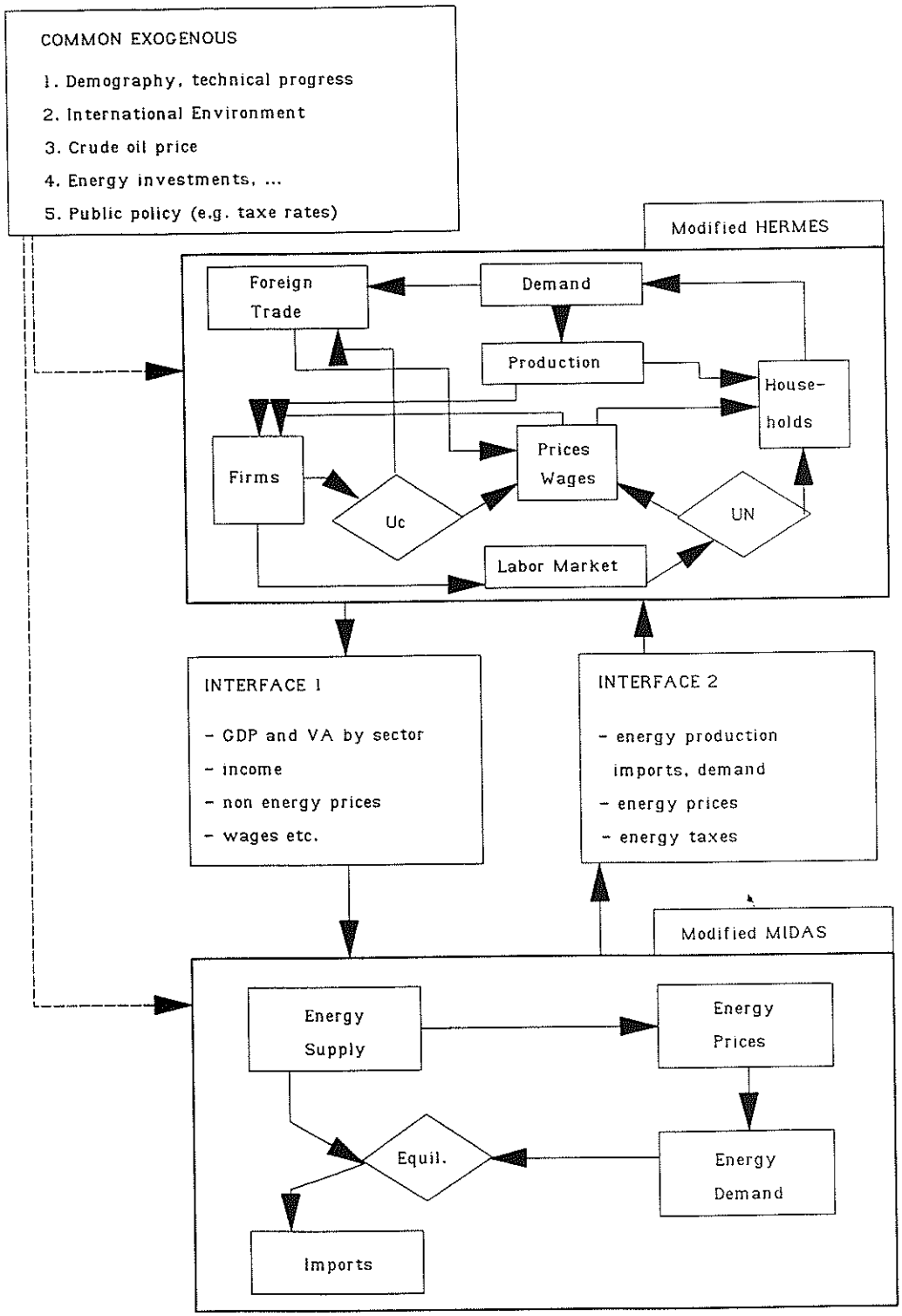


Figure 6. Simplified Representation of the Hermes-Midas Linkage

The problem has been studied both on a theoretical and an empirical basis; see Capros, Karadeloglou, Mentzas (1987 and 1989) for details. The research included the following:

- in a first step an aggregate core model of Hermes was constructed, i.e. model (1)' to (25)'; its steady-state solution was analytically obtained, as well as the effects on its properties when the energy demand equations were replaced. Moreover, the conditions for the preservation of the qualitative nature of Hermes' results were established;

- In a second step the same problem was studied empirically by injecting exogenous shocks to modified versions of Hermes and by comparing the results with those of the original model.

The main conclusion drawn from this study states that the alteration of Hermes behavior is not qualitatively significant if the Midas energy demand equations share some common properties with the equations eliminated from Hermes and if Midas demand elasticities are within a range from the ones of Hermes. The common properties include:

- the use of relative prices as main explanatory variables in the demand equations;

- the application of substitutions only to the latest vintage of capital; and

- the separability assumption concerning the allocation of total energy demand in energy products.

All three conditions hold for the demand equations found in Midas, as can be verified from the presentation in section 4.1.

A comparison of the simulation properties of the original Hermes model with those of the linked Hermes-Midas system has been performed under several exogenous shocks. This comparison justified our choices for the design of the linkage, the replacement of all energy equations of Hermes with the Midas model.

The conclusions drawn from the simulation exercises can be summarized as follows:

- in the case of exogenous shocks affecting energy-related variables the behavior of the linked system presents significantly different results compared to those of the original Hermes model; this is due to four types of reasons: (i) the price elasticities of energy demand are different, since they are separately estimated; (ii) demand-side substitutions between energy products are allowed only in the linked system; (iii) supply-side substitutions between energy forms and technologies, as well as within the energy import/export system are represented only in the linked system; (iv) national cost elements influence the consumption prices of energy only in the linked system;

- when the exogenous shocks affect capital within the capital-energy complex of the production function block, there is still a difference in the results but it is less important than in the previous case;

- if the shock is of a "pure" macroeconomic nature and does not affect directly the capital-energy complex, only slight differences are observed; this holds true even for variables affecting the third production factor, i.e. labor;

- however, if the shock affects variables entering the equations determining the national production cost of energy (e.g. wage rate, exchange rate, interest rate) the energy production, imports and exports are affected; this has economy-wide effects via the supply-side of energy; finally,

- all differences in the properties of the two systems (i.e. the original HERMES and the linked system) are only of quantitative nature, without significant qualitative discrepancies, especially for the short-term.

5. CONCLUSIONS

The Hermes-Midas linkage is a new modeling device for performing medium/term Keynesian energy-economy interactions analysis for European countries. It integrates into a common computational environment a complete representation of macroeconomic mechanisms and complete econometric/engineering energy models. Although it links different models, mathematically it should be considered a single model since the linkage has been constructed to be formal, numeric and dynamic. The whole system is composed of about 2500 simultaneous equations for each country and is running in the TROLL environment installed either on a personal computer or a mainframe.

Unlike the great majority of energy-economy modeling systems, the whole methodology on which the Hermes-Midas linkage is based, is adapted to the requirements of the short/medium term analysis. Macroeconomic mechanisms are represented following the neoKeynesian theoretical framework which allows for disequilibrium pressures in the labor and goods markets. It includes, also, feedback effects from prices-wages to domestic and foreign demand, as well as putty-clay technology factor demand equations. Concerning energy representation, details on pricing and taxation policies are incorporated into the energy price formation equations and techno-economic and process analysis aspects are specified into the system of energy demand formation and energy supply system simulation.

The Hermes-Midas linkage has important policy analysis and forecasting simulation capabilities for the medium term. It can be used for studying the effects of international events, namely oil price shocks, variation of exchange rates etc., on both energy and the economy. It can also evaluate the consequences of energy taxation and pricing policies, as well as energy investment schedules, and shifts in energy demand efficiencies and techniques. It can also perform all traditional macroeconomic policy analyses for the medium-term.

REFERENCES

Ackley, G. (1961)."Macroeconomic Theory", New York: MacMillan.

d'Alcantara, G. and A. Italianer (1982). "European project for a Multinational Macrosectoral Model", CEC, MS11, XII/759/82.

Ando, A., F. Modigliani, R. Rasche and S. Turnovsky (1974). "On the Role of Expectations of Price and Technological Change in an Investment Function", International Economic Review, Vol. 15, No 2, June.

Artus, P. et al (1979). "METRIC, Une modelisation de l'economie Francaise", Paris: INSEE.

Artus, P. and C. Peyroux (1981). "Fonctions de production avec facteur energie: estimations pour les grands pays de l' O.C.D.E.", Annales de l'INSEE, No 44, Oct-Dec.

Askin, A.B. (1978). "How Energy Affects the Economy", Lexington: Lexington Books.

Basile, S.P. (1980). "The IIASA Set of Energy Models: its Design and Application", IIASA, RR-80-31, Laxenburg, Austria.

Berndt, E. and D. Wood (1977). "Engineering and Econometric Approaches to Industrial Energy Conservation and Capital Formation: A Reconciliation", MIT, Working Paper, No MIT-EL 77-040 WP, November, published also in American Economic Review, Vol. 69, No 3, September, 1979.

Brillet, L.J. et al (1982). "Energie et economie: le modele mini-DMS energie", 23Economie et Statistique, No 146.

Capros, P. et al (1984). "Le modele macroeconomique MIEL", Revue Generale Nucleaire, No 2, Mars-April, p.166-182.

Capros, P., P. Karadeloglou and G. Mentzas (1987). "Linking Hermes with Midas", National Technical University of Athens, Report to the Commission of the European Communities, DG/XII, March.

Capros, P., P. Karadeloglou and G. Mentzas (1988). "The Energy Supply Model MIDAS", National Technical University of Athens, Report to the Commission of the European Communities, DG/XII, February.

Capros, P., P. Karadeloglou and G. Mentzas (1989). "Energy Policy Extensions of KLE-based Macroeconomic Models", Journal of Policy Modeling, (forthcoming)

Capros, P. and J.-E. Samouilidis (1988). "Energy Policy Analysis", Energy Policy, February.

Dantzig, G. et al (1978). "Stanford Pilot Energy/Economy Model", EPRI EA-626, May.

Detemmerman, M.V., Y. Guillaume and M. Lecloux (1988). "Midas Demande", Les modules Industrie et Transports, Report to the Commission of European Communities, DG/XII, April.

Eckstein, O., E.W. Green, A. Sinai (1974). "The Data Resources Model: Uses, Structure and Analysis of the U.S. Economy", International Economic Review, 15, 595-615.

ECOSIM SARL (1986). "The Midas Energy Model", DG/XII, Commission of the European Communities, March.

Ek et al (1979). "ENOR-An Energy Model for Norway", Institut for Atomenergi, Norway, International Conference on Energy Systems Analysis, Oct. 9-11, Dublin, Ireland, Reidel Publ. Co.

Evans, M.K. (1970). "Econometric Models" in W.E. Butter, R.A. Karesh, R.B. Platt (ed.), "Methods and Techniques of Business Forecasting", Englewood Cliffs, N.J., Prentice Hall, Inc.

Fouquet, D. et al (1978). "DMS, Modele Dynamique Multisectoriel", Collections de l'INSEE, No C64-65, Paris.

Fromm, G. and L.R. Klein (eds) (1975). "The Brookings Model: Perspective and Recent Developments", Amsterdam: North Holland.

Hogan, W.W. and J.P. Weyant (1982). "Combined Energy Models", in Advances in the Economics of Energy and Resources, Volume 4, pages 117-50, JAI Press Inc.

Hogan, W.W. and J.P. Weyant (1983). "Methods and Algorithms for Energy Model Composition: Optimization in a Network of Process Models", in B.Lev (ed.): "Energy Models and Studies", pages 3-43, North Holland Pub. Co.

Hudson, E. and D. Jorgenson (1974). "U.S. Energy Policy and Economic Growth, 1975-2000", The Bell Journal of Economics, Vol. 5, No 2.

Italianer, A. (1986). "The Hermes Model: Complete Specification and First Estimation Results", CEC, EUR 10669 EN.

Klein, L.R. (1947). "The Keynesian Revolution", New York: MacMillan.

Klein, L.R. and A.S. Goldberger (1955). "An Econometric Model of the United States, 1929-1952", Amsterdam: North Holland.

Kydes, A.S. et al (1983). "A New Energy-Economy System", in A.S. Kydes et al. (ed.) "Energy Modeling and Simulation", IMACS, North Holland Publishing Company.

Lucachinski, J. et al (1979). "An Integrated Methodology for Assessing Energy-Economy Interactions", International Conference on Energy Systems Analysis, Oct. 9-11, Dublin, Ireland, Reidel Pub. Co.

Manne, A.S. (1978). "Energy-Economy Interactions: An Overview of the ETA-Macro Model", presented at a Symposium on "Energy Modeling and New Energy Analysis", Chicago, pp. 341-351.

Oettinger (1982). "A MSG-model for Sweden", Draft, Energy Systems Research Group, French-Swedish Conference, June, Stockholm.

Ogawa, H. and M. Uchida (1979). "Modeling Japan's Energy-Economic Systems", in K.B. Haley (ed.): "O.R. '78", North Holland.

Osterrieth, M. (1979). "The Belgian Energy Models: A General Introduction", International Conference on Energy Systems Analysis, Oct. 9-11, Dublin, Ireland, Reidel Pub. Co.

Preston, R.S. (1972). "The Wharton Annual and Industry Forecasting Model", Wharton School, University of Pennsylvania.

Schmitz, K. et al (1981). "Simulation Techniques in Energy Analysis", in I.Kavrakoglou (ed.): "Mathematical Modeling of Energy Systems", Sijthoff & Noordhoff.

Valette, P. and P. Zagame (ed.) (1989). "HERMES Model", Commission of the European Communities, DG/XII, book in preparation, DG-XII, forthcoming.

Weyant, J.P. (1985). "General Economic Equilibrium as a Unifying Concept in Energy-Economic Modeling", Management Science, Vol. 31, No 5, May, p. 548-563.

WHARTON EFA (1979). "The Wharton Annual Energy Model: Development and Simulation Results", EPRI, EA-1115.

Energy Vol. 15, No. 3/4, pp. 325–339, 1990
Printed in Great Britain

0360-5442/90 $3.00 + 0.00
Pergamon Press plc

3.2. MULTISECTORAL PLANNING MODELS

XIA SHI

Department of Engineering-Economic Systems
Stanford University, Stanford, California
U.S.A.

YINGZHONG LU*

Institute for Techno-Economics and Energy System Analysis
Tsinghua University, Beijing
CHINA

Abstract - The evolution of input-output techniques and the associated linear and non-linear programming models are first introduced. The basic structures and mechanisms of multisectoral input-output planning models are then discussed in details, including the objective function and various types of constraints. Moreover, The advantage of using the planning models relative to computable general equilibrium models is discussed. Since the solution to planning models can also represent a competitive eqilibrium under certain specifications, the planning model technology is flexible to be applied to both market and centrally planned economy. Two examples of the applications of such multisectoral planning models are given, i.e., the Stanford PILOT macroeconomic model and a multisectoral model of China, the latter being believed to be in great need in the current Chinese economic reform process.

1. INTRODUCTION

Input-output techniques have been used for planning and economic analysis since Leontief originated his input-output pioneering work during the 1940s. The core of input-output techniques is that they provide a consistent framework to study intersectoral transaction relationship of productions and consumptions. The earliest economic planning models using a multisectoral input-output framework were the consistency models based on the assumption that the growth of production must be balanced with domestic demands to some extent. Static and dynamic models were both used in such planning processes, depending on the nature of the problems under study. These models were mainly used to verify the proposed plans of economic growth so that the final demands could be consistent with the quantity available of some primary resources; they were also used to make the desired sectoral allocation of investment goods. Since international trade has played an increasingly important role in the process of economic development for a country, many of these models have elaborately incorporated foreign capital inflows and comparative analysis of international trade. However, these models contain two major shortcomings: the lack of price induced substitution and the absence of economic criteria, thus leading to the further development of the linear programing activity analysis models.

* Visiting Professor at the Economics Department of Stanford University.

The linear programming models provide the linkages among social objectives, structural limitations, and policy instruments for the countries studied. The solution to such models gives the optimal allocation of productive resources among sectors. Many substitution schemes for different production technologies can be incorporated into these models. For foreign trade sectors, The linear programming models offer the optimal choice between producing or importing a given commodity, and between expanding exports or replacing imports as well as the optimal foreign borrowing schemes. Solution to a linear programming model also gives estimates of accounting prices as the opportunity cost of the primary factors in the system as well as of the constraints specified. By correctly choosing the social welfare function and the constraint set of the linear programming model, the dual prices system will represent a competitive market equilibrium. Although it was observed that some of these prices were quite sensitive to the particular specifications, the shadow pricing approach has been generally accepted as a practical tool for the appraisal of investment projects, providing an essential link between a macroeconomic plan and decentralized decisions on the investment projects. The linear programing models are more useful in exploring the implications of interdependence in formulating detailed plans and in expanding the view of development planners, although they have not been widely used in actual planning exercises.

As computer technology and computational algorithms have developed, nonlinear programing models have emerged, which increase the possibility for substitution within the models and hence reduce the instability of shadow prices. A more complex welfare function can be used that may involve distributional as well as growth objectives and lead to social welfare prices resulting in a different allocation of resources. A complex nonlinear general equilibrium system can be incorporated as the constraint set in a nonlinear optimization model in which one seeks to maximize a planner's objective function. Such nonlinear optimization models are useful in exploring the best choice among policy instruments in the presence of various distortions within an economy.

The size of nonlinear programming models is still limited by the current computation abilities especially for the dynamic version. Linear activity analysis has been proved to be effective for a large-scale system with tremendous numbers of engineering parameters. The Stanford PILOT model is a notable example in successfully using a nonlinear social welfare function within a linear activity analysis framework and providing a systematic way in analyzing the effects of technological changes on the US economy (McAllister, Stone and Dantzig, 1988). The PILOT methodology seems to be appropriate for both market and centrally-planned economies where a bulk of engineering information is available and is useful for analyzing the effects of technological changes in the medium- or long-term development.

The socialist countries have a long tradition of using input-output techniques for micro- and macro-modeling, national and regional planning, the determination of prices, the management of production resources, and a great variety of other analytical and administrative purposes. Dynamic Multisectoral input-output models, some quite elaborate, have also been used both for medium-term five-year plans, and for long-run prospective planning. Input-output analysis has dominated the economic scene in the Soviet Union since the 1950s, both in terms of resources devoted to construction of numerous tables and models, and in terms of their applications. Multisectoral input-output models have long been used in various forms to solve the "material balances" problem in quantitative planning. The People's Republic of China has also adopted such a technology since 1980, and many interesting applications could be found in both national and regional planning and forecasting exercises. These works flourished in an environment where direct quantitative controls were the major policy instruments and where the price system was not viewed as an important tool in economic development.

In this paper, we focus on the input-output based multisectoral planning model methodology and applications. Section II presents the basic structures and specifications of the

input-output programming models. In section III, two applications of the multisectoral planning model are presented: the Stanford PILOT model and a multisectoral planning model of China.

2. PLANNING MODEL METHODOLOGY

The purpose of this section is to describe the basic structures and mechanisms of the multisectoral input-output planning model. Any model must be based on a particular underlying economic system theory that can best represent a realistic picture of a nation's economy. The planning model here does not necessarily mean that the economy modeled is centrally planned or directly controlled by a single decision maker. However, we do interpret that the economy is coordinated through a consistent plan jointly made by all the actors within the economy. Thus, for the presentation, we assume that information about future development plans will be available to and implemented by all members within the economy. Therefore, we could deliberately exclude from the model the nature of the ownership of the means of production as well as the individual monopolistic behaviors. The planning model's purpose is to build a linkage between the macroeconomic plan and various decisions of particular sectors or agents, under the constraints of available natural resources so as to achieve the social development objectives agreed upon to some extent by all agents within the economy. The economy could be decentralized and we will show at the end of this section that under a certain optimization framework, the solution to the multisectoral planning model is also a competitive equilibrium.

The multisectoral planning model based on the input-output structure can be formulated in a compact form:

$$\begin{aligned} &\text{Maximize} && W(Z) \\ &\text{Subject to} && (I-A)X - Y \geq 0 \\ & && F(Z_1) \geq 0 \\ & && M \cdot Z_2 \geq b \end{aligned} \tag{1}$$

where $W(Z)$ is a social welfare function to be maximized; X and Y are the gross sectoral output vector and final demand vector; Z_1 and Z_2 are vectors of endogenous variables; while the vector Z is a vector of certain endogenous variables (sometimes called target variables) based on which the social welfare function is being maximized. The first equation in the constraint set represents the material balance constraint with A denoting the input-output matrix. The function $F(Z_1)$ and matrix M represent some nonlinear and linear constraints in the economic system. They may represent nonlinear or linear activity analyses of production or represent some forms of demand systems. The basic requirements for this nonlinear planning model are that the utility function being maximized should be concave and the constraint set as a whole should be convex. It is easy to see that the linear model is a special case of the above formulation. The majority of multisectoral planning models are in physical volume terms. However, prices can be incorporated exogenously or calculated endogenously on a cost markup basis for accounting purposes.

The Objective Function

The objective function is the maximand in the models and it is typically a concave utility function of social welfare, which is usually a continuous function of certain endogenous variables, or target variables. There are many forms of utility functions that can be found in the planning modeling research literature. However, the utility function chosen must be consistent with the relevant social development objectives. It is also possible to specify the objective function as some sort of planner's preference function that need not be based on welfare theory. Depending on the modeling purpose, the target variables could be household consumption, the aggregate income incorporated with employment or the distribution of income, or optimal foreign borrowing. One of the most important development objectives is to maintain a reasonably high rate of economic growth without significant fluctuations. Searching for the optimal economic growth path by selecting the "best" pattern of final demand and resource allocation has been the main feature of most dynamic multisectoral planning models.

One type of widely used objective function is to maximize some form of the welfare function based on household consumption, which is usually based on a declining marginal within period utility and is discounted for each period. The within period utility could be the Geary-Stone type utility function corresponding to well-known linear expenditure systems. The objective can be expressed as

$$\max \sum_{t=1}^{T} \frac{U(t)}{(1+\delta)^{t-1}} + \frac{V(T)}{(1+\delta)^{T-1}} \tag{2}$$

$$U(t) = \sum_{i} \beta_i \log(C_i(t) - \gamma_i)$$

where β_i and γ_i are parameters and $C_i(t)$ is the consumption of the ith commodity, and the social discount rate δ reflects the present value of future consumption. The term $V(T)$ is the terminal utility, which is added in order to take into account the contribution of the present generation to the welfare of the future generation. The specification of $V(T)$ involves what is called the terminal condition.

The System Constraints

There are a variety of constraints that could be imposed on the optimization schemes to capture some technological and institutional limitations on economic development. The technological constraints reflect real limitations on economic growth that could be modeled easily and realistically. Examples are input-output material balance constraints, bounds on total factor uses, the balance of payment constraint, and absorptive capacity restrictions on the amount of investment that can be undertaken in a given industry at a given period. The institutional constraints reflect the action of government or organizations in fiscal and monetary matters, as well as some political and external restriction on economic growth. Examples are upper bounds on savings, lower bounds on the level of employment, requirements for certain regions to grow at specified rates, and upper bounds on imports for protecting domestic industries. For some linear specifications, additional constraints are included for purely technical reason on ad hoc bases, in order to avoid overspecialization such as in trade, "flip-flop" consumption patterns, and other forms of extreme behavior to which linear systems are prone. There is plenty of literature on linear planing models and their shadow price systems. In this paper, we concentrate on the methodology of multisectoral planning models based on dynamic nonlinear programming and leave out any discussion of the dual systems.

Material Balance Constraints. Almost all multisectoral planning models start with the basic material balance equations of the input-output system, which specify that supply equals demand for the output of each sector. The input-output framework is generally based on the assumption that each sector produces a single homogenous commodity and that the gross output is proportional to intermediate inputs from all sectors $i = 1, ...n$. For a particular sector, we have

$$X_i(t) - \sum_{j}^{n} a_{ij}(t) X_j(t) \geq Y_i(t) \tag{3}$$

where a_{ij} is the (i, j) elements of matrix A, defined as the intermediate requirements from sector i per unit of output of sector j at the time period t.

The material balance constraints provide market clearing behavior for all the product markets. Although each of them is written as an inequality that in principle permits free disposal of excess supply, an acceptable solution to the model should contain no excess supply. The material balance constraints usually are stated in constant-price terms (in quantity indexes), sometimes called the real terms, although they hold true in current prices as well.

The matrix A is usually exogenous and derived from an empirical intersectoral transaction table. All elements of matrix A are nonnegative and the sum of each column will be strictly less than one. The coefficients, $a_{ij}(t)$ will be changing over time as the production technology is upgraded during economic development. The coefficient $a_{ij}(t)$ could be generated endogenously and in some models the Cobb-Douglas specification is used so that the a_{ij}

stays constant in value, not in volume, terms. Therefore the technology choice can be made by price-induced substitution.

The method of using engineering information to describe alternative technologies has been adopted in many planning models by adding some alternative columns for the same sector in the input-output matrix. The Stanford PILOT model has elaborate specifications using this method for the technology choice. This method has also been used in some models to study possible employment-creating effects of alternative techniques in labor intensive sectors such as the construction sector and some subsectors in agriculture.

The final demands, $Y_i(t)$, in the material balance equation are normally disaggregated into private consumption, government consumption, and investment requirements. These different final demands either can be forecast exogenously or computed endogenously in a variety of ways with different degrees of details, depending on the issues the model addresses. For the model of an open economy, the simple input-output framework can be extended to include foreign trade in various ways. At one extreme, imports are assumed to be perfect substitutes for the domestically produced goods and hence exports and imports are simply treated as additional sources of demand and supply. At the other extreme, imports are viewed as completely non-competitive and hence can not be included with domestically produced goods in the material balance constraint for a given sector since they are by definition not comparable.

Production Capacity Constraints. The levels of production performed in an economy with certain technology are mainly decided by available production resources or factor inputs. Beside the limitation of natural resources, capital and labor are two important factors that determine the development level of an economy. The production technology describes the relation between sectoral output and factor inputs. We can present production technology in a transformation function, $F(K, L)$, if only capital and labor are considered. Thus, we have

$$F_i[K_i(t), L_i(t), t] \geq X_i(t) \tag{4}$$

where $K_i(t)$ and $L_i(t)$ are capital and labor available for sector i at the time period t. The function $F(K, L)$ is supposed to be homogenous of degree one (constant return to scale) and to satisfy various regularity conditions. The adoption of nonlinear production functions makes price-induced substitution possible. Although some doubt the reliability of this non-technological relationship, many modelers have increasingly chosen the nonlinear continuous production function in their models. The commonly used production functions are the Cobb-Douglas and constant elasticity of substitution (CES) production; both have price-sensitive substitution effects with unitary price elasticity for the former and a wide range of positive constant elasticity for the latter.

In many planning models for developing countries, Leontief production technologies are used, that is, the sectoral output is proportional to available capital and labor:

$$\frac{K_i(t)}{a_{K_i}(t)} \geq X_i(t) \tag{5}$$

$$\frac{L_i(t)}{a_{L_i}(t)} \geq X_i(t) \tag{6}$$

where $a_{K_i}(t)$ and $a_{L_i}(t)$ are capital-output and labor-output ratios respectively of sector i at the time period t. Surplus of labor is a popular phenomenon in developing countries; many models use labor-output ratio to determine industrial employment with fixed wage rates and remaining labor is placed in the agriculture sector. If planners want to investigate the effects of education reform and skill updating, they can do so by subdividing labor into skill groups in the production function. This kind of scheme is not appropriate for a country with an extreme surplus of labor during the development planning periods.

Foreign Trade and Balance of Payment Constraints. Foreign trade problems are important and they almost always existe in developing countries that are often short of the foreign exchange necessary to expand their production capacity and improve the production technology.

It is often assumed that the model economy trades with a single entity called "the rest of the world," allowing unique import and export prices to be defined. In models of developing countries, the "small country" assumption is often adopted, which means that the shares of the country's imports and exports in the world trade volume are too small to affect the world price, that is, world prices of imported or exported goods can be fixed for the economy modeled. However, some models specify constant elasticity of export demands in order to reflect a realistic picture of foreign trade, thus introducing nonlinear effects. Furthermore, the supplies of imports are often classified into competitive imports and noncompetitive imports in many models of developing countries. Competitive imports are completely perfect substitutes for domestically produced goods, while the non-competitive imports are the goods for intermediate as well as final use for which no domestic capacity existes, and for which no substitution by domestic output is possible, whatever the relative price and domestic outputs may be. The treatment of competitive and non-competitive imports makes it possible to recognize the technical, political, and administrative restrictions on the use of foreign imports. The dollar prices for these different kinds of imports as well as exports in the same sector are usually different due to the fact that the commodity composition of imports and exports can differ or that the average tariff on imports can be different than the average duty or subsidy on exports.

The foreign exchange constraint requires that the foreign exchange that can be spent on imports in any time period be less than or equal to foreign exchange revenue generated in that same period through exports and total foreign capital inflow, $F(t)$. In equation terms:

$$\sum_{i=1}^{n}\Big[PM_i^c(t)M_i^c(t) + PM_i^{nc}(t)M_i^{nc}(t)\Big] \leq \sum_{i=1}^{n} PE_i(t)E_i(t) + F(t) \tag{7}$$

where $M_i^c(t)$ and $M_i^{nc}(t)$ are competitive and non-competitive imports; $PM_i^c(t)$ and $PM_i^{nc}(t)$ are the dollar prices of competitive and non-competitive imports; and $E_i(t)$ and $PE_i(t)$ are the exports and their dollar prices.

The non-competitive imports are complementary to domestic production; the requirements for such imports must be satisfied before other supplementary imports are allowed. In planning models, the noncompetitive imports are either exogenous or linked to output and/or investment levels by fixed coefficients. However, these coefficients can be changed period by period to take account of planned programs of import substitution.

While competitive imports are the substitutes, the foreign exchange available to purchase competitive imports can be limited. Competitive imports are usually generated in the model according to comparative advantage. In the absence of diminishing returns, however, the model solution would tend to concentrate the use of competitive imports in only a few sectors. To the extent that diminishing returns existed, this would unduly restrict domestic capacity formation in those sectors. To avoid the possibility of overspecialization, the maximum use of foreign exchange for competitive imports is often stipulated for each sector in the planning model. However, in some situations, the above constraints can be released so that the model has complete freedom to allocate foreign exchange for competitive imports.

Exports are often fixed exogenously or set to grow at least at certain minimum rates in each period. Capital inflow can be either given exogenously or formulated elaborately to consider various kinds of capital inflows. Some planning models focus on foreign borrowing strategy analysis by formulating a different debt structure in a dynamic context. Thus, foreign capital inflow in each period can be generated as an optimal solution for using external resources.

Capital Stock Formations. The popular specification for capital stock formation is using a matrix of capital share coefficients to translate demands for investments by sector of destination into demands for investment goods by sector of origin, $H(t)$, a column vector of final demands in the material balance equation. Thus, we have

$$H(t) = S(t) \cdot D(t) \tag{8}$$

where $D(t)$ is a vector of capital formation and $S(t)$ is the capital coefficient matrix, the elements of which s_{ij} represent the proportion of capital stock in sector j originating in sector i and $\sum_i s_{ij} = 1$ for all j. It is usually assumed that coefficients s_{ij} do not change in response to changing capital goods prices.

The above specification assumes that the capital stock in each sector is a well defined combination of various commodities and that its composition differs across sectors. Therefore, the capital stock in each sector is a special aggregate commodity that can not be directly shifted to other sectors without losses (in recombination). The "putty-clay" assumption of capital stock is often used for most dynamic models, which means that once capital stock is installed in any sector it is "fixed" in place. However, some models assume homogenous capital stock composition for all sectors and assign newly produced capital to each sector according to certain kinds of investment reward schemes; in this case, sectors should have been disaggregated in a limited pattern so that there would be no pure sectors.

The capital investment gestation lags could be incorporated in the above specification. This is important for a model dealing with short period dynamics where a realistic two- or three-year structure of gestation lags could provide detailed growth prospects with stocks of capital goods-in-process.

Dynamic Linkage Equations. The dynamic model connects each period by defining intertemporal equations that serve to update some endogenous variables. One important dynamic linkage between periods is the capital stock updating equation:

$$K_i(t+1) = (1-\delta_i)K_i(t) + D_i(t) \tag{9}$$

where δ_i is the fraction of capital stock depreciated each period for the ith sector; $D_i(t)$ is the capital stock newly formed at the end of period t in the ith sector.

Other dynamic linkage equations could be population growth, foreign debt updating etc., depending on whether the model specifies these flow variables endogenously or not.

Initial and Terminal Conditions. The initial conditions describe the endowments of the economy at the beginning of the planning horizon. The endowments constrain production capacity and consumption during the planning period. In most dynamic models, the endowments are capital stocks and foreign debt which should be repaid during the planning horizon or in the future. Therefore, the initial conditions are simply the value assignment of exogenous variables that are crucial to affect the dynamic path. The terminal conditions specify how much wealth should be left and how it should be composed for the future generation. The need for terminal conditions arises out of the necessity to prevent the depletion of the economy's endowments due to the nature of the programming models. In other words, the terminal conditions are those specifications that ensure the availability of capital and limited foreign debt to future generations.

As we are dealing with finite horizon applied planning models, what would occur after the terminal period will not be explicitly covered. We are mainly concerned with these dynamic linkages between the post-terminal period and the period covered by the planning model. In particular, we have to specify certain flow variables of the post-terminal period whose values are decided by certain updating equations of the terminal period. This can be done either by restricting these flow variables, by setting some lower bounds, or by putting valuation on these flow variables and then attaching them to the social welfare function being maximized. Examples of these flow variables are the sectoral capital stocks and the level of foreign debt. The welfare weight given to these flow variables may acquire special meanings. They can serve as the price of the terminal capital stock or the cost of the terminal foreign debt. By varying these weights, alternative optimal trajectories can be generated. The planner could set these terminal weights at such a high level so that all the efforts during the plan period would be devoted to building up the capital stocks or to lowering the level of foreign debts.

One alternative scheme is to set these updating variables to grow at least at a certain minimum rate for each period to ensure that the terminal capital stock or foreign debt remain within a certain range. There are many discussions and suggestions in the literature on the setting of the terminal condition, but little agreement on how terminal conditions actually ought to be set. We have hence here discussed only the methods commonly used in many planning models.

Programming Models and General Equilibrium

In the past decade, multisectoral computable general equilibrium (CGE) models, introduced by Adelman and Robinson (1978), have come into wide-spread use in developing countries to analyze many policy issues such as income distribution and structural adjustment. Technically, the CGE models consist of a set of simultaneous equations that are formulated by simulating the functions of a market economy in which suppliers and demanders interact across markets in response to price signals. The input-output techniques are used in most of these CGE models and have been extended to form a framework of social accounting matrix (SAM) that forms a linkage between sectoral material production and national income distributions and expenditure flows. In its purest form, a CGE model provides an empirical implementation of the Walrasian model of production and exchange under perfect competition; prices and quantities of traded goods and factors are typically determined simultaneously. The CGE models for developing countries have incorporated many structuralist features and imposed many rigidities and policy instruments, such as tariff, subsidies, and limitation on factor mobilities.

The CGE models have been increasingly used in many developing countries for the analysis of a variety of development policy issues. It seems that with the emergence of the CGE models, we have been, to some extent, the passing of the planning models based on linear programming that were quite fashionable a decade ago. The strong linearity and the weak theoretical underpinnings are the main drawbacks of these planning models, which make it difficult for them to provide a reasonable representation of an actual market economy.

However, multisectoral planning models based on optimization framework do have some unique features that the CGE models cannot easily overtake. The CGE models have proved to be powerful tools for developing policy analysis, but their applications have been limited mainly to market-oriented or mixed developing economies. For developing countries with a socialistic economic system, the basic assumption that underlies the CGE models is questionable. In its purest form, CGE models represent a Walrasian equilibrium that reflects the competitive market behavior of decentralized economic agents, while in most developing socialist economies production is characterized by the monopoly of state enterprises. Although economic reforms such as market mechanism and decentralization have been introduced in many socialistic countries in order to improve their economic efficiency, the application of CGE models for these countries are still in the initial development stages. In the pre-reform socialist developing countries, the central planning authorities have strong control over investment and production as well as income distribution; physical planning and administrative allocations are basic characteristics. Before the reform of the Chinese economy, public ownership predominated throughout the country and production was almost completely organized by public sectors; hence, there is no interaction between suppliers and demanders in response to prices. The government was the only decision maker with absolute power to affect production and income distribution. Therefore the programming model could better fit the past planned economy in China.

In addition, the multisectoral planning models can incorporate large portions of technical data for engineering analysis of technology development. Unlimited inequality constraints can be imposed on the model to reflect many technological as well as institutional restrictions on economic development. Many planned economies have accumulated many physical flows and technological data that could be better used in programming models. In recent years, there have been major advances in computational algorithms that facilitate the solution of very large scale linear and moderate large nonlinear programming problems. Thus the role of planning

models based on optimization has been reappraised by many economists and modelers. Many nonlinear specifications have been incorporated as a constraint set to nonlinear programming models or have been linearized together with a linear objective to form a large scale linear programming. Therefore, substitution effects can be greatly enhanced in the planning models to analyze of certain development issues.

A planning model with certain specifications can give the solution that represent a competitive general equilibrium. It has already been proved by Negishi (Negishi, 1960), that a Walrasian competitive equilibrium can be reproduced as the solution of the maximization of a linear social welfare function defined as the weighted sum of the utilities of the consumer groups in the model, where the weights on the welfare function are strictly positive and must be chosen endogenously. Stone has developed and assessed a number of equivalent mathematical formulations of general equilibrium and proved that the traditional representation of general equilibrium can be alternatively represented as an nonlinear optimization problem with linear production (Stone, 1988). He found also that these results can be extended to the case of nonlinear production with constant return.

Starting from Koopmans's activity analysis, Ginsbergh and Waelbroeck (1981) have developed a complete set of elaborate methodologies with a clear theoretical foundation based on Negishi's welfare optimum to solve economy-wide competitive equilibrium, which they have called the AGE (activity analysis general equilibrium) approach. They have constructed a well-known GEM model of the world economy using this approach, based on a linear programming with a linearized objective function. The Stanford PILOT model is also a general equilibrium model using a large scale dynamic nonlinear programming based on linear activity analysis. The PILOT model formulations demonstrate that for a correctly chosen maximand, the shadow prices from a properly specified nonlinear programming are equivalent to the market prices of general equilibrium. It is therefore very reasonable to expect that a multisectoral planning model based on nonlinear programming can be further developed that will fit both market and planned economies.

3. APPLICATIONS OF MULTISECTORAL PLANNING MODEL

Stanford PILOT Macroeconomic Model

PILOT, an acronym standing for Planning Investment Levels Over Time, is a dynamic general equilibrium model of the United States economy developed over several years by the System Optimization Laboratory at Stanford University. The model was first constructed in 1975 based on linear programing, starting from a quite modest multi-period process model of the energy sector coupled in each period with a 12-sector aggregation of the Leontief input-output framework with some additional activities for foreign trade and capital formulation.

Through many years of development, the model has been transformed into a full blown macro model of the general economy with detailed representation of consumption, investment, and production activities. The model is now a large-scale dynamic general equilibrium model of the U.S. economy. Solutions to the model provide a coherent overall picture of the long-term consequences of innovation, modernization, conservation, and foreign market conditions in the context of technological change. In particular, the model can be used to assess the effects of different time paths of availability and/or technical characteristics of new technologies.

The Model Structure. The model is essentially a nonlinear programming model with a quadratic objective function, which represents the per capita utility based on consumption over time. The utility function is derived from a system of integrable consumer demand functions, which is linear to income and price-sensitive. The model maximizes this concave utility function subject to a convex set of the physical flow constraints, yielding an optimal solution that is consistent with the general equilibrium conditions. The input-output framework is the core of the model's formulation. In order to take full advantage of engineering analysis as the

basis for defining new or hypothetical technological alternatives, the production sector is modeled using the proven techniques of linear activity analysis (or linear process modeling). The general equilibrium framework integrates such a representation of production with a smooth consumer demand function. Actually, the activity analysis representation of production goes beyond the traditional fixed coefficient input-output analysis by defining multiple columns for each sector representing alternative production technologies from new investments. The energy-producing sector of the economy as well as energy-use decisions in manufacturing and in residential and commercial building are represented via modular process models. The productive capital stock is vintaged and unmalleable once in place. The domestic industries compete with imports (whose prices are given exogenously) and face a downward-sloping export demand schedule (also exogenous). The exchange rate is determined endogenously. The final demand is represented via the government input-output sector, endogenous investment activity in all sectors, and consumer demand for 13 aggregate commodities. The per capita demands for five of these consumption aggregates are specified exogenously. The demand for the other eight aggregates are functions of prices and income, modeled via an aggregate quadratic utility function.

As PILOT is currently formulated, it directly represents economic activities and relations in each of nine discrete time periods five years in length, covering the time horizon 1975-2015. The model contains a lot of engineering data and divides the US economy into 40 production sectors, each having several columns in each time period in order to allow for a free choice of various degrees of modernization and innovation. A given activity representing a long-lived investment could have coefficients in the relations of many subsequent periods. It also has coefficients in relation to one or more previous periods to the extent that the investment requires substantial construction lead time. Since decisions on undertaking these long-lived investments are made by perfect foresight, the equilibrium solution must be calculated simultaneously across all time periods. The model is now used for assessing the long-term effects of innovation, modernization, conservation, foreign competition, and imported energy prices on the growth of the economy and its various sectors.

A Multisectoral Planning Model of China

The economic development of China has followed a different trajectory than that of other developing countries in the world. The Soviet-type economic planning dominated in China from the early 1950s until the late 1970s. It is characterized by exclusively over-centralized planning. Since the late 1970s, the economic reforms have brought the Chinese economy to a somewhat decentralized structure, and a so-called planned market economy is on its way to being realized. The market mechanism has begun to play to some extent its role in supply-demand equilibrium of commodity production and consumption, especially in the agricultural sectors. However, in general, the macro-economy is still controlled by the central government and planning will continue to be an important task of the whole economy.

We have developed a multisectoral planning model for the Chinese economy based on the input-output framework using a nonlinear programming technique. The household consumption demands are the main driving force of the model, which is achieved through the maximization of a social welfare function. The model is based on the current Chinese economic structure, which assumes consumers and producers are interacting through the product markets. The market clearing conditions are achieved by the government's ingenious planning, which will maximize the "standard of living" (in terms of a social welfare function) over the planning horizon. In China, formal although formal capital and labor markets have not yet been established, product markets do exist with certain kinds of distortions. Under such specifications the model's solution will satisfy competitive equilibrium conditions with the dual prices of sectoral outputs as the market clearing prices of the product markets. The model exclusively contains the specifications of national income generation and distribution; consistency has been imposed between the real sphere and control sphere. Many rigidities are imposed on wage rate, sectoral bonus, and employment. Capital is the only factor that affects the growth and its accumulation is generated by the model according to the optimal level of domestic saving and foreign capital inflow. Thus, the model solution for product markets does

not represent a neoclassical free-market equilibrium. But it does represent a market equilibrium solution constrained by technical and institutional specifications believed to represent a realistic picture of the Chinese economy.

The model can provide the central planner a possible way to use the market clear prices as accounting indexes for the allocation of production resources and consumption goods. Economic growth can be assessed using the model by varying different technological parameters or institutional policy instruments according to different scenarios. The model's solution provides the optimal sectoral allocation of production and consumption and the optimal level of foreign borrowing. Many structural parameters can be adjusted to reflect the government preference on the pattern of economic growth. The model would be a useful tool for the analysis of structure changes in the real economy as well as providing a guidance for macro level economic control over aggregate investment and consumption.

The overall model consists of four modules: a social welfare module, an input-output physical flow module; a foreign trade module; and a national accounting module. A flow chart of the integrated economic planning model is shown in Figure 1. Each module is described in the following section.

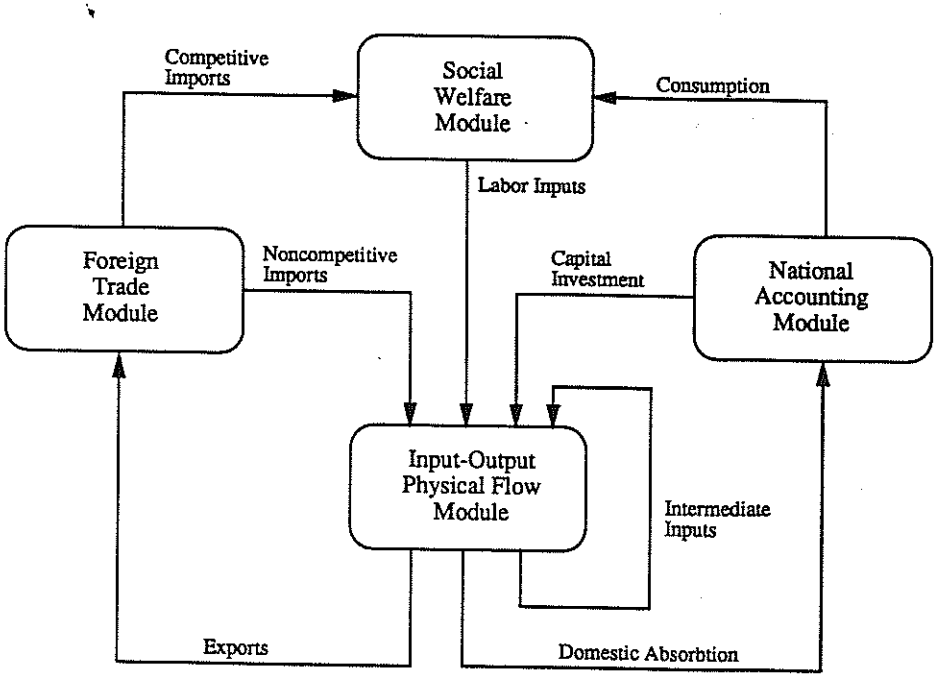


Figure 1. Integrated Economic Planning Model: Diagrammed Representation

Social Welfare Module. The social welfare module consists of an objective function of social welfare as well as the specifications of population and labor force growth, household consumption determination, and the terminal conditions regarding the welfare of the future generation. This module is designed to assist the planners in analyzing different social development schemes.

Since household consumption is one of the most important indicators of the social welfare, we choose the objective function as a discount sum of within period utility of household consumption plus the valuation of the total terminal capital stock and the terminal foreign debt. The within period utility function is a quadratic function of average household consumption

derived by Dantzig et all (1988) from an integral demand system that is linear to income and price-sensitive. Professor Dantzig has shown that a general equilibrium problem in a market economy can be converted to one that maximizes a nonlinear (utility) function subject to physical flow constraints by carefully choosing the normalization rule for the price vector used to define the aggregate demand function. This innovation makes it possible to use the the general equilibrium mechanism in modeling centralized planning economy. A preliminary investigation of Chinese household consumption patterns has proved that the aggregate demand function is approximately linear on a broad range of household incomes at fixed prices. Therefore, we could estimate such a quadratic utility function that represents price-sensitive expenditure choices among a class of consumer goods and services.

The household consumption categories in the utility function are classified as food, clothing, household operation, personal care, and service. These categories differ from the sectoral classification of production, and a transformation matrix is used to convert the household consumption demand vector to the sectoral consumption demand vectors. The consumers are divided into two groups: the urban households and the rural households, since the consumption behaviors of these two groups are quite different in China.

The population growth is strongly regulated by the various levels of government. The population growth rates in each period are set exogenously to reflect government decisions on population control. The total labor force in each period is defined as a function of population and population age structure.

Input-Output Physical Flow Module. The input-output physical flow module contains all the physical flow specifications of the planning model. The material balance constraint based on input-output technique is the core of the module and its dual prices are used for national accounting. The factors' requirements of production as well as their accumulation are formulated to reflect the special features of China's production systems.

The production sectors, especially the industries', are exclusively dominated by the state and collective ownership. The factor markets do not traditionally exist. Therefore, factors input demands are basically determined on the basis of engineering analysis. Following the Chinese classification, the model distinguishes two kinds of capital stocks: fixed and circulating capital stocks. In each period, the gross outputs of each sector are bounded by the total sectoral capital stocks available at the beginning of the period according to fixed capital-output ratios which vary across sectors. The sectoral labor inputs for production are rigidly related to sectoral output. Specifically, employment in each sector except the crops sector is determined by the gross output and an exogenous labor productivity coefficient (labor-output ratio). However, the employment in the crop raising sector, where labor is assumed to be in surplus, is the difference between the total labor force and the employment in all other sectors.

In all sectors, production requires not only capital and labor but also specified amounts of intermediate inputs that are determined by input-output coefficients. Input-output coefficients reflect the state of technologies in different stages of development. It is therefore reasonable to expect some changes in these coefficients during the course of economic development and reform. However this possibility is exploited in the model only in a few sectors (specifically, energy-related sectors) for the lack of adequate information.

The sectoral fixed and circulating capital accumulations are needed if production in the next period is to be expanded. The dynamic linkage equations are used to update the sectoral capital stocks. The total sectoral capital stock accumulations are determined by the intertemporal optimization conditions. The capital stock accumulations generate the demand for investment goods of origin that are obtained through the use of capital coefficient matrices (for fixed and circulating capital investments respectively).

Foreign Trade Module. The foreign trade module is formulated based on China's foreign trade structure, incorporating the current "Open Door" policy. Following the small-country assumption, the world prices of Chinese imports and exports are assumed to be fixed by world market

conditions. The supply of imports to China is infinitely inelastic at that price. Sectoral exports are set to be at least at certain amounts in each period. This minimum amount of export is proportional to the gross sectoral outputs; however, the proportion will grow during each period at a specified rate to reflect the planner's bottom line expectation of the domestic export supply growth potential and the absorption by the world market for China's export. Sectoral imports are divided into two portions: competitive and non-competitive. Imports are non-competitive in the sense that these are the goods for intermediate as well as final use for which China has limited capacity to produce in the current stage of development. Thus, we do assume that the non-competitive imports are perfect substitutes. The amount of noncompetitive imports are set to be related to gross output level by fixed coefficients. Competitive imports are goods for which domestic capacity can be created. They are, therefore, allocated among consuming sectors. The choice between the export promotion and competitive import substitution is made within the model's solution to achieve the optimal level.

The total amount of imports in each period is limited by the availability of foreign exchange. This in turn depends on the total amount of exports and foreign capital inflows in the same period. The model distinguishes four different sources of capital flows: net foreign transfer (including remittance flow); net foreign borrowing (foreign borrowing minus amortization); interest payments of foreign debt; and changes in official foreign exchange holdings. The foreign borrowing is referred to as foreign aid from government sources, private foreign investments and so forth. The foreign debt is updated each period as the result of foreign borrowing and interest payments. We assume that the total amount of foreign debt that is controlled by the government in each period to be no more than a fraction of the national income. This fraction reflects the government attitude to acceptable foreign debt. The solution to the model will give the optimal level of foreign borrowing for each period.

National Accounting Module. The national accounting module consists of several central policy constraints on income distribution and capital accumulation as well as on consumption and savings determination. In this module, the general equilibrium prices, i.e. the dual variables corresponding to material balance constraints, are used to calculate national income and its distribution. This specification will facilitate planners to make decisions on the divisions of investment and consumption according the the product market equilibrium conditions, thus maintaining consistency between the real sphere and the control sphere of the economy. This specification would help the planners in designing price control policies as well as achieving realistic macro-monetary planning objectives.

Total national income is divided into household consumption, government consumption, and domestic savings. Household income consists of two portions: wage income and bonus income. The urban and rural household savings rates are given exogenously, while the levels of sectoral consumption by urban and rural households are achieved by utility maximization. Total government consumption is assumed to be a fixed proportion of national income. The sectoral consumption by government is determined by constant expenditure shares. The discrepancy between total national income and household and government consumption is used as domestic saving. Finally, aggregate domestic savings and net foreign capital inflow form the total investment funds that should be enough to meet the needs of investment goods for the requirements of new capital formation.

The module has elaborate specification on the sectoral wage rate and bonus determination. The wage rates in each sector depend on the overall level of wages in the economy. The wage ratios of different sectors are fixed , and such rigidities are likely to persist during the plan period to reflect roughly the existence of labor skill differentials among sectors. However, the real wage rates of all sectors are specified to grow endogenously at the same rate relative to the initial period. The sectoral profits are defined as the gross sectoral output minus the costs of intermediate inputs and labor inputs. We assume that each sector can take a fixed fraction of its profit as bonus to the workers and that this fraction is controlled by the government. In the crops sector, the bonus could be its entire profit, since its profit is solely the farmers' self-employment income. The endogenous wage growth rate is an essential equilibrating variable for achieving the intertemporal efficiency on division of consumption and investments. One can

also fix the wage growth rate to reflect the government planned wage level and let the sectoral bonus vary to achieve the desired consumption growth. The sectoral wage and bonus are split among urban and rural households according to specified shares. These wage and bonus share parameters essentially reflect the government policy on income distribution among rural and urban households as well as the economy's structure regarding the scope of the economic activities performed in the rural area.

The model described above is still in its early stage of development. We have run an initial experiment model with limited equations and available data. The experiment model is tested with a simplified structure on the IBM 370 machine using GAMS/MINOS software, which combine MINOS nonlinear programming algorithm and the GAMS high level modeling language (Meeraus, 1985). The solution is achieved by iterating on the sectoral product price vector. At the first stage, GAMS/MINOS is employed to find the optimal solution on an arbitrary accounting price vector, and then, the dual prices generated from the solution are modified and plugged into the model instead of the original price vector and the model is run again. This procedure is repeated until the changes in dual price variables from one step to the next satisfy a break-off criterion. We are going to enlarge the data base required to run the complete model and to design some scenarios so that the model can be used to analyze the growth and policy issues of development, which are believed to be in great need in the current Chinese economic reform process.

REFERENCES

Adelman, Irma and S. Robinson (1978). Income Distribution Policy in Developing Countries: A Case Study of Korea (Stanford University Press, Stanford).

Chenery, Hollis B. (1984). "The Evolution of Development Planning", Journal of Policy Modeling, Vol. 6, no. 2, pp. 159-174.

Chow, Gregory C. (1985). The Chinese Economy (Harper & Row, New York).

Ciaschini Maurizio (ed.) (1988). Input-Output Analysis: Current Development (Chapman and Hall Ltd., New York).

Cohen, S. I., P. A. Cornelisse, R. Teekens, and E. Thorbecke (1984). The Modelling of Socio-Economic Planning Process (Gower Publishing Company, Vermont).

Dantzig, George B., Patrick H. McAllister and John C. Stone (1988). "Deriving a Utility Function for the U.S. Economy", Technical Report 88-6, Systems Optimization Laboratory, Department of Operation Research, Stanford University.

Dervis, Kemal, Jaime De Melo, and Sherman Robinson, (1982). General Equilibrium Models for Development Policy (Cambridge University Press, Cambridge).

Ginburgh, Victor and Jean. Waelbroeck (1981). Activity Analysis and General Equilibrium Modeling (North-Holland Publishing Co., Amsterdam).

Ginsburgh, Victor and Sherman Robinson (1984). "Equilibrium and Prices in Multisector Models", in M. Syrquin, L. Taylor, L. E. Westphal (ed.), (1984) Economic Structure and Performance (Academic Press, New York).

Li, Xuezeng, S. Yang, and J, He (1985). "The Structure of China's Domestic Consumption", World Bank Staff Working Paper No. 463 (The World Bank, Washington, D.C.).

Lluch, Constantino, Alan A. Powell, and Ross A. Williams (1977). Patterns in Household Demand (Oxford University Press, Oxford).

Lu, Yingzhong (1985). "An Introduction to the Intergal Approach of Energy-Economy Model of PRC", Working Paper, Institute of Nuclear Energy Technology, Tsinghua University, Beijing, China.

Manne, Alan S. (1974). "Multi-Sectoral Models for Development Planning: A Survey", Journal of Development Economics, Vol. 1, pp. 43-69.

Manne, Alan S. (ed.) (1984). Economic Equilibrium: Model Formulation and Solution, Mathematical Programming Study No. 23 (North-Holland Publishing Co., Amsterdam).

McAllister, Patrick H., John C. Stone,and George B. Dantzig (1988). "Analyzing the Effects of Technological Change: A Computable General Equilibrium Approach ", Technical Report SOL 88-12, Systems Optimization Laboratory, Department of Operation Research, Stanford University.

Meeraus, Alex (1985). "An Algebraic Approach to Modeling", Journal of Economic Dynamics and Control, Vol. 5, no. 1, pp. 81-108.

Nigishi, T. (1960). "Welfare Economics and the Existence of an equilibrium for a Competitive Economy", Metroeconomica, Vol. 12, pp. 92-97.

Pyatt, Graham and Jeffery I. Round (ed.) (1985). Social Accounting Matrices: A Basis for Planning (The World Bank, Washington, D.C.).

Scarf, Herbert and John B. Shoven (1984). Applied General Equilibrium Analysis (Cambridge University Press, Cambridge).

Shi, Xia (1988). Modeling Chinese Economy - A General Equilibrium Approach (forthcoming Ph.D. dissertation), Stanford University.

Stone, John C. (1988). "Formulation and Solution of Economic Equilibrium Problems", Technical Report 88-7, Systems Optimization Laboratory, Department of Operation Research, Stanford University.

Taylor, Lance (1983). Structuralist Macroeconomics: Applicable Models for Third World (Basic Books, New York).

Weyant, P. John (1985). "General Economic Equilibrium as a Unifying Concept in Energy-Economic Modeling", Management Science, Vol. 31, no. 5, pp. 548-563.

World Bank (1985). China: Long-Term Issues and Options (Washington, DC 20433).

Zalai, Erno (1985). "A Nonliner Multisectoral Model for Hungary: General Equilibrium Versus Optimal Planning Approaches", in A.C. Kelley, W.C. Sanderson, and J.G. Williamson eds., Modeling Growing Economies in Equilibrium and Disequilibrium (Duke University Press, Durham).

Energy Vol. 15, No. 3/4, pp. 341–344, 1990
Printed in Great Britain

0360-5442/90 $3.00 + 0.00
Pergamon Press plc

3.3. INTEGRATED MODELS+

WALTER C. LABYS

Resource Economics Department, West Virginia University
Morgantown, West Virginia
U.S.A.

and

JOHN P. WEYANT

Energy Modeling Forum, Department of Engineering-Economic Systems, Stanford University
Stanford, California
U.S.A.

Abstract - Today there is an increased tendency to integrate different kinds of modeling methods so as to provide a more complete and accurate model of the total system being considered. Integrated modeling theory is discussed and examples provided.

We have discussed the growing tendency among engineering-economic modelers to use two or more of modeling methodologies in combination. In particular, as the modeling process has become more sophisticated, analysts have begun to borrow skills from one another, so that the methodologies which have emerged are more difficult to classify. A glance at several recent model anthologies such as that by Lev (1983), Lev et al. (1984) and Quirk et al. (1982) will confirm the extent to which such integration has become widespread. This phenomenon has arisen not only because of the need to couple economic and engineering considerations, but also because of the interrelationships that exist between the energy sector and the macroeconomic characteristics of national economies. Although this integration has taken a variety of forms, it usually proceeds by model hybridization or model linkage.

Hybridization basically involves constructing models that combine two or more modeling methodologies. For example, it is not uncommon for the demand side of an energy model to be constructed econometrically, while the supply side might be principally engineering in character. In such cases the output of several different models might be needed to provide a comprehensive analysis of a particular energy problem. The models to be combined have tended to be of a programming nature, often also linked to some form of econometric model. The most comprehensive effort to develop a methodology to deal with hybrid energy models is the "combined" energy model approach of Hogan and Weyant (1982 and 1983). A theoretical as well as a computational approach is presented for reaching equilibrium solutions with a combined set of models. Each set of the processes or submodels are placed within an optimization framework with the inputs of each process being the outputs of others. All processes are operated simultaneously to achieve the greatest value of some objective function. The framework is organized in a way that promotes a consistent theory, natural data organization modular design, decentralized implementation, and efficient computation. The solution of the model system is also aided because the submodels provide natural limits for employing model decomposition methods. The actual model or system solution takes place with an interactive method that moves

+This chapter represents an update of an earlier paper published by W. C. Labys together with D. O. Wood, "Energy Modeling," in W.A. Vogely (Ed.) Economics of the Mineral Industry, Society of Mining Engineers, Littleton, CO., 1985.

through the network of process models, solving and resolving the optimality conditions for each of the component models, while checking and updating the values of the variables in the connecting links.

The straightforward pursuit of the interative process, sequentially visiting the component models while updating the demands for inputs and the prices of outputs (known as the method of successive approximations), is the network process model generalization of the simple cobweb search suggested by Figure 1.

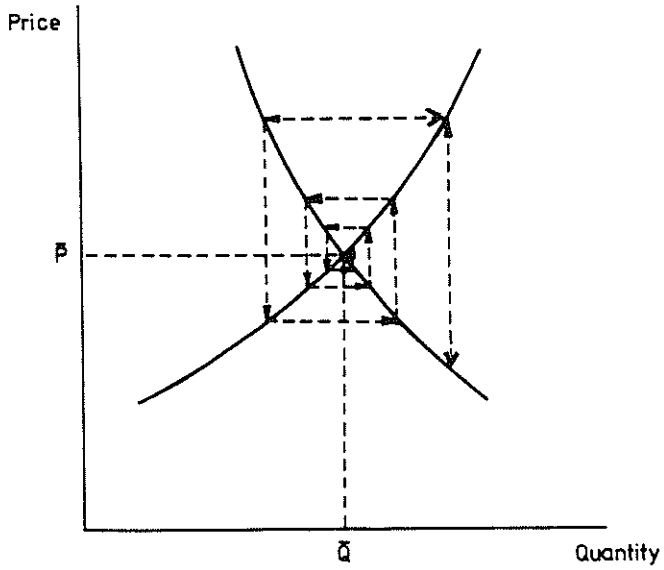


Figure 1. Cobweb Search for Equilibrium

An early application of this approach can be seen in the Project/Independence Energy Evaluation System (PIES) constructed for the FPA; this model consists of a macroeconometric model, an econometric demand model, and a programming model explaining fuel supplies, conversion and shipments (see EIA, 1979). An integration procedure was employed to solve over the three submodels to compute what would be equilibrium levels of energy demand and supply consistent with related macroeconomic activity. This model exists in the form of the Mid-range Energy Forecasting System (MEFS) and more recently the Integrated Fuel Forecasting System (IFFS) (see IEA, 1982). The MEFS model includes econometric models of energy demand, linear programming models of refinery operations and electric power generation, network optimization models of the transportation of energy fuels, engineering process models of investment and production for energy supply development, and an interactive procedure for achieving equilibrium of the overall network. The IFFS model in contrast, is organized more about the markets for individual fuels, with the equilibria in other fuel markets.

The above applications of hybrid programming appear to have resulted in closely related energy equilibrium models. The Canadian BALANCE model constructed by Damel and Goldberg (1981) has dynamic convergence properties similar to MEFS and these are enhanced by the potential of increasing own-price demand elasticity approximations to account for the effects of likely price changes in prior periods. The SRI-Gulf model and its descendants as developed by Decision Focus Inc. (1979) utilized this approach to solve an extensive system of process models.

One of the best known examples of the integration of an energy model with a model of a complete economy is that constructed by Hudson and Jorgenson (1974). This model consists of

a macroeconometric growth model of the U.S. economy integrated with an interindustry energy model. The growth model consists of submodels of the household and producing sectors, with the government and foreign sectors taken to be exogenous, and it determines the levels and distribution of output valued in constant and current dollars. The model determines the demand for consumption and investment goods, the supplies of capital and labor necessary to produce this level of output, and the relative equilibrium prices of goods and factors. The model is dynamic and has links between investment and changes in capital stock and between capital service prices and changes in prices of investment goods. It has been used to forecast long-term developments in energy markets within the framework of a consistent forecast of macroeconomic and interindustry activity. Another application has been to analyze the impact on energy demands of alternative tax policies, including a uniform Btu tax, a uniform energy sales tax, and a sales tax on petroleum products.

Among other energy-economy models, Groncki and Marcuse (1980) report on the BESOM model which combines a programming model of energy supply with a long-run macroeconomic growth and input-output model. One example of the use of the model has been to analyze the economic impacts of fuel scarcities. Constraints can be placed on the availability of fuels and resources, and the required fuel substitutions are determined. Coefficients in the input-output model are revised to reflect the new fuel mix, and the model is again solved with the revised mix. Several interactions are required between the two models in order to obtain a solution in which the energy demands and fuel mix are consistent in the two models. Impacts in the macroeconomic growth of various industrial sectors in the economy are then evaluated.

While the above models analyze energy-economy interactions in a domestic context, research has also taken place in modeling international interactions. The most well known example of such a combined model is the Project LINK system developed by Klein which attempts to evaluate trade and payments adjustments at the global level (see Ball, 1973). Economies of individual countries constitute the submodels and overall trade equilibrium interaction enables as to predict changes in trade patterns and economic growth based on changes in world economic conditions. Other international integrated modeling efforts have emphasized energy-economy interactions. Such models typically generate scenarios which explore the impact on international trade and growth of changes in energy conditions. The latter include rising oil prices as well as increased oil depletion and energy conservation. Some of these models include energy as only a minor sector of the whole model while others consider energy as the driving sector of the model. Examples of these include the Hughes and Mesarovic (1978) world integrated model (WIM), the Herrera and Scolnik (1976) version of the Bariloche Model, and the Linneman (1976) MOIRA Model.

REFERENCES

Ball, J. (1973) Models of Project Link, North-Holland Publishing Co., Amsterdam.

Decision Focus Inc., 1979, "Modeler's Guide to Building Energy-Economy Models Using the Generalized Equilibrium Modeling System (GEMS) and the Model Data Management System (MDMS)," Report to the Energy Information Administration, Washington, DC.

DeGenring, P. C., and Jackson, J. R., 1980. "An Econometric-Engineering Analysis of the Impacts of the National Energy Act on U.S. Commercial Sector Energy Demand," Energy Modeling Studies and Conservation, ECE, ed., Pergamon Press for the United Nations, New York, pp. 303-18.

Energy Information Administration, 1979, Documentation of the Project Independence Evaluation System, Vols. I-IV, Department of Energy, Washington, DC.

Energy Information Administration, (EIA), 1982, "Review and Critique of the Department of Energy's New Integrated Forecast System," Presented at the Symposium on EIA Models, Washington, DC.

Exxon Corporation, Public Affairs Department, 1977, "World Energy Outlook,: New York, January.

Farina R., and Glover, F., "The Applications of Generalized Networks to Choice of Raw Materials for Fuels and Petrochemicals," In B. Lev (ed.) Energy Models and Studies, North-Holland Publishing Co., Amsterdam.

Groncki, P. J., and Marcuse, W., 1980, "The Brookhaven Integrated Energy/Economy Modeling System and Its Use In Conservation Policy Analysis," Energy Modeling Studies and Conservation. ECE. ed., Pergamon Press for the United Nations, New York, pp. 535-556.

Herrera, A. O., and Scolnik, H. D., 1976, "Catastrophe or New Society: A Latin American World Model," International Development Research Center, Ottawa.

Hogan, W. H., and Weyant, J. P., 1980, "Combined Energy Models," Advances in the economics of Energy and Resources, J. R. Moroney, ed., JAI Press., Greenwich, CT, pp. 117-150.

Hudson, E. A., and Jorgenson, D. W., 1974, "US Energy Policy and Economic Growth, 1975-2000," Bell J. Econ. Manage. Sci., Vol. 5, pp. 461-514.

Hughes, B. B. and Mesarovic, M. D., 1978, "Analysis of the WAES Scenario Using the World Integrated Model," Energy Policy, pp. 129-139.

Lev, B., Murphy, F. H., Bloom, J. A., and Gleit, A. S., (eds.), 1984, Analytical Techniques for Energy Planning, North-Holland Publishing Co., Amsterdam.

Lev, B., (ed.), 1983, Energy Models and Studies, North-Holland Publishing Co., Amsterdam.

Linneman, H., 1976, "MOIRA: A Model of International Relations in Agriculture - The Energy Sector," Working Paper, Institute for Economical Social Research, Free University, Amsterdam.

Quirk, J., Terasawa, K., and Whipple, D., 1982, Coal Models and Their Use in Government Planning, Praeger Publishers, New York.

Energy Vol. 15, No. 3/4, pp. 345–362, 1990
Printed in Great Britain

0360-5442/90 $3.00 + 0.00
Pergamon Press plc

3.4. SET OF MODELS

GALE BOYD[+], JOHN FOX, and DONALD HANSON[++]

Policy and Economic Analysis Group
Environmental Analysis and Information Systems Division
Argonne National Laboratory
Argonne, Illinois
U.S.A.

Abstract - The use of disparate and detailed engineering and economic models is often necessary for environmental policy analysis and forecasting. However, to conduct consistent forecasting and policy analysis over all economic sectors these disparate models must be coordinated into a consistent model set. One approach to such a modeling effort is illustrated by the NAPAP Integrated Model Set, a collection of engineering, emissions-forecasting and energy-market models that is driven by and interacts with other energy-market and economic models.

1. INTRODUCTION

To make forecasts and analyze policies that involve complex systems, like those involving interactions among energy, environmental, and economic factors, a set of highly specialized models may be desirable. The difficulty comes in integrating this set of models into a consistent framework where important interactions can be represented. This is a different type of modeling task than the development of a single model for a single purpose. In particular, the genesis of these specialized models may be different, as may be the modeling approach that each one takes. The task of integrating these disparate models into a consistent set, eliminating overlap, and establishing critical information flows and feedbacks is therefore an interesting project in itself.

As is clear from the variety of models represented in this volume, models may have an engineering, economic/statistical, or hybrid approach, depending on the focus of the forecasting and analysis they were developed to address. For example, the specific task of modeling and analyzing policies designed to control environmental emissions must be closely tied to the state of current and emerging technology. Therefore models that consider the costs and consequences of pollution controls tend to be highly specialized, with detailed engineering specifications.

[+] To whom all correspondence should be addressed

[++] The authors wish to acknowledge the support and assistance of the Assistant Secretary for Fossil Energy, Office of Planning and Environment, U.S. Department of Energy under contract W-31-109-Eng-38 and the guidance of the project manager, Edward C. Trexler. Greatly appreciated has been the excellent work on various parts of the Integrated Model Set (IMS) by Gordon Lurie, C. Yuen, Ron Fisher, Amit Bando, Brian Edwards, Cary Bloyd, Dan Miller, and Guy Pandola in the Energy Policy Section at Argonne. Comments by David G. Streets, Director of the Policy and Economic Analysis Group, and Marylynn Placet have also been helpful. The models of the utility sector (AUSM) and industrial sector (ICE, PROMPT, VOC) were developed by contractors to the U.S. Environmental Protection Agency, Larry Jones, manager. An earlier version of this paper was presented at the 27th Joint National meeting of CORS/ORSA/TIMS.

Nevertheless, these models require driver data (i.e., linkages) from energy market and economic models in order to run.

Such is the case for the set of models collected for use in the United States National Acid Precipitation Assessment Program (NAPAP).[1] These models are mostly engineering-process oriented or use other engineering notions to develop environmental emissions forecasts, control strategies, and cost estimates. In some cases the models may play a key role in the energy market (e.g., represent a specific demand sector). In one important case, the model that is the focus of critical environmental policy analysis addresses a significant part of the energy market (the electric generating sector).

The NAPAP Integrated Model Set (IMS) has been developed for use in the NAPAP assessment planned for 1990 (NAPAP 1989). This assessment must be credible and based on a consistent set of assumptions. This consistency is derived from an integrated approach, in which a common set of assumptions is used to drive the components of the IMS. The components of the IMS include economic and energy scenario driver modules, sectoral energy-use and emissions models, feedback loops, and integration connections.

Section 2 of this paper begins with a review of approaches that have been used for integrated long-term modeling. Section 3 is an overview of the configuration of the NAPAP IMS, including a description of this application and how it is planned to be run for the 1990 NAPAP assessment. Sections 4-7 provide a description of the four steps in the IMS: (1) scenario preprocessors, (2) annual recursive state loop, (3) emissions postprocessor models, and (4) aggregation, feedback, and report writing. The computer implementation of this system is summarized in Sec. 8. A brief summary is provided in Section 9. It is hoped that after the NAPAP 1990 assessment is finished the tools that have been developed for it will be useful for other assessment and policy issues like the growing debate over the production of greenhouse gases that contribute to global warming.

2. A REVIEW OF APPROACHES USED IN INTEGRATED NATIONAL LONG-TERM MODELING

Three approaches have been used to model national energy-market, engineering, and environmental systems: general equilibrium models, disequilibrium or simulation models, and programming or optimization models. This section is not a complete list or review of energy and emissions models; it simply describes these approaches and provide some examples.

General equilibrium models use a mathematically computable form to describe various sectors in the economy. Some formulation of the mathematical fixed-point theorem is implemented to find a set of prices in the economy that "clear the sector markets" such that every traded commodity has neither excess supply nor demand. Two examples of general equilibrium models are the TESOM/9DGEM and the PC-AEO models.

The TESOM/9DGEM model is a general equilibrium model of the U.S. economy with an explicit treatment of the energy sector. It consists of a set of submodels describing the household

[1] NAPAP is a 10-year multiagency effort and its organization includes an Interagency Science Committee, an Interagency Policy Committee, and seven task groups. Task Group 1, on Emissions and Controls, is chaired by the U.S. Department of Energy (DOE) and charged with three areas of responsibility: (1) Provide accurate estimates of emissions of precursors of acid deposition and historic emission trends. (2) Provide an integrated model set capable of projecting future emissions and costs under a variety of assumptions and policy alternatives. (3) Assess and help advance emission control technology.

sector, producer sector, investments, exports, and government expenditures. The producer submodel determines input purchases, which minimize average production costs for the prevailing set of input prices, while the household model determines the demand for goods and services and the supply of labor based on a model of consumer utility maximization. Private investment is determined by capital market conditions that reflect private domestic saving and government and foreign deficits. Government expenditures and exports are determined exogenously. The submodels are solved repeatedly until equilibrium is reached in input and output markets and until consumers and producers have optimized their behavior subject to their financial constraints.

The PC-AEO modeling system of DOE's Energy Information Administration (EIA) is a series of linked spreadsheets used to forecast national energy trends for various sectors and fuels in the United States. The system includes a macroeconomic model and models describing the residential, commercial, industrial, transportation, and electric utility sectors and the fuels petroleum, coal, oil, and natural gas. Each spreadsheet model can be operated independently, or the models can be integrated by exchanging price and quantity data repeatedly among them until energy balance and economic equilibrium are attained. Although it is not a true general equilibrium model since only the energy sector is represented and convergence is based on a heuristic approach, the PC-AEO model is in this class of modeling approaches.

The disequilibrium model approach allows for the possibility of non-clearing markets. This is usually attributed to structural inflexibilities in the energy or economic system. The disequilibrium approach may also be described as a dynamic simulation approach (i.e., a simulation of observed energy-market and economic behavior). The FOSSIL2 modeling system is a dynamic disequilibrium model of the U.S. energy system. It is based on a number of exogenous decision rules that determine the flow of resources, investments, and energy-consuming goods. The model uses difference and integral equations to describe the dynamic behavior of four major energy-consumption sectors and four major energy-production sectors, based on forecast inputs of key macroeconomic variables.

The programming or optimization models use some assumption about the optimizing behavior of the engineering and economic system to predict the future outcome of important decision variables. Two examples of programming models are the PILOT energy modeling system and the Market Allocation (MARKAL) Model. The PILOT system is a time-phased (dynamic) linear programming model that consists of several submodels. These submodels are input/output models that describe consumer behavior and production in seven nonenergy and five energy sectors. The model optimizes over planning periods for up to 100 years in five-year increments by maximizing a linear objective function. This objective function is usually the present value of personal consumption over the planning period. MARKAL is a multiperiod, linear programming energy supply model. The model contains a detailed description of existing and emerging energy supply technologies. It optimizes a set of preprogrammed objective functions subject to technological constraints and a set of exogenously specified demands. The model is most useful as a technological assessment tool and for analyzing tradeoffs such as emission levels and cleanup costs.

All three national modeling approaches use a common framework and solution approach to forecasting and policy analysis. The integration of a model set involves drawing from disparate, (often standalone) models to obtain scenarios as consistent as possible. An earlier effort at collecting several models under a common framework and set of assumptions was the IIASA model set. The IIASA model set was designed to operate on the global level (Basile et al. 1980). For this purpose, the world was divided into seven geopolitical regions, and individual scenarios were developed for each region. These were then aggregated to yield global scenarios (Hapele et al., 1981). The set is flexible enough to operate on the national level and was used for national projects.

The iterative procedure begins with initial assumptions about the state of economic growth. The aggregated macroeconomic model had been used for generating the gross national product

(GNP). The population projection was the input to the labor force calculation. The first formal model was the demand model MEDEE. (This methodology has been described by Chateau and Lapillone in their Chapter on the accounting and end-use models). In addition to macrofactors, MEDEE is driven by lifestyle parameters, technical efficiencies, etc., to produce projections of energy consumptions. These are fed into a model of the energy supply system and it general environmental impacts (MESSAGE). This is the optimization model that computes the least-expensive supply strategy that meets the specific consumption levels subject to constraints on resources or technology. The MESSAGE model is an extension of the initial IIASA energy-supply model developed by Hafele and Manne (1974). This methodology has been adopted in all intertemporal energy models (see detailed description by Kleemann and Wilde in their Chapter on intertemporal models). The linear programming algorithm that has been used in these type of models generate the marginal cost for energy forms. These prices can be fed into the model at several points to verify the preliminary assumptions. The energy-supply strategies (from the MESSAGE model) are fed into an economic IMPACT model that calculates the requirements for capital investment, labor, land, water, materials, equipment, and additional energy (for IMPACT methodology see Kononov in Chapter 3). In the initial assumption the IMPACT mode was chosen to close the loop.

Critical opinions about the IIASA set of models (Keepin and Wynne 1984) refer first to the way in which the models have been applied and not to the method as such. While recognizing the criticism of the model sets, one should still realize that in many cases these sets may be applied successfully. A very important advantage of this approach is that it forms a good basis for analyzing most of the important interactions.

3. OVERVIEW OF THE INTEGRATED MODEL SET

Summary Description

Figure 1 presents a simplified view of the structure of the IMS. A model run consists of four steps: Step I is a set of scenario preprocessors, Step II is the annual recursive state loop, Step III is the postprocessor sectoral emission models, and Step IV is aggregation, feedback, and report writing.

A model run begins in Step I when it uses data on a target GNP growth path, exogenous energy prices, and other variables from DOE's National Energy Policy Plan (NEPP). Emission constraints, if applicable, are also specified for an emissions-control-policy scenario. This GNP and some energy information from NEPP are used in the Data Resources, Inc. (DRI) annual macroeconomic model to provide a consistent macroeconomic simulation. The DRI macro model forecasts industrial output by sector, housing starts, interest rates and other variables used in the IMS. The macroeconomic simulation is disaggregated in the DRI regional model to determine the regional activity data required by the sectoral energy-demand modules and emissions modules (Bando, et al. 1988).

The Advanced Utility Simulation Model (AUSM) national loop creates a preliminary plan for the electric utility system. The national loop calculates a capacity expansion plan for each state (new units by technology type and fuel type), interstate transfers of electricity, and state emission caps. The national loop's Coal Supply Model (NCSM) determines mine-mouth prices and the price of the least-delivered-cost coal to each state by sulfur content. The Gas Supply Model (GSM) determines a wellhead price of natural gas, given gas demand. Final-user gas prices are constructed from wellhead prices by adding transmission and distribution costs to each region. The exogenous world oil price is disaggregated into fuel types and regionalized based on an analysis of historic patterns.

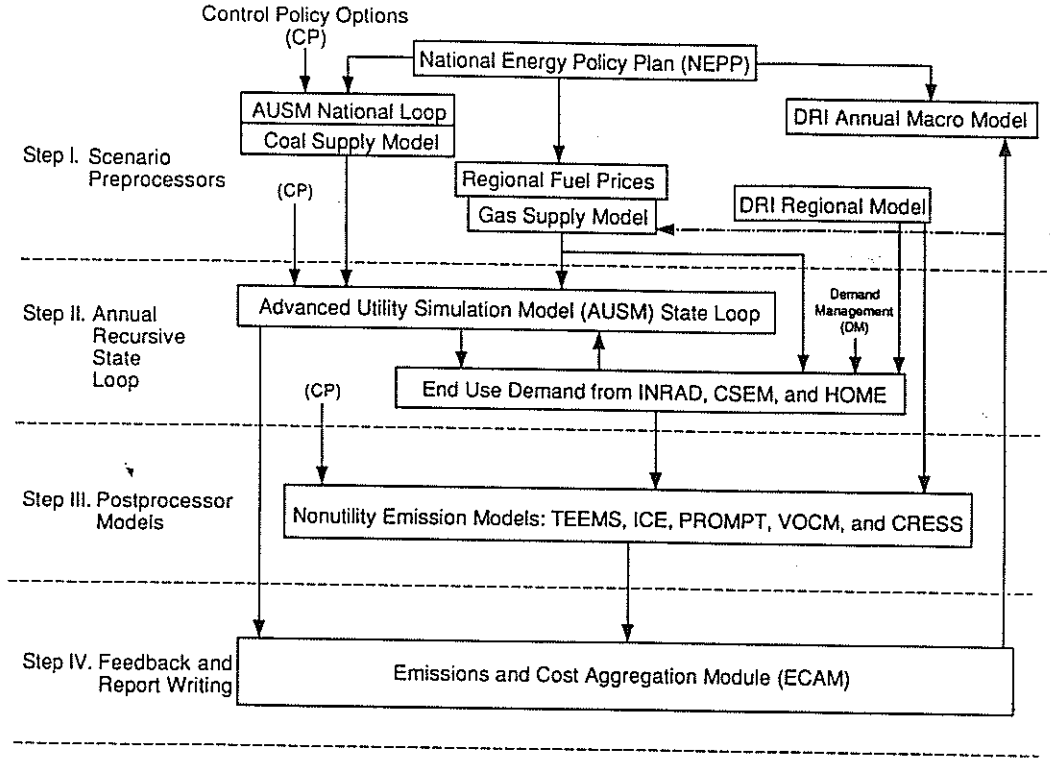


Figure 1. Overview of the Integrated Model Set

In Step II, this information is used by the AUSM state loop to annually model the electric utility industry in each state by determining utility fuel use, electricity output, electric rates, and utility emissions. Electric rates and activity data for each year are passed to the Industrial Regionalization and Disaggregation (INRAD) model, the Commercial Sector Energy Model by State (CSEMS), and the Household Model of Energy by State (HOMES) to determine the demand for electricity and fuels by these sectors. The annual electricity demand data are sent to the AUSM state loop to form the annual recursive loop. Electric rates are calculated by AUSM using a revenue requirements approach.

When Step II of a simulation has been completed for all states, the energy demands from Step II and other activity and price data from Step I are sent to the postprocessor models in Step III to calculate nonutility emissions. The Transportation Energy and Emissions Modeling System (TEEMS) is used to determine energy use and emissions from the transportation sector. The Industrial Combustion Emissions (ICE) model is used to determine fuel choice and emissions from industrial boilers. The Process Model Projection Technique (PROMPT) accounts for emissions from process energy use. The Volatile Organic Compounds (VOC) model calculates industrial emissions of nonmethane hydrocarbons. Commercial and residential emissions are calculated by the Commercial and Residential Emissions Simulation System (CRESS).

The sectoral data from Steps II and III are aggregated and summarized by the Emissions and Cost Aggregation Module (ECAM) in Step IV. In this step, reports on activity levels, emissions, costs, energy and other resource use are prepared. This information can be returned to the preprocessor models in Step I to provide recursions if necessary to attain a consistent forecast.

NEPP and the NAPAP Reference Scenario

Forecasts of energy prices and economic activity forecasts from NEPP provide many of the required scenario inputs for the NAPAP Reference Scenario (NRS). Sources of the input assumptions for the NRS are listed in Table 1. The provision of driver data on supplemental electricity supply at the state level is described in Hanson, et al., (1988b) for use in AUSM testing. Many of the NRS electric utility supply factors are determined by a joint DOE/EPA agreement. Sensitivity cases address alternative economic growth rates, power plant retirement ages, nuclear growth, penetration of clean coal technologies, vehicle miles traveled (VMT), world oil prices, and domestic gas resources.

Table 1
Sources of Input Assumptions for NAPAP Reference Scenario (tentative)

Category	Source
GNP and Demographics	
GNP	NEPP-7 GNP path until 2010 GNP extension until 2030 provided by Chris Caton, DRI as embodied in our 2030 version of the DRI Annual macro model.
Population and demographics	U.S. Bureau of the Census middle population growth series as used by NEPP and DRI
Energy Prices and Demands	
World Oil Price	NEPP-7 and long-term extension
Ave. U.S. oil product prices	NEPP-7 (extended to 2030 by regression analysis with world oil price as independent variable)
Nuclear fuel prices	DOE/EIA
Fuel and electricity demand	Reference Case Calibrated to NEPP-7 for Analysis
Nonfossil electricity supply	
Net electricity imports	NEPP-7 and NERC reports
Industrial cogeneration	FOSSIL2 model
Nuclear capacity and generation	NEPP-7 and EIA
Renewable electricity generation	NEPP-7, DRI regional shares
Air Pollution Regulations	
NSPS, proposed ozone standards	EPA
Electric Utility Factors[2]	Joint DOE/EPA agreement

[2] These include assumptions regarding repowering, life extension, retirement age, performance characteristics, minimum capacity factors, penetration of clean coal technologies, and independent power producers

Proposed emission control policies include various ways to reduce emissions. They include (1) fuel switching vs. forced scrubbing, (2) adoption of clean coal technologies, (3) alternatives that control different combinations of sources and sectors, (4) mandated retirements in noncomplying units, (5) emission taxes or tradeable permits, (6) alternative financing, (7) strategies relying on increased natural gas use either by co-firing or seasonal fuel switching, (8) policies implemented in two or more phases, and (9) scenarios that give different weights to reductions in sulfur dioxide, nitrogen oxides, and volatile organic compounds (SO_2, NO_x and VOC).

Determining which sources and sectors should be controlled involves many options in itself. Controls could affect a wide range of sources: utilities, transportation sector, industrial boilers, industrial processes, solvent use, and commercial fuel use. An advantage of the economic-incentive approaches, such as an emissions tax, is that they may be cost effective over a wide range of sources. Controls can also be targeted to those sources deemed to be causing the greatest damage. (Streets et al., 1984).

It is intended that the IMS be capable of simulating many of the policies listed above. Some of these policies, such as intrastate or interstate emission trading, will require some postprocessing of model runs to calculate efficient, least-cost trades among emission sources (e.g., see Streets 1983).

4. SCENARIO PREPROCESSORS

The first step involves defining the scenarios. Forecasts of key macroeconomic and energy variables are provided from the NEPP simulations. Corresponding economic and demographic variables forecast from the DRI macro and regional models. The AUSM national loop is a linear program that is run as part of Step I to provide an overall plan for the electric utility sector. Key energy prices are also calculated in Step I for the coal market and the gas market.

DRI Annual Macroeconomic Model

The DRI annual macroeconomic model of the U.S. economy is being used to simulate the details of the economic scenario. Being an annual model, it is appropriately suited for the long-term projections (up to the year 2030) required by NAPAP. It also provides sufficient industry-specific detail to conform to the input requirements of the remaining components of the IMS.

NEPP specifies information on the growth in GNP, world oil prices, and other items related to the energy sector of the economy that are used as inputs to the DRI macroeconomic model. The macroeconomic model is then used to generate a set of macroeconomic and industrial forecasts that are consistent with the economic growth provided by the NEPP forecast. The main variable that DRI uses to adjust its GNP forecast is the outlook for the labor force participation rate. The money supply is also adjusted in the macro simulation to hold the nominal short-term interest rate constant.

Since the other modules of the IMS require regional and industry-specific drivers, the output from the DRI macroeconomic model is fed into the DRI regional model, which will be discussed in the following section.

DRI Regional Information Services Model

The DRI Regional Information Service (RIS) model provides economic forecasts and historical statistics for 20 industries in nine regions and in 50 states and the District of Columbia. More than a decade of historical data and up to 25 years of consistent forecast data are available. For the IMS, time-series data for several regional variables were used. DRI/RIS

provided an output file from its January 1989 simulation containing the historic and projected values of the variables in its multisector regional econometric model.

The variables from the January 1989 regional simulation have to be scaled to the forecasts from the DRI annual macroeconomic model. The regional data contain forecasts on two types of variables: quantity variables and index number variables. The scaling algorithm for the regional projections of the index number variables required adding the difference between the growth rates from the current macro model run and the January 1989 run of the DRI macro model to the growth rate of the corresponding variable from the regional file. For quantity variables, the regional variables were scaled to match totals provided by the DRI macro model run. Finally, the scaled variables were extended to 2030 using the growth rates obtained from the macro model. Figure 2 shows a representation of this scaling methodology.

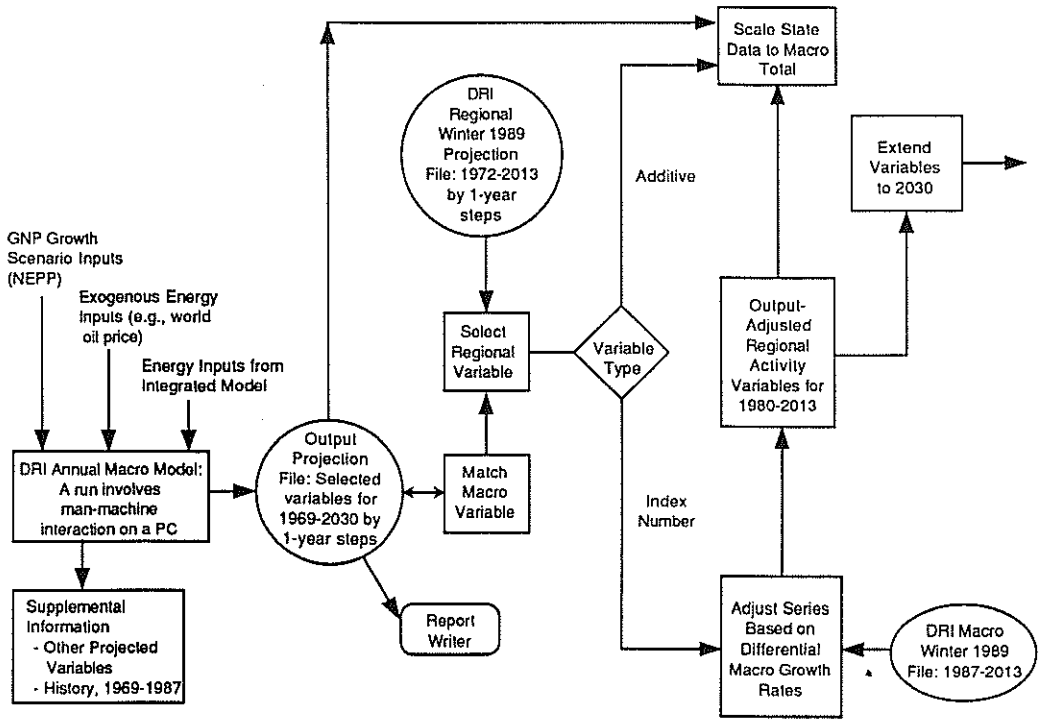


Figure 2. Scaling Methodology for Regional Variables

AUSM National Loop

The AUSM national loop is an linear-programming-based simulation model that develops a general utility strategic operating plan that is used by the AUSM state loop. The information passed to the state loop consists of projected capacity additions, interstate power transfers, and future coal prices.

The national loop is composed of two basic modules: the multiperiod, multistate (MPMS) programming module and the national coal supply module (NCSM). A general flow diagram of

the national loop showing its relationship to the state module is given in Figure 3. The MPMS module is a large linear program that performs a least-cost optimization of electricity production subject to emission and demand constraints. The optimization is carried out on a regional basis over multiple states, time periods, and technologies. In the base case, there are 13 regions, 12 time periods, and 20 technology categories. No parameters are passed between regions. Thus, a single MPMS execution would provide information on capacity planning and interstate power transfers for over the complete modeling time horizon for all states within the specified region. Each region is solved in turn as a separate problem formulation. After runs for all regions have been executed, coal usage information from all the regions is aggregated to provide the input to the NCSM.

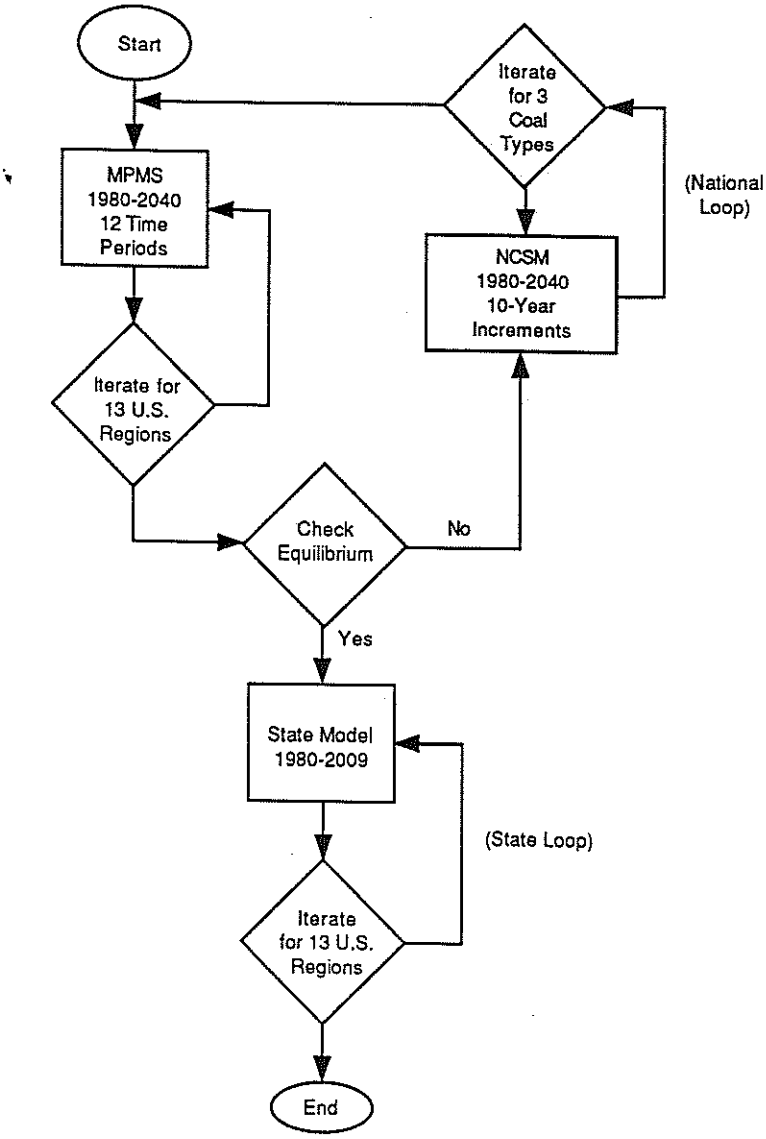


Figure 3. General AUSM Model Structure

The NCSM consists of a large coal-supply model and an equilibrium testing model. The coal-supply model estimates coal prices on a national basis over a 1980 to 2070 time frame in 10-year increments. Prices are given for 35 supply regions, 48 demand regions, 11 coal sulfur categories, and 3 coal heating values. The coal model is executed three times, once for each coal heating value (bituminous, subbituminous, and lignite). An equilibrium model compares the coal demand before and after an MPMS execution; if equilibrium conditions are not met, the coal file is updated to reflect the new demand, and another execution of MPMS takes place. At equilibrium, the resulting coal file becomes the input to the AUSM state loop. The formulation of the NCSM is discussed in more detail in the next section.

National Coal Supply Model

The NCSM is run before the detailed AUSM state model simulation. The coal supply model simulates costs of coal production, washing, and transportation and computes delivered least-cost coal as a function of time.

The existence of both near-term econometric and long-term model mine projections allows AUSM to deal realistically with near-term conditions as well as factors that determine coal prices over the long term. The AUSM capability for short-term coal-supply projection is based on an extensive, detailed data base of information on current coal production and prices and current coal movements. This short-term capability is linked to the standard long-term mine-costing approach so that the overall coal supply module more accurately reflects "inertia" built into the current coal supply and transportation system. The mine-costing algorithms used in the long-term show the effects of escalation in major components (e.g., wages) as well as the effects of cost increases due to severance taxes or other policies.

A major improvement over existing coal supply models is the AUSM coal supply module's incorporation of a detailed coal-cleaning model. The general objective of coal-cleaning processes is to improve the value of the coal produced by reducing the presence of impurities such as ash, sulfur, and moisture. The reductions are obtainable only at a cost, which comprise both the cost of constructing and operating a processing plant and the price of combustible material that is discarded with the refuse from the preparation process.

The final component in the NCSM is a set of coal-transportation costs linking each of the supply regions to the 48 demand regions. These origin-destination costs are computed for rail-water and rail-only routings using the U.S. Department of Agriculture's Transportation Analysis System, probably the most detailed representation of transportation networks available. Effects of changes in the system (e.g., escalation of various cost components) on future coal movements and prices can be analyzed.

The Gas Supply Model

The purpose of the Gas Supply Model (GSM)is to integrate the U.S. domestic natural gas market into the NAPAP Integrated Model Set. This model determines an endogenous wellhead price for natural gas.

The GSM is an adaptation of EIA's Oil and Gas Spreadsheet (OGS) model, which is a submodel of the PC-AEO spreadsheet model used to prepare the Annual Energy Outlook 1989 with Projections to 2000 (DOE 1989). The OGS model is a supply and demand model for natural gas that reads in world oil prices and natural-gas demands from other spreadsheets in the PC-AEO system. Then the model calculates domestic crude-oil and natural-gas production, imports, and the equilibrium wellhead price of natural gas.

The spreadsheet model has been integrated into the model set being recoded into Fortran 77 and by inputting natural-gas demand from the other modules as a substitute for the OGS demand functions. Thus in the integrated system, natural-gas demand is the sum of electric-

utility demand from AUSM, industrial demand from the industrial sector models, commercial demand from CSEM, residential demand from HOME, and gas-industry demand from within the model.

The supply of natural gas in the GSM is calculated using the same three-step process and the same econometric functions as those used in the OGS. First, the number of new exploratory gas and oil wells drilled is estimated as function of the world oil price and as a function of an initial assumed gas price. Second, the level of oil and gas reserves is revised based on the proportion of successful oil and gas wells. Third, domestic gas production is calculated as the sum of gas produced from base reserves, new reserves, reserve revisions, and gas associated with oil production. Total gas supply is determined by adding net imports and net withdrawals from storage.

To determine the equilibrium price, the quantity of gas supplied is compared with the quantity of gas demanded. If the quantities are within a tolerable range of 1%, the price remains at the initial assumed price. If the quantities are outside the tolerable range, the price is changed using a sequential interval bisection technique until the tolerable range is reached. The wellhead price of natural gas determined within the GSM will be used in the price module to determine natural-gas prices for each state and for each final-user class by using transmission and distribution costs forecast by the Gas Research Institute (Holtberg et al., 1988).

Other Fuel Prices

Prices for fuels other than coal or natural gas are determined exogenously. Historical state prices of all fuels used in the residential, commercial, industrial, transportation, and utility sectors were obtained from EIA's State Energy Price and Expenditure Report 1986 (SEPER) (DOE 1986). Historical prices for some electric utility fuels for 1987 were obtained from EIA's Cost and Quality of Fuels for Electric Utility Plants 1987 (DOE 1987). Projected national average sector prices for petroleum prices were obtained from the Office of Technology Policy's National Energy Policy Plan Projections. Draft 2 (DOE 1987). Projected national average nuclear fuel costs were obtained from an input to EIA's Annual Energy Outlook for 1989 (DOE 1989).

These projected national average sector prices were converted into projected state sector prices using Eq. 1:

$$P_{i,j,s} = P_{i,j} + \mu_{i,j,s} \times W \tag{1}$$

where:

$P_{i,j,s}$ = projected state price for fuel i, sector j, and state s

$P_{i,j}$ = projected national price for fuel i in sector j

$\mu_{i,j,s}$ = arithmetic mean difference between the state and national price for fuel i, sector j, and state s between 1980 and 1986

W = weighing factor

The weighing factor is 1.0 in 1987 and declines to 0.7 in 2010, thus assuming a regression of these state prices toward the national mean.

5. ANNUAL RECURSIVE STATE MODEL

The annual recursive state model, described in this section, is based on the AUSM state model and forms the core of the IMS. As diagrammed in Figure 4, the price and activity data

from the preprocessors in Step 1 are provided to the annual recursive loop in Step 2. Step 2 begins by initializing the national data by reading national files on global scenario parameters, demand data, planning data, finance data, and fuel data. Then the AUSM state loop begins; it runs sequentially for each of the contiguous 48 states with annual recursion until the end of the selected period.

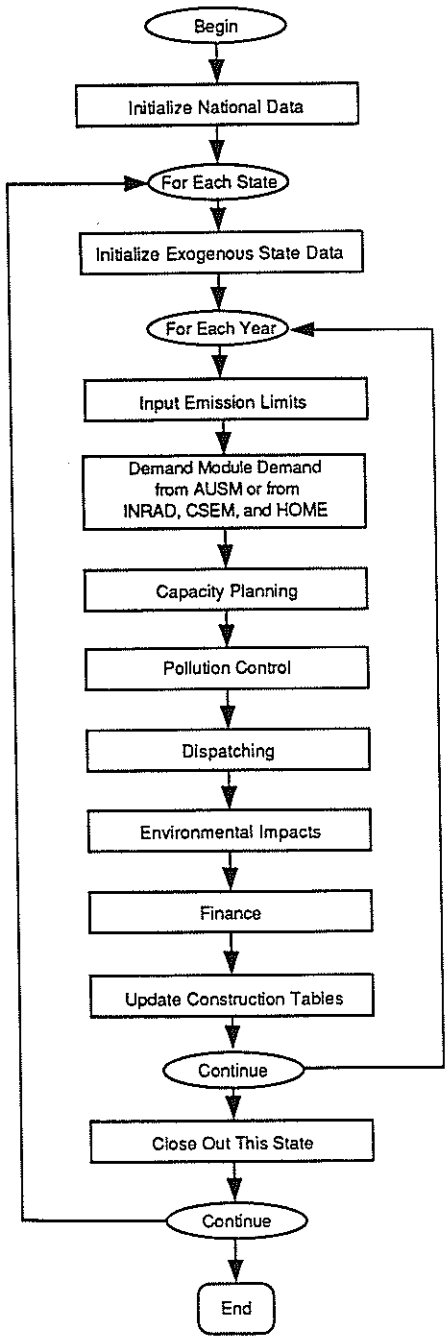


Figure 4. Annual Recursive LOOP-Based on AUSM State Model

AUSM State Model Annual Recursion

The first step within the AUSM state loop is to initialize the state data. This step includes reading state files on fuel cost, fuel characteristics, projected demand, planning data, and scenario parameters including emission limits.

The next step is to determine energy demand. The demand module begins by determining rate schedules for electricity. These rate schedules are based on historical data in the first year and are updated using information from the finance module in subsequent years. The AUSM demand module allows the user to choose among three methods of determining demand: (1) endogenously using a system of linear logit demand functions, (2) exogenously, or (3) exogenously with annual recursion. The last option calculates price induced changes about an exogenous baseline. A fourth option has been added to AUSM by Argonne to complete the IMS. This option specifies the historical demand over the period 1980-1986 exogenously and calculated demand endogenously thereafter based on the energy driver for the NAPAP sector models. These energy drivers are INRAD, CSEMS, and HOMES for the industrial, commercial and residential sectors, respectively. These models provide not only electricity demand to AUSM but also fossil fuel demand to the sectoral emissions modules ICE, PROMPT, and CRESS (to be discussed in Sec. 6). Hence, overall consistency in the energy forecast is maintained for the purpose of computing emissions.

Electricity demand goes to the capacity-planning module to determine a schedule for the construction of electric-utility power plants. Then the pollution-control module determines the optimal pollution-control strategies at existing plants and those under construction. In addition, this module supervises the construction of pollution-control equipment at existing plants, assigns capacity penalties and compliance costs to these plants, and upgrades their fuel and pollution-control data.

The dispatching module allocates the demand for electric energy among the generating plants available to meet the demand. This module is usually used to find the least-cost allocation of generation that meets emissions limits. The module can be used, however, to determine the allocation that minimizes emissions of a particular pollutant while complying with limits on other emissions and total cost. Then the environmental impacts module calculates the emissions from fossil-fuel plants and the variable cost of pollution control in each year of the simulation.

The finance module has two main functions. First it calculates the electric-utility revenue requirements based on the cost of existing plants plus the cost associated with new capacity or emissions controls. These revenue requirements are used in the demand module to determine electric rates. In addition, the finance module prepares reports on representative state-level utility financial statements.

At the end of the forecast period, the state file is closed out. Summary statistics are printed, and the program loops to the next state until all 48 state runs are completed.

Commercial and Residential Sector Energy Demand

The objective of modeling the commercial and residential sector demands is to develop a module capable of preparing energy-use projections by major fuel types for the two sectors on a state-by-state basis. The models should be capable of simulating alternative baseline scenarios and policy analyses. The models selected to represent these sectors are two EIA models: CSEM and HOME. These models need to be consistent with the AUSM state loop and with the input requirements of the emissions module that is described in Sec. 6.5. As part of the IMS, these models provide sectoral electricity-demand projections for the AUSM state loop and fossil-energy-use projections for the CRESS emissions module. Revisions had to be made to these models (1) to increase the time horizon to 2030 as required by the IMS, (2) convert them to operate for individual states rather than at the regional level, and (3) enhance their fuel-

substitution specification. To project fuel consumption for each end-use, the models follow three basic steps:

(1) update the basic sectoral activity variable, floorspace in the commercial sector, and housing stock in the residential sector;

(2) adjust the share-of-activity variable using each fuel type; and

(3) apply a utilization rate (fuel use per unit of activity).

Industrial Demands

Industrial energy demand plays a dual role in IMS. Industrial electricity demand provides a portion of the driver data for AUSM. Industrial fossil-fuel demand for boilers and process heaters is required as input for ICE and PROMPT. The extent to which these demands are interconnected is modeled by the INRAD module. INRAD combines and extends the previous National Adjustment and Disaggregation (NAD) module (Boyd, et al. 1986) and the industrial-sector regionalization, the Argonne Regionalization Activity Module (ARAM), used to provide test data for the development of the individual emissions models (Hanson et al. 1988a).

INRAD is implemented using two equations based on a constant return to scale (CRTS) generalized Leontief factor-demand system (Deiwert and Wales 1987). The two factors in INRAD are fossil fuel and electricity. The CRTS generalized Leontief approach yields equations that predict the factor intensity; i.e., the ratio of the factor input to a measure of output. The variables used to predict the factor intensities are the square root of the ratio of the factor prices, capacity utilization, and technical change variables. The prices included in INRAD are for fossil fuel and electricity, and price indices for capital, labor, and materials are included as well.

The equations are estimated from national energy accounts data for the period 1958-1985. A separate equation is estimated for eight energy-intensive industry groups; only the energy-intensive, basic-materials portion of the two-digit sector is used. A single equation is estimated for all other nonenergy-intensive industries.

The national-level equations are implemented at the state level by assuming common price elasticities in each industry, across all states. To preserve the elasticities, a multiplicative benchmark of the national-level equation must be computed based on state-level energy intensities, which depend on state-level energy prices. State-level, industry-specific energy prices are input into the INRAD energy intensity equations to compute a state-level energy intensity. Since the goal is to forecast energy use, the benchmarking is done by taking the forecast growth rate in industry output for a state and the growth rate for the corresponding predicted energy intensity and applying these two growth rates to a base-year state-level energy use. The use of growth rates is equivalent to computing a multiplicative benchmark for each state.

INRAD obtains the forecast of state-level, industry-specific prices from other modules in the IMS. When running in the annual recursive mode, INRAD takes the average industrial electric prices from AUSM. This procedure is done to account for industry-specific and state-specific base-year differentials and forecast trends in prices, respectively. The industry- and state-specific fossil-fuel price index is constructed by taking 1980 industry-specific expenditure weights for residual oil, distillate oil, natural gas, and coal consumption and weighting the corresponding state-level price forecasts from the IMS price module.

6. POSTPROCESSOR SECTORAL MODELS

These model activities and emissions for the residential, commercial, industrial, and transportation sectors using output from the preprocessors and the annual recursive loop. The industrial sector is by far the most diverse sector modeled by the NAPAP IMS. The variety of industrial activities and the ways in which this sector affects the energy drivers and emissions are significant. For this reason, the industrial sector is broken down into several component models: the ICE model, VOC model, and PROMPT. These three components of the IMS were originally developed as standalone emissions-forecasting models. They all calculate emissions and control costs for regulatory scenarios.

The ICE model is a forecasting model for boiler fuel choice and emissions. The basic drivers in the model are regional boiler-fuel and a fossil-fuel price forecasts from INRAD and the price modules, respectively. The industrial VOC model transforms inputs of growth rates for 126 categories of regional industrial activity into levels of VOC emissions. These activities are largely derived from the DRI macro and regional forecasts, with some appropriate adjustments. PROMPT is an emissions accounting model that transforms a forecast of industrial process-energy use into the corresponding SO_2 and NO_x emissions. The fuel-use inputs are based on the ISTUM II model of industrial process-energy use.

The commercial and residential emissions module CRESS is designed to project emissions estimates for SO_2, NO_x, and VOCs from 1985 to 2030. The module uses fuel-use projections from CSEMS and HOMES, economic-activity projections from DRI, the NAPAP emissions data base, and assumptions about engineering controls and emission rates to forecast these pollutants.

TEEMS is a model system that forecasts the transportation sector activity and the impacts of these activities on emissions of SO_2, NO_x, and VOCs. Price and activity data from the IMS are inputs to the TEEMS submodels and emissions and fuel-use data are output. TEEMS incorporates submodels to forecast (see Saricks, 1985, and Vyas and Saricks, 1986):

- All local personal travel, both that occurring within standard metropolitan statistical areas (SMSAs) as defined in 1980 and other (nonSMSA) travel;
- Intercity personal travel by ground and air (business and nonbusiness);
- Commercial and rental automobile and light-truck travel;
- Interurban (including port-to-port) goods movement (coupled with intraurban distribution);
- Other aviation (including general, military, and international travel); and
- The emissions from each of the above activities.

The TEEMS emissions module is MOBILE4. Considerable work has gone into the preparation of speed, temperature, and other inputs to MOBILE4.

7. FEEDBACK

In the integrated model, as in any free-market system, prices play an important role in resource allocation. As shown in Figure 5, the price of coal is determined endogenously by AUSM's NCSM, and the wellhead price of natural gas is determined endogenously in the GSM. The prices of petroleum products such as gasoline, residual oil, and distillate oil are determined exogenously by the world oil price. The price of nuclear fuel is also determined exogenously.

These prices, together with activity variables, technology factors, other input prices, and emission regulations, determine the supply of electricity and the demand for utility fuels in the AUSM state model. The demand for electricity and other fuels by the industrial, commercial, and residential sectors is determined by INRAD, ICE, CSEM, and HOME, using these fuel prices and consumer-activity variables such as consumer income and the output level of final products.

Equilibrium is attained in the coal, natural-gas, and electricity markets as the prices of these fuels are determined endogenously and will adjust within the model by recursion until the quantity supplied equals the quantity demanded in each market. Since the prices of petroleum products and nuclear fuel are determined exogenously in the model, these markets may not attain equilibrium. The IMS seeks a partial-equilibrium solution in this sense.

Because the prices of coal, natural gas, and electricity are determined endogenously, the model can be used to determine the impact on these industries of changes in emissions or other policies in terms of both the prices and the consumption of these fuels. Since the prices of petroleum products and nuclear fuels are determined exogenously, the impact on the consumption of petroleum and nuclear fuel would be determined within the model, but the effect on prices would not.

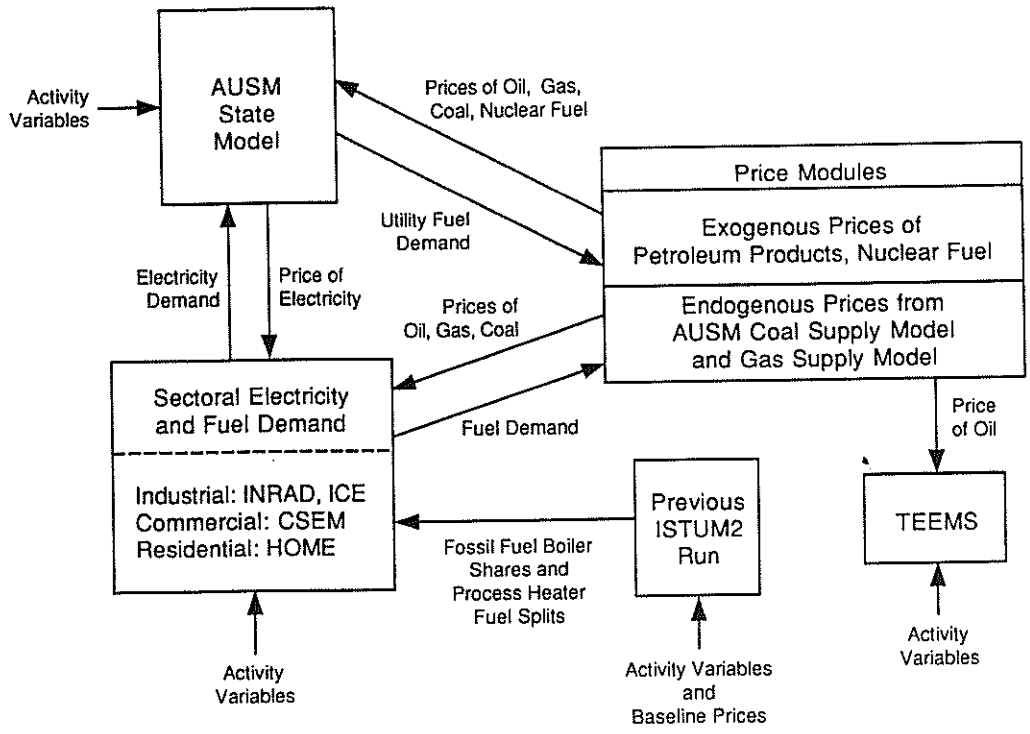


Figure 5. Energy-market Balance and the Role of Prices

The national impacts of any specific scenario may often mask the regional disparities. Hence the regional impacts are considered crucial and are modeled explicitly through the DRI/RIS model, which, in turn, depends on the values of selected output variables from the DRI annual macroeconomic model.

8. OVERVIEW OF COMPUTER IMPLEMENTATION

The implementation goal of the IMS is to provide a portable system that can be used and run by members of the emissions modeling community. The IMS has been assembled from several independently developed activity and emissions models. Computer development has been proceeding in three primary phases. Phase 1 encompassed the initial model development that took place under several language and operating-system environments over the past 10 years. In Phase 2, the present state, most of these models have been ported to the UNIX operating system. The models were then modified to provide for either simultaneous or iterative interaction between models during simulation runs. At present, the integrated system operates primarily on a SUN UNIX workstation. In Phase 3, those remaining models not operating under the UNIX operating system will be ported to it, allowing the complete system to run on a single, independent UNIX machine.

9. SUMMARY

This paper describes the implementation and development of an integrated set of models through the example of the NAPAP IMS. This model set has been constructed to provide a consistent assessment of the impacts of control policies for the precursors of acid deposition. It is intended to replace myopic sectoral studies with a model system that will determine the effects of control policies throughout the economy.

REFERENCES

Bando, A., D.A. Hanson, and D.G. Streets (1988). "The Role of Regional Activity Forecasts in Projecting Emission Trends of Acid Deposition Precursors", Argonne National Laboratory, presented at the Second Joint Conference on Integrated Data for Decisionmaking, National Governor's Association, Washington, DC.

Basile, P., et al. (1980), "The IIASA Set of Energy Models: It's Design and Application". International Institute for Applied Systems Analysis, Loxenburg Austria.

Boyd, G.A., et al. (1986). "Disaggregation of Industrial Fossil Fuel Use in the 1985 National Energy Policy Plan: Methodology and Results," Argonne National Laboratory Report ANL/EES-TM-303.

Cambridge Systematics, Inc. (1988). "Evaluation of the Transportation Energy and Emissions Modeling System (TEEMS) Final Report," Berkeley, CA.

Curkendall, S.M. (1988). "Review of the Economic Driver Models," prepared by National Economic Research Assoc., Inc., Cambridge, MA.

Diewert, E. and T. Wales (1987). "Flexible Functional Forms and Global Curvature Conditions", Econometrica, Vol. 55, No. 1, pp 43-68.

Hapele, W., et. al. (1981). "Energy in an Finite World, Vol. 2: Global Systems Analysis," report by the Energy Systems Program of the International Institute for Applied Systems Analysis, (Bollinger), Cambridge, MA.

Hanson, D.A., and K.H. Alfsen (1986). "Why an SO_2 Emission Tax is an Unpopular Policy Instrument: Simulation Results from a General Equilibrium Model of the Norwegian Economy," Proc. 7th Annual North American Conference of the International Association of Energy Economists, Boston, MA.

Hanson, D.A., D.W. South, and W.H. Oakland (1988a). "A Regionalization Methodology for Sector Model Input Data: Derivation and Applications," Argonne National Laboratory Report ANL/EES-TM-301.

Hanson, D.A., et al. (1988b). "Advanced Utility Simulation Model (AUSM): State-Level Projections of Required Electrical Generation and Sources of Supply," Argonne National Laboratory Report ANL/EES-TM-357.

Hogan, W., and A. Manne (1977). "Energy-Economic Interactions: The Fable of the Elephant and the Rabbit" EMF working Paper 1.3, Energy Modeling Forum.

Holtberg, P.D., et al. (1989), "Baseline Projection Data Book, 1988 GRI Baseline Projection of U.S. Energy Supply and Demand to 2010", Gas Research Institute. Chicago.

Keepin, B., and B. Wynne (1984). "The Technical Analysis of the IIASA Energy Sciences", Nature, Vol. 312 20/27/, 691-693

National Acid Precipitation Assessment Program (NAPAP), (1989). "Plan and Schedule for NAPAP Assessment Reports 1989-1990, State of Science, State of Technology, Integrated Assessment", Washington, D.C.

Saricks, C.L. (1985). "The Transportation Energy and Emissions Modeling System (TEEMS): Selection Process, Structure, and Capabilities," Argonne National Laboratory Report ANL/EES-TM-295.

South, D.W., M.J. Bragen and C.M. Macal (1985). "Industrial Combustion Emissions (ICE) Model: Regionalized Projections of Demand for Purchased Industrial Boiler Fuel," Argonne National Laboratory Report ANL/EES-TM-302.

South, D.W., et al. (1988). "Industrial Volatile Organic Compounds (VOC) Model: Regionalized Projections of Uncontrolled VOC Emissions by Source Category," Argonne National Laboratory Report ANL/EES-TM-306.

Streets, D.G., et al. (1983). "A Regional, New-Source Bubble Policy for Sulfur Dioxide Emissions in the Eastern United States," Argonne National Laboratory Report ANL/EES-TM-239.

Streets, D.G., D.A. Hanson, and L.D. Carter (1984). "Controlling Acidic Deposition: Targeted Strategies for Control of Acidic Deposition", Journal of Air Pollution Control Assoc. 34 (12)

Vyas, A.D., and C.L. Saricks (1986). "Transportation Energy and Emissions Modeling System (TEEMS): Configuration for Forecasting Transportation-Source Emissions for the 1985 Test Runs," Argonne National Laboratory Report ANL/EES-TM-321.

U.S. Department of Energy, (1986) "National Energy Policy Plan Projections - Draft 2", Office of Technology Policy.

U.S. Department of Energy (1987). "Cost and Quality of Fuels for Electric Utility Plants," Energy Information Administration Report DOE/EIA-0191(87).

U.S. Department of Energy (1989). Annual Energy Outlook 1989 With Projections to 2000, Energy Information Administration Report DOE/EIA-0383(89).

U.S. Department of Energy (1986). State Energy Price and Expenditure Report 1986, , Energy Information Administration Report DOE/EIA-0376(86).

Energy Vol. 15, No. 3/4, pp. 363–378, 1990
Printed in Great Britain

0360-5442/90 $3.00 + 0.00
Pergamon Press plc

3.5. SIMULATION MODELS

NAZLI CHOUCRI and CHRISTOPHER HEYE

Department of Political Science
Massachusetts Institute of Technology
Cambridge, Massachusetts
U.S.A.

Abstract - This paper summarizes some basic assumptions and procedures for simulation approaches. A brief discussion of system dynamics is presented. A description of the International Petroleum Exchange Model (IPE) developed at MIT is presented. As a simulation model of the world oil market, IPE shows the interactions, adjustments, and behavior of major actors and agents, and of the underlying supply, demand, and price relationships. Illustrations of model results are presented along with comparison of alternative simulation scenarios.

1. INTRODUCTION

A wide range of simulation approaches are available for modeling markets, transactions, and interactions. To be effective for policy purposes, simulation models require specification of criteria to differentiate among options and select among alternatives. Two of the most widely used are system dynamics simulations and econometric-based simulations. This paper addresses some broad features of system dynamics, and then contrasts these with econometrics simulations. We then present the characteristic features and selective behavior of the International Petroleum Exchange Model (IPE), a system dynamics simulation model of the world oil market determining price endogenously.

The objectives of simulation, in general, are to understand the structure of a system, the processes generating behavior, and the outcomes that result from these processes. Policy simulations have the added objective of serving as a tool for purposive or instrumental analysis, i.e., where decisions have to be made and choices among options identified.

2. SYSTEM DYNAMICS SIMULATION

System dynamics is a theory of system structure and consists of a set of tools for identifying, depicting, and analyzing multiloop, nonlinear feedback relationships. Its theoretical foundations are in control theory and systems engineering. The objectives of a system dynamics simulation are as follows:

1. To understand the structure of the system under consideration;

2. To clarify prevailing theories regarding the system in question and provide the basis for the comparison of alternative theories;

3. To identify and resolve differences in assumptions, perspectives, and views regarding the structure or processes under consideration;

4. To isolate sensitive parameters in a model whose modification would affect overall system behavior; and

5. To identify policies that might propel a system toward the desired objectives, or assist in avoiding negative outcomes.

Given the complexities of markets, the salience of nonlinearities, and the difficulties of identifying, measuring, and predicting the future behavior of key variables, system dynamics as a methodology has important potential applications. Many of the problems confronting policy-making involve a host of factors over which actors do not have direct control and whose interactions they do not fully understand. The system dynamics paradigm is particularly well suited for the organization of information, theories, and expectations regarding policy problems, and the simulation algorithms allow for ready investigation of the consequences of alternative policies. Thus, a decision-maker can define the boundaries of his problem, specify in even loose fashion the relationships among key variables, and articulate his expectations regarding their future behavior and observe the consequences of alternative policies.

Assumptions

As a simulation approach to energy modeling, system dynamics is based on the following assumptions:

First, cause and effect are often separate in time and space, suggesting delays in information gathering, policy implementation, and policy impacts.

Second, most key relationships are nonlinear, discontinuous, or have no adequate linear approximation.

Third, policy analyses are at best ones of satisficing, rather than optimizing (as is more conventionally the case with other formal modeling approaches).

Fourth, feedback structures shape behavioral tendencies of a system, depending on the balance of positive and negative loops. The basic structure of a feedback loop is "a closed path connecting in sequence a decision that controls action, and the level of the system, and information about the level of the system, the latter returning to the decision-making point" (Forrester, 1968, pp. 1-7, and 1971). Boundary definition is a necessary prerequisite for modeling.

Fifth, the model is based on specification of functional relationships in a feedback system and not estimation of stochastic relationships. Emphasis is on interlocking feedback with non-linear relations in a system of differential equations rather than (1) best fit criteria; (2) simultaneous solutions of algebraic equations; or (3) optimization of key parameters.[1]

Sixth, simulation is for purposes of conditional analysis, not point prediction. Policy modeling involves simulation of system behavior over time. Policy testing identifies impacts of interventions in system behavior.

[1] There are no methodological reasons why stochastic analysis cannot be undertaken within a system dynamics paradigm or, alternatively, why the parameters of a system dynamics model cannot be estimated from empirical data.

Procedure

A typical system dynamics simulation design involves the following steps:

- Define the problems to be investigated;
- Specify system boundary (delineating endogenous, exogenous, and excluded variables and relationships);
- Identify the reference mode of the system, i.e., the behavior to model and reproduce initially as the first formal representation of the problem at hand;
- Formulate the causal loop diagram representing the feedback structure of the system;
- Formulate the computer flow diagram, an unambiguous graphical specification of the system structure;
- Write the equations;
- Examine the simulation output, noting the model behavior for changes in the structure or in the parameters and undertaking systematic experiments with alternative policy;
- Make policy recommendations based on the results of the sensitivity runs.

Simulation language

The major elements of the simulation language are: (1) levels, which are the state variables in the system and the major structural determinants; (2) rates, which control the fluctuations in the state variables; and (3) auxiliaries, which help define the rates. The key mathematical tool is integration, through which a quantity (level) at any given moment is related to the rate of change of that quantity over time. Rates represent "decision," and levels are the accumulation of all past decisions. In reality, rates are action streams that may be determined by conscious human intervention (e.g., all policies enter the system as rate changes), but may also be determined by structural factors in the system. Auxiliary variables are expansions of the rate equations to assist understanding of the relationship between levels and rates. Auxiliaries and rates may change rapidly; levels change more slowly.[2]

To write the equations for a system dymanics simulation, the modeling language includes: delays, table functions, and constants. Delays represent material lags of material flows in the system and lags in flows. Delays may enhance system stability since they can dampen the effect of an exogenous "shock." Table functions are the functional relationships between two variables as a way of incorporating nonlinear relationships. They are depicted as graphical representations of bivariate relations, and the shape of the graph provides the information necessary for specifying the value of a variable at any given moment.

Two important graphic tools of system dynamics -- the causal loop and the flow chart -- serve as important aids to theoretical formulation and to computation. The causal loop (see example in Figure 1 below) is designed to articulate the theoretical linkages of the structure, process, and outcomes in the system modeled. The flow chart is an explicit representation of the major elements of the simulation language utilized (see example in Figure 3 below). The

[2] The computer language for system dynamics is known as DYNAMO. A simplified version of the language developed for personal computers is known as STELLA. A version of DYNAMO is also available for personal computers.

computer equations can therefore be derived explicitly from the flow chart (by specifying the required parameter inputs and values of functional relationships).

In Section 4 below we present the theoretical framework, structure, processes, and outcomes of a major system dynamics simulation of the world oil market. The detailed description is designed to help the reader appreciate the characteristics of system dynamics, its usefulness for policy purposes, and the advantages of specifying interlocking feedback relations in an intertemporal framework.

3. COMPARISON OF METHODS: ECONOMETRICS AND SYSTEM DYNAMICS

This comparison involves key generic issues at the base of all simulation techniques:

System boundary

Econometric (estimation and simulation) models are open systems, including many factors that are, relative to the time horizon under consideration, not determined by the system modeled. Exogenous variables are needed to estimate the parameters of the model, and there are specific assumptions about the nature of these exogenous variables, and restrictions on their use. By contrast, a system dynamics simulation seeks to represent a closed system and generate reasonable system behavior without recourse to exogenous variables. It is assumed that if all the key relationships are included in the feedback structure, then the behavior modeled will be adequately simulated.

Stochastic factors

Econometric models assume that a simulation (or a forecast) is not meaningful unless accompanied by a statement regarding level of uncertainty. An econometric model makes use of information contained in the level of uncertainty and in the distribution of residuals. By contrast, system dynamics assumes that random factors are not strong enough to determine system behavior.

Validation

Econometrics adopts a statistical perspective on validation independent of the purposes of the analyst. For system dynamics validation is defined not by external criteria but by the extent to which the model replicates historical behavior and is useful in decision-making.

Uses of data

Econometric models require data to estimate the parameters and evaluate the validity of the model based on statistical requirements and constraints. Data issues are not as central in system dynamics simulation as in econometrics. System dynamics emphasizes specifying functional relations among variables rather than identifying their statistical relationship.

Reference mode

System dynamics simulation defines the system behavior in equilibrium as its basic reference. Alternative simulations are then compared with the reference or base case. Econometrics simulation based on parameters estimated empirically are used as the base case.

Time frame

System dynamics simulations generally adopt a long-range perspective. Econometrics conventionally focuses on the shorter range, since the parameters estimated from empirical data

are viewed as valid only over the data base period. (The algorithims of system dynamics can adapt to any time perspective with any time interval.)

Feedback

Feedback relations are built into system dynamics simulations by assumption, algorithm, and functional specification. Econometric models are generally less dynamic and less prone to feedback specifications, due partly to the nature of the algorithms, and partly to established practices in the field.

On balance, each of these issues is of importance for policy (or engineering-economic) modeling. The choice of simulation approach should ultimately be dictated by the objectives of the analyst. The following section shows the structure and behavior of a system dynamics simulation model of the world petroleum market.

4. INTERNATIONAL PETROLEUM EXCHANGE MODEL (IPE)

The International Petroleum Exchange Model (IPE) was developed at MIT in order to analyze the behavior and interaction of structure, processes, and actors in the world oil market.[3] The IPE model provides an integrated view of the market and key actors to ensure (and enforce) internal consistency as planners or analysts engage in assessments or forecasts of the world oil market. The first version was completed by 1980.[4] The model has subsequently been respecified and disaggregated on both the "demand" and "supply" sides.[5]

As a simulation model IPE focuses on price determination and its worldwide effects. The parties in the exchange are aggregate buyers (the oil-consuming and importing countries of the OECD), aggregate sellers (the oil-exporting countries of the Gulf region in the Middle East), and international managers (the major oil companies). Non-Gulf producers are assumed to play a role in market equilibration in the short run.[6]

The model is based on key equations that represent dynamic behavior and a set of accounting equations that monitor the effects of this behavior. The core features are demand, supply, and price relationships. The price of oil (anchored to the world prices) is endogenous, computed as a function of actual capacity utilization, desired capacity utilization, and costs, which are related to the proven reserve decline rate.

---

[3] The focus of this summary is on the demand side; however, the IPE model is of the entire market, and the supply side (which is not described at any length here) is specified in considerably greater detail than the demand side.

[4] For a rationale, description, and illustration of the early, complete version, see Choucri (1981).

[5] An early version of the disaggregation was developed as part of a Stanford University exercise at the Energy Modeling Forum. IPE model re-specifications have involved a number of researchers at MIT. Foremost include David Scott Ross, Brian Pollins, Michael Lynch, and, in the 1987 version, Christopher Heye. Michael Lynch was instrumental in the re-specifications of the model for the EMF6 exercise. For results, see "World Oil," EMF Report #6, Vol. 2, 1982.

[6] The model is initialized at 1970 values -- to test for robustness of behavior over a known set of values from 1970-1980 -- and computations are made to 2020. This has helped develop confidence in the analytical structure, as the price changes of the 1970s decade were ill-simulated. The aggregate actors are disaggregated into individual groups and/or states in recent versions of the IPE model.

Price changes affect both the quantity demanded and the amount supplied. In turn, supply and demand adjust to price. There are time lags involved on both sides. On the supply side there are the lags associated with investment delays. In the short run demand adjusts to price, and supply from the Gulf is relatively unresponsive. Non-Gulf supplies adjust to meet demand at the prevailing price. Over the longer run both supply and demand adjust to price and in turn influence the final determination of price.

Table 1 presents a simplified description of the IPE model, for descriptive purposes only rather than to depict the precise relationships. The left-hand variables are endogenous, as are some of the right-hand variables. Once quantities supplied and demanded are generated -- at a particular price -- consumer oil imports, producer investments and imports of goods and services, and profits from oil sales and investments in the oil industry are computed.[7] Figure 1 shows, in simplified form, key functional relationships.

Demand

Quantity demand is modeled as a dynamic process, adjusting to prices, to substitutes, and indirectly to income. Total demand is a sum of demand of (i) buyers and (ii) sellers. In the case of buyers, the oil consuming countries of the West, demand is a function of forecasted OECD demand, adjusted by substitute availability, domestic oil production, and price. The OECD calculations are themselves based on income assumptions and expected price.

The effect of energy substitutes upon demand is thus: we compare the oil price with the projected substitute price at that point in time; substitute availability is determined by the price of substitute energies in any given year. The higher the price of oil, the greater the percentage of oil demand covered by substitutes. Imports demand is computed residually, as the difference between demand and domestic production.

For demand of the sellers, the oil-exporting countries, regression estimates have been obtained based on the impact of population and capital stock on oil consumption. Estimates were derived from pooling cross-sectional and longitudinal data for key Gulf countries.

Supply

Supply is stipulated as (i) Gulf production; (ii) production in consumer countries; and (iii) non-Gulf LDC production. The entire production process in the Gulf is modeled explicitly, taking into account physical conditions, technological factors of exploration and development, and economic ones of costs and price.[8] Domestic supply in the consumer countries is briefer, but takes into account price and depletion.

[7]This specification originates with the initial versions of the model when written in the early 1970s, namely, representing interaction and interdependence among major actors (and agents) in the oil market.

[8] The model specifies the production process as described in Adelman (1972).

Table 1. Simplified analytical representation of the IPE model.

	Key equations	Variables
Supply	$TS = S_p + S_c + S_R$	TS = total supply
	$S_p = f(PC, R)$	S_p = producer supply from Gulf
	$S_c = f(\hat{S}, P, DR)$	S_c = consumer supply
	$S_R = TD - S_p - S_c$	PC = production capacity
		R = reserves
		$\hat{S}$ = base OECD series
		P = price
		DR = decline rate
		S_R = output of residual sellers, inventory changes, etc.
Demand	$TD = D_c + D_p$	TD = demand
	$D_c = f(\hat{D}, P, E)$	D_c = consumer demand
	$D_p = f(Pop, K)$	D_p = producer demand
		$\hat{D}$ = Base OECD series
		E = alternative energy sources
		Pop = producer population
		K = producer capital stock
Price	$P = T + C + MK$	T = tax rate
		C = cost
		MK = corporate markup
Markup	$MK =$	$f(MK_0, PCU, DR, \Lambda COMD)$
		MK_0 = base markup
		PCU = production capacity utilization
		$\Delta COMD$ = change in consumer import demand
Consumer import demand	$COMD = D_c - S_c$	
Market clearing	$S_R = COMD + D_p - S_p$	

Source: Choucri (1981: 25).

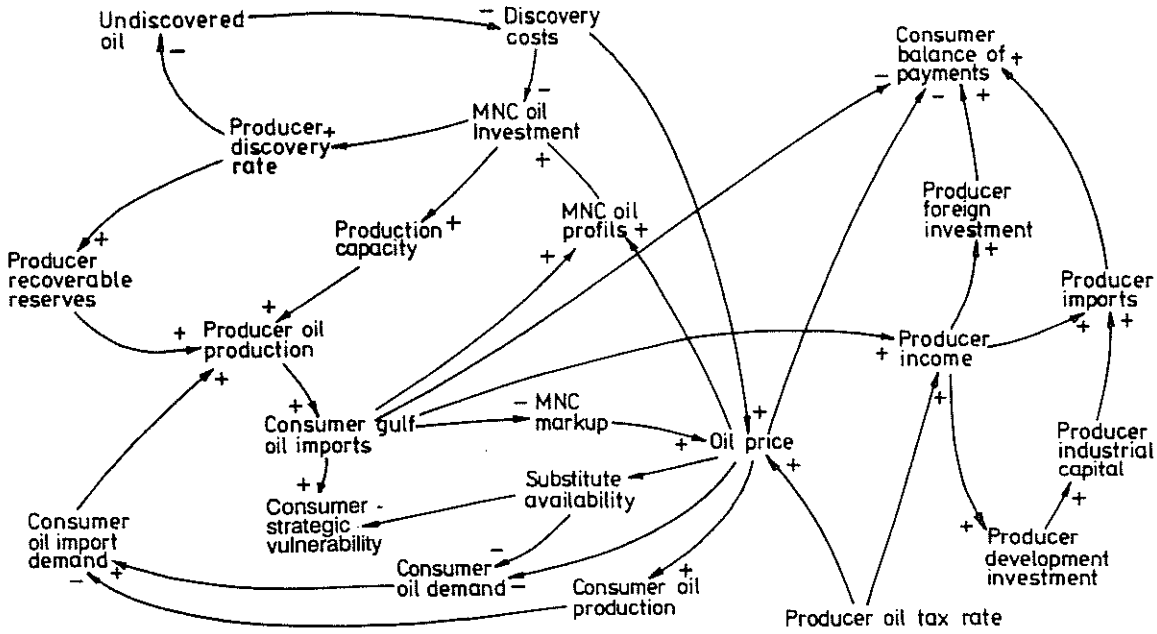


Figure 1. Simplified Model Overview of Major Causal Relations.

Source: Nazli Choucri with David Scott Ross. International Energy Futures: Petroleum Prices, Power, and Payments. Cambridge, Massachusetts: The MIT Press, 1981, p. 30).

Production in oil-exporting countries is affected by demand, given known reserves and production capacity. While reserves of oil in the ground are in producer countries, investments in exploration and development are undertaken by oil companies. Effective exploration investments determine the oil discovery rate that transforms undiscovered oil (initialized at 500 billion barrels for the Gulf fields) into recoverable oil.[9] Changes in recoverable oil-in-place in producer countries varies inversely and non-linearly with output. Investments in development and production costs determine production capacity. The ratio of production to production capacity sets the amount of capacity utilized.

Supply from domestic sources in consumer countries draws on base OECD projections adjusted by price and the rate of depletion. Supply elasticity, indicating the responsiveness of quantities offered for sale to a change in price, is a long-run coefficient set at 0.3. In the short term, of course, supply is relatively inelastic. (The current version of the model specifies both a long-term and a short-term price elasticity.)

Supply from non-Gulf sources is modeled in the same way as consumer oil production, i.e., as responsive to price and to reserves. Non-Gulf suppliers clear the market at each period.

[9] The precision of this estimate is not critical to the model behavior or output generation. It serves only to "anchor" production in some referent.

Price determination

Price is a function of producers' base price, production costs, and a markup factor. Producer price is determined by the capacity utilization rate. Its influence on final price and on the decline rate depends on its initial size as well as on market conditions. The markup factor takes into account capacity utilization in the production process, the impact of the decline rate, and the gap between the consumer countries' demand for imports from the Gulf and the Gulf's ability to meet that demand. Markup is a dynamic, endogenously determined influence on price. Through markup, the system-wide supply and demand conditions influence price, and in turn prices influence the prevailing conditions in the oil market as a whole. The ratio of markup to total price differs under different market conditions and size of the producers' base price.

Adjustment mechanisms

The IPE model combines the characteristics of two types of economic models -- the dominant firm model for a short-run analysis and a longer-run adjustment process of quantities to price. The dominant firm model applies to the Gulf producers who make price and quantity decisions in the short run by setting the base price of crude and/or the amounts to be produced or capacity to be utilized and in the longer run on the size of the residual market (where excess demand is met in the short run by non-Gulf producers) and on the price responsiveness of demand.

In the short run consumers can influence the size of the residual supply, increase domestic production, and to some extent cut their imports. In the longer run they can adjust demand and expand the use of alternative sources of energy.

As the markup factor represents the immediate adjustment to prevailing market conditions, it is intended to reflect the oil industry's (the international corporations') influence on and responses to market conditions. In the longer run the oil industry influences exploration and development through their investments in oil.[10]

Computational structure

The IPE model is composed of seven sectors, designed to represent the physical characteristics of oil production, the economic context and constraints, and the international financial exchanges that ensue from the trade in oil. See Figure 2 for sectoral relationships.

- the supply sector represents the physical stages of oil production, tracing the process from exploration for oil-in-place and the development of recoverable reserves to the installation of productive capacity and actual production.

- The finance sector calculates oil import expenditures for consumer countries, corporate profits and oil investments for the oil companies, and oil revenue for the producer countries.

- The management sector specifies investment decisions affecting the supply of oil. For analytical purposes, the model specifies the major investments of the multinational corporations in development and exploration, based on information drawn primarily from the supply sector, in conjunction with considerations of oil demand from the consumer sector. (Investments are separated from the agent, i.e., countries vs. corporations.)

[10] Due to changing jurisdictional conditions, this markup factor can be incorporated in producer country behavior. Originally, however, it was designed as a means of explicitly incorporating the international oil companies as distinct actors in the world oil market.

- The price sector calculates oil price based on inputs from other sectors. Producer price is endogenous; production costs and the corporate markup are also taken into account. Once calculated, the effects of price are then transmitted throughout the model to compute its financial and security implications for producers and consumers.

- The producer sector models the process of industrial development that generates demand for development investments and for imports of goods and services.

- The consumer sector computes demand for oil imports and monitors consequences for the consumers' strategic vulnerability and dependence upon external sources of supply. Domestic supply and demand and energy substitutes are specified.

- The international economic sector calculates the consumer balance of payments and traces the foreign investments of the producing states; it registers and links the consequences of actions taken by the consumer countries, the producers, and the international oil companies. (This sector helps achieve "closure.")

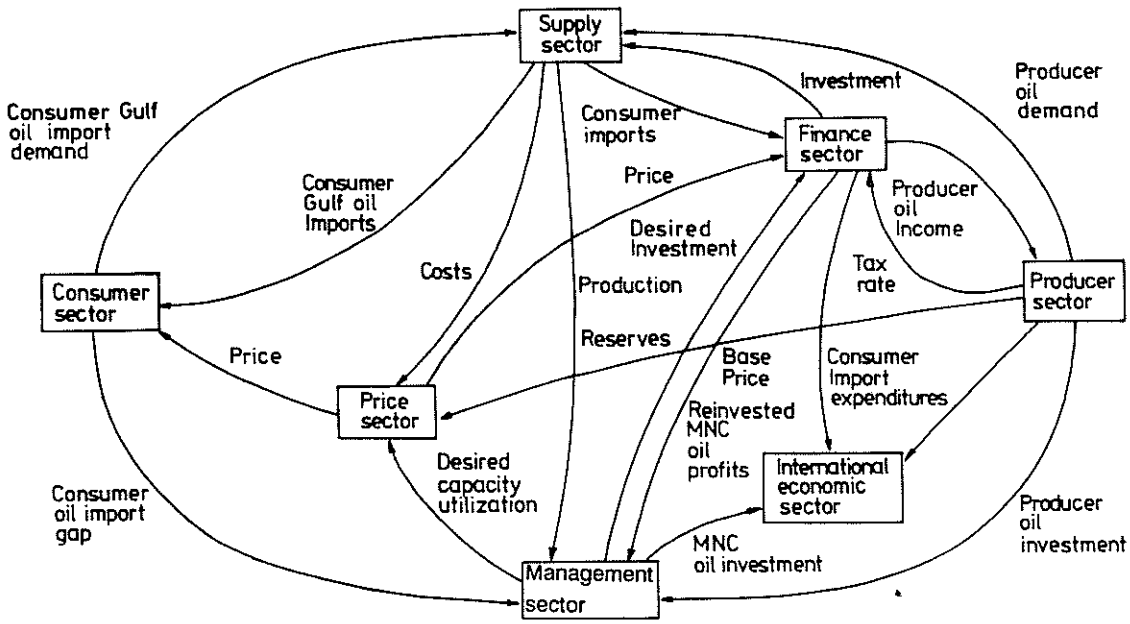


Figure 2. Select linkages -- sector link diagram.

Note: Only one component of markup is derived in the management sector.

Source: Nazli Choucri with David Scott Ross: International Energy Futures: Petroleum, Prices, Power, and Payments. Cambridge, Massachusetts: The MIT Press, 1981, p. 38.

The time frame of the IPE model is 1970-2000. Few experiments have been run beyond 2000, although they are technically feasible. Results are presented at annual intervals; however, the actual computations are done over 10 increments per year. The demand, supply, and price formulations are computed iteratively. The flow chart diagram for the consumer sector is given in Figure 3 to show the DYNAMO conventions for representing feedback and functional relationships.

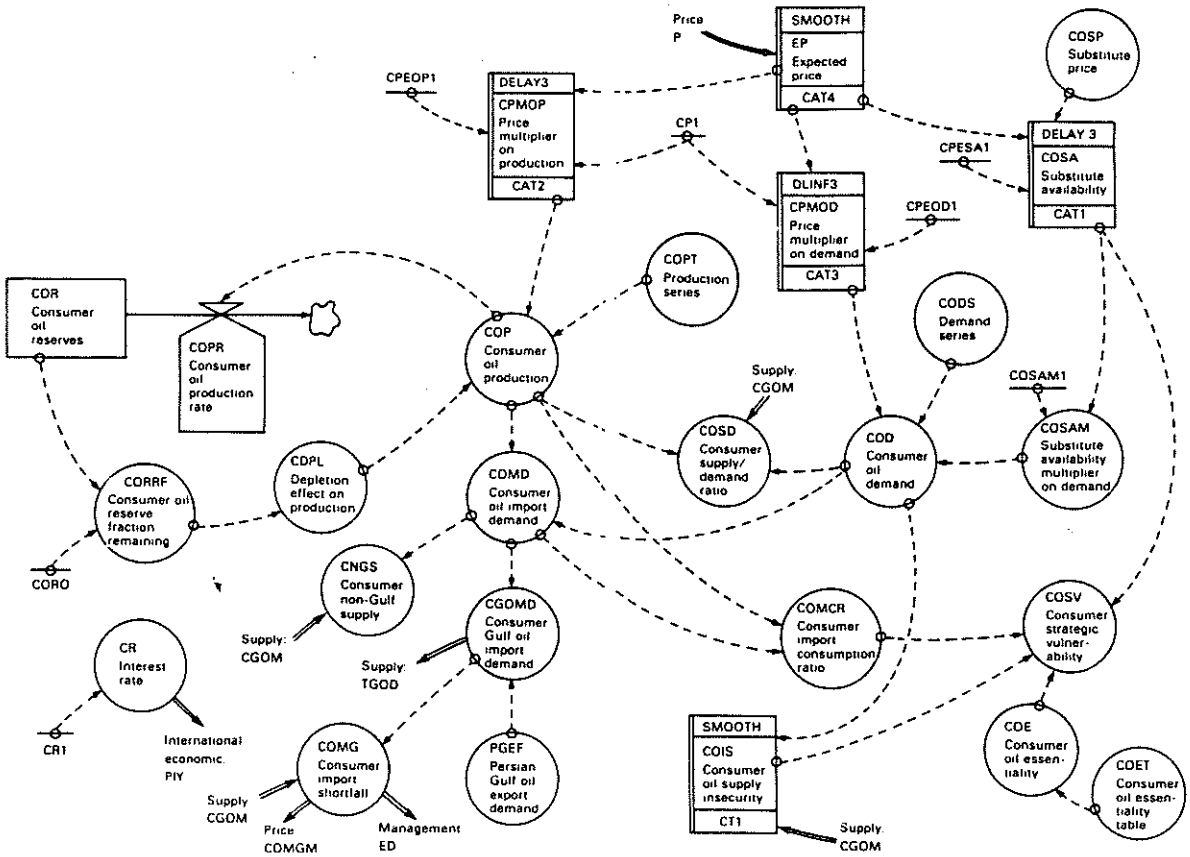


Figure 3. Computer flow diagram of the consumer sector.

The IPE model has been disaggregated on both the supply and demand sides to enable analysis and simulation of the behavior of individual agents and actors in the world oil market. The initial disaggregations were for the EMF 6 exercise as indicated above, and further refinements are currently being made for the EMF 11 exercise. Space limitations preclude a detailed discussion of the disaggregation strategy, results, and implications for policy analysis.

Source: Choucri (1981: 106).

Re-specification of the IPE model

The following regional sectors and the individual disaggregated components have been in the IPE model:

- OECD, aggregated from individual sectors representing the United States, Western Europe, Japan, and Canada-Australia-New Zealand (CAO);
- OPEC, aggregated from Middle Eastern (including Algeria and Libya) and non-Middle Eastern sectors;
- other Third World countries.

This disaggregation is designed to test for adjustments of, and effects on, individual (national and regional) behavior and to improve policy analysis by embedding the characteristic features of different buyers and sellers in the model (via functional relationships and parameter values).

Re-specification of supply and demand

Demand estimation. Three types of demand functions are used in the IPE:

- OPEC demand is estimated by means of regression estimates determining the response of demand to change in oil income, population levels, and non-oil GDP.

- Mexico demand is estimated using initial (1970) demand, assuming a certain price-independent growth rate and modifying this by means of a long-term price elasticity.

- OECD and other Third World demand is based on a more complicated growth mechanism with feedback from oil prices, modified by long- and short-term demand elasticities.

Supply estimation. Oil production is determined as follows:

- For Canada-Australia-New Zealand, production is equal to the minimum of 95 percent of capacity or total demand. Production capacity is derived from initial capacity and a price multiplier.

- Non-OPEC Third World production is set at maximum levels, 90 percent of capacity. The difference between production and demand is exported.

- Mexican production is determined directly from a base series, modified by a price elasticity.

- U.S. production is determined by price and depletion, modifying the initial production series; Western Europe production uses price and depletion, modifying a production series which increases to show the effect of the North Sea coming on-line in the 1970s; Japan production is set at zero.

- Capacity utilization for the two OPEC groups (Middle East and non-Middle East) is determined in part concurrently; actual capacity is determined separately and with a different specification. Middle Eastern capacity is determined by investments based on estimated demand. Non-Middle Eastern capacity is based on an initial level, modified by a price multiplier and reserve depletion. Capacity utilization is calculated by: (a) setting aside sufficient capacity to meet domestic demand, then (b) sharing the remaining capacity equally between Middle East and non-Middle East producers.

- Synfuels, defined as liquid fuels derived from coal or oil shale, are produced only in the U.S. sector, using a price-adjusted (0.3) base production series. The base series assumes 50,000 barrels per day will be produced by 1990 and 5 million barrels per day by 2000.

Economic growth

Base economic growth for the major consuming regions is supplied exogenously (see above demand analysis). For most regions, growth is estimated to range in the 2 to 3.5 percent interval.

Economic growth in OPEC countries is estimated in the form of non-oil domestic capital, which is a function of investment and depreciation. Investment is derived from regression estimates on population, oil income, and existing non-oil capital. Capital is assumed to be 2.5 times total GDP, a rough indicator commonly used in development economics.

Price specification

Price is generated endogenously; alternatively it can be used as a policy variable (by modifying the base price exogenously). Computationally, the base world price is with production costs and markup added to yield final price. The endogenous mechanism represents the tightness of the market (as measured by the capacity utilization in the Middle East) and the decline rate. For simulation purposes the capacity utilization rate is specified as a basis against which adjustments occur.

The higher the capacity utilization (over the specified percent and subject to change for policy analysis), the greater will be the price increase (the lower the rate, the less the impact on price). The model is designed so that different degrees of responsiveness can be examined. (The decline rate of reserves has a relatively small impact overall in the market.)

5. RESULTS OF IPE SIMULATION

Of the many ways to illustrate the behavior and validity of simulation models, we have undertaken three strategies: (1) comparing model results for the reference (base case) to actual known values (here, 1970 to 1986); (2) comparing IPE model results with those generated by other world oil models; and (3) examining behavior for internal consistency and plausibility, given known parameters of the world oil market.[11] Here we present the results of the first type of validity -- a comparison of major IPE outputs with the historical record.

Table 2 shows the actual vs. IPE values for Central Planned Economies (CPE) world oil supply, demand, and prices, 1970 to 1986.

Table 2. IPE Base Case Results and Historical Record.

	Non-CPE World Supply		Non-CPE World Demand		Price -- 1980$'s	
Date	Actual	IPE	Actual	IPE	Actual	IPE
1970	38.1	39.1	39.0	39.6	4.30	4.13
1971	40.0	41.3	41.1	41.9	4.81	4.20
1972	42.1	43.6	43.9	44.2	5.42	4.45
1973	45.9	45.6	47.4	46.4	5.47	4.67
1974	47.6	46.8	45.8	47.7	17.83	11.34
1975	43.9	44.3	44.6	45.2	16.68	16.78
1976	47.4	45.1	47.4	46.1	16.84	16.90
1977	49.1	46.3	48.9	47.7	17.11	17.06
1978	48.9	48.4	50.3	49.5	16.23	17.24
1979	51.3	49.9	50.9	51.0	18.23	17.42
1980	47.9	48.3	48.3	49.5	28.00	29.17
1981	44.5	45.4	46.5	46.9	29.80	28.90
1982	42.1	43.7	45.1	45.5	30.47	28.67
1983	41.5	42.9	44.9	44.9	26.82	28.65
1984	42.9	42.7	45.5	45.0	25.07	28.18
1985	42.5	42.9	45.3	45.4	24.07	27.39
1986	44.6	44.5	46.4	46.9	13.00	16.76

Sources: British Petroleum, BP Statistical Review of World Energy, various years, and 20th Century Petroleum Statistics, 1986).

[11] See Choucri (1981) for detailed results of the initial IPE validation analysis. The model is updated approximately every three years. The first version was written in two years. A typical updating effort takes about 3 man-months.

6. SIMULATION OF ENERGY DEMAND

To demonstrate the demand side of the IPE model, two price simulations are shown below: the base case vs. the high price scenario.

The base case price is endogenously determined by the IPE model. The IPE makes two assumptions regarding the future availability of crude oil reserves and the pace of energy conservation which tend to distinguish it from other forecasts: (a) it holds a somewhat "optimistic" outlook for future crude oil discoveries and (b) it assumes that conservation and fuel switching will persist, as will further development of oil-saving technologies. IPE therefore generates both lower demand and greater non-OPEC supply than most other models of the world market. The IPE currently forecasts (endogenously) real oil prices to remain flat in the $18-$20 per barrel range (in real terms) for the foreseeable future (i.e., the year 2000).

The high price case is based primarily on existing conservative estimates of reserves of non-OPEC producing countries and lower economic growth among industrial nations, roughly in the area of 2.5 percent.[12] Implicit in the high price case is the assumption that the energy to GNP ratio will end its current slide and that oil demand in the future will be highly inelastic. A corollary is the presumed ability of OPEC to regain almost complete control over all pricing, including a capacity to significantly raise real prices over the long run, in spite of substantial gains in energy conservation. A price comparison is given in Table 3.

Table 3. World Oil Price: Two Contending Forecasts.

	(1986 Dollars)	
	IPE Base	High Price
1986	20.10	14.02
1987	18.79	18.50
1988	18.30	18.51
1989	18.35	20.53
1990	18.36	22.55
1991	18.39	23.58
1992	18.44	24.59
1993	18.51	25.55
1994	18.61	26.53
1995	18.72	27.45
1996	18.99	28.56
1997	19.12	29.58
1998	19.21	30.57
1999	19.28	31.58
2000	19.38	32.58

Figure 4 compares the results for energy demand of the alternative price scenarios. The demand for petroleum is lower in the high price case over the long run. Though all regions show some demand increases under this higher price scenario, for some regions, this growth is minimal. Not shown here, for example, Western European demand is projected to increase only .3 percent (total, not per annum) between 1986 and 2000 in high price case. U.S. oil demand, after moderate increases through the late 1980's, shows virtually no growth throughout the 1990's under

[12] U.S. Department of Energy (1987). (Since this DOE scenario only extended to 1995, prices in the IPE beyone 1995 were extrapolated.)

this scenario. Most other regions (not shown here) display increases of only 8-10 percent (total) over the interval from 1986 to 2000.

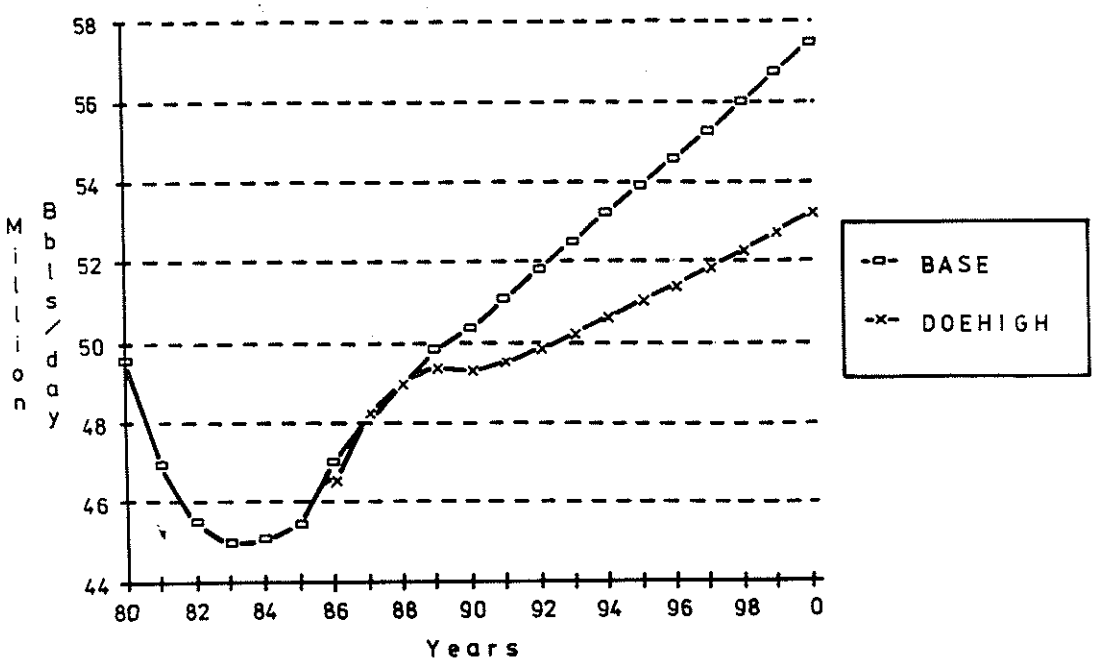


Figure 4. Scenario Comparisons: World non-CPE Oil Demand

7. CONCLUSION

This discussion of simulation approaches to policy modeling focuses on system dynamics, with a partial comparison with econometric simulation. The structure, respecification, and system behavior of the International Petroleum Exchange model were presented to illustrate the design of a simulation model of the world oil market. Finally, a simulation of the demand side was undertaken showing model behavior in two cases: the base case vs. the high price case.

The IPE model simulates all major CPE suppliers, thereby tracking the interdependence between OPEC and non-OPEC suppliers. Although the primary focus of this discussion has been the model's restructured demand specifications, the supply and price sectors of the IPE model are modeled in considerable detail. The supply sector, particularly that of the OPEC producers, has greater empirical grounding than the demand specifications. The IPE model is distinctive, even unique, among energy models in the degree of attention it pays to behavioral characteristics of the OPEC suppliers. The financial aspects of production, including both producer government and multi-national oil company activities, are particularly well specified.

Because we regard OPEC behavior as crucial to the determination of world oil prices, we have recently embarked on a new effort to model the internal activities of OPEC in even greater detail.[13] Specifically, we are attempting to model the decision rules for production sharing within OPEC itself. Price changes can best be understood by examining the production arrangements conferred and abrogated by OPEC. Market structure since 1985 reflects the importance of non-OPEC countries and the extent to which OPEC behavior affects global production and supply.

[13] Currently being completed is a detailed re-specification of the supply-side to differentiate among the behavioral "rules" used by different OPEC countries in their determination of production. These "rules" are embedded in differences among them with respect to reserve, production capacity, and capacity utilization. This module can be run autonomously (and is labelled CAPRO model).

REFERENCES

Adelman, M. A. (1972). The World Petroleum Market, Baltimore: Johns Hopkins Press.

British Petroleum Review of World Energy, various issues.

Coucri, Nazli with David Scott Ross (1981). International Energy Futures: Petroleum Prices, Power, and Payments, Cambridge, Mass.: MIT Press.

Choucri, Nazli and Thomas W. Robinson, eds. (1978). Forecasting in International Relations: Theory, Methods, Problems, Prospects, San Francisco: W. H. Freeman.

Choucri, Nazli, Christopher Heye, and Michael Lynch (1989). "Analyzing Oil Production in Developing Countries: A Case Study of Egypt," MIT: Department of Political Science.

Energy Modeling Forum (1982). "World Oil," EMF Report #6, Vol. 2.

Forrester, Jay W. (1961). Industrial Dynamics, Cambridge, Mass.: MIT Press.

Forrester, Jay W. (1968). Principles of Systems, Cambridge, Mass.: MIT Press.

Forrester, Jay W. (1971). World Dynamics, Cambridge, Mass.: Wright-Allen Press.

Goodman, Michael R. (1974). Study Notes in System Dynamics, Cambridge, Mass.: Wright-Allen Press.

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

Petroleum Economist, various issues.

Pindyck, Robert S. (1979). The Structure of World Energy Demand, Cambridge, Mass.: MIT Press.

Richmond, Barry, Steve Peterson, and Peter Vescuso (1987). An Academic User's Guide to STELLA, Lyme, N.H.: High Performance Systems Inc.

Shearer, J. Lowen, Arthur T. Murphy, and Herbert H. Richardson (1971). Introduction to System Dynamics, Reading, Mass.: Addison-Wesley Publishing Company.

U.S. Department of Energy (1987). International Energy Outlook 1987, Washington, D.C.: U.S. Government Printing Office.

Energy Vol. 15, No. 3/4, pp. 379–386, 1990
Printed in Great Britain

0360-5442/90 $3.00 + 0.00
Pergamon Press plc

3.6. CROSS-IMPACT MODELS

ANDREW J. LIPINSKI

Institute for the Future (ret.)
Menlo Park, California
U.S.A.

1. INTRODUCTION

In trying to simulate features of a complex, dynamic situation which has not been modeled before and for which no reliable historical information exists, a tool is needed that would capture the essential interactions between variables and produce forecasts of the variables of interest to decision makers.

One of the options open to the analyst is to start modeling from scratch, beginning with a concept of the structure of the problem, deciding which variables to include, and inserting the functional interactions from whatever source -- intuition if need be. Many new models start this way.

Another option is to employ a ready-made modeling structure and limit the contribution of the modeler to specifying the interrelations between variables. The examples are the spread-sheet models and the System Dynamics models for which standard software is available. They belong to a class of structural models. (Cearlock, 1977; Lendaris, 1980).

Whichever the choice, the analyst has to make another set of decisions. The first decision is whether to specify the variables in the model or to seek the advice of experts. The second decision is whether to generate the interrelations between the variables on his/her own or whether to rely on the judgment of experts, familiar with the nature of the interrelation.

The cross-impact family of models, now three decades old, was born within the futurist fraternity, alongside the development of Delphi. The foundation on which Delphi was based was that information sharing between experts would produce more "accurate" forecasts. We believe that it is as valid today as it was then, except that today we would apply the adjective "accurate" not so much to the value of the variable being forecast but, in a probabilistic interpretation, to the parameters of the distribution that is being forecasted. Thus a panel of experts may produce a better estimate of the mean of the distribution, or of its measure of variability. For example, a panel of experts may judge the roulette wheel to be running true and therefore producing a uniform probability distribution over a range of 0-36 or 00-36.

Given the background of the inventors, it was natural that from the very beginning, cross-impact models relied on expert judgment for both the selection of the variables and the specification of the interrelations between variables. In early versions of cross-impact the interrelations were between probabilities of binary events rather than variable values but difficulties with operating on probabilities soon forced a change to the values of the variables (or their time differentials). One feature however remained constant throughout the development and right up to the present. All cross-impact models generate dynamic, deterministic forecasts by taking into account interaction of the variables two at a time.

DYCIM (Dynamic Cross-Impact Model), the modeling technique described here, is the last in the long evolution of the cross-impact approach to simple modeling of complex situations where no historical data is available. DYCIM's immediate predecessor is KSIM, a simulation model of the cross-impact type (Kane, 1972).

The following material on DYCIM extensively quotes from the publication (Institute for the Future, 1983) and the paper (Lipinski and Tydeman, 1979).

We will first describe the DYCIM methodology, then the procedure involved in building the DYCIM model and, finally comment on the model's size and cost.

2. THE METHODOLOGY OF DYCIM

In all cross-impact models the variables can be imagined to be arranged in a square grid, each variable appearing in both columns and rows.

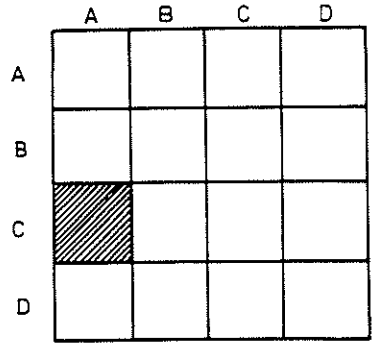


Figure 1. The Cross-Impact Matrix

Just as they could in a spread-sheet, the entries in the cells of a cross-impact model can be thought to contain a formula describing the interaction between the column and the row variable.

We will start by describing KSIM and then show how DYCIM differs from KSIM.

Kane proposed a cross-impact method that, although restricted to the cross-impact of trends on trends, applies to many real situations. The trend variables behave in a logistic fashion, and the net effects of the cross-impact coefficients go to zero as trends reach their maximum or minimum values.

In Kane's formulation, trends are scaled to have values lying between zero and unity. The impact of one trend on another is then given by

$$x_i(t_n) = x_i(t_{n-1})^{\pi_i} \tag{1}$$

where the power π_i is a quotient of two terms. The numerator is the sum of the negative impacts on trend i, while the denominator is the sum of the positive impacts. In particular,

$$\pi_i = \frac{1 + \frac{\Delta t}{2} \sum_{j=1}^{N} (|Q_{ij}| - Q_{ij})}{1 + \frac{\Delta t}{2} \sum_{j=1}^{N} (|Q_{ij}| + Q_{ij})} \tag{2}$$

Although the cross-impact Q_{ij} can have any functional form, in Kane's original formulation it is usually restricted to terms proportional to the value and the derivative of the impacting variables. Conceivably Q_{ij} could include more complex interaction terms -- for example, proportional to the product of the impacting variables -- but the problems of estimating the coefficients are large, especially if they have to be derived from subjective judgments. In their modification of KSIM Lipinski and Tydeman replaced Kane's cross-impact matrix with a more flexible functional form.

In DYCIM, this matrix, called the D(for dynamic) cross-impact matrix, has coefficients that are proportional to the derivative of the impacting variable but may vary over time: Thus, there might be a D-matrix in the current year D(0), another for the previous year D(1), another for two years ago D(2), and so on. It may vary over time.

More formally,

$$Q_{ij}(t_n) = \sum_{m=n}^{1} D_{ij}(n-m)\frac{dx_j(t_m)}{dt} \qquad n = 1, 2, \ldots \tag{3}$$

For example, a price change in a commodity, is estimated to cause a time-stream response in the impacted variable (e.g., sales). The response may be of any shape, can change sign and can be delayed. Therefore, variables can impact on themselves (after a delay). Figure 2 shows a typical impact.

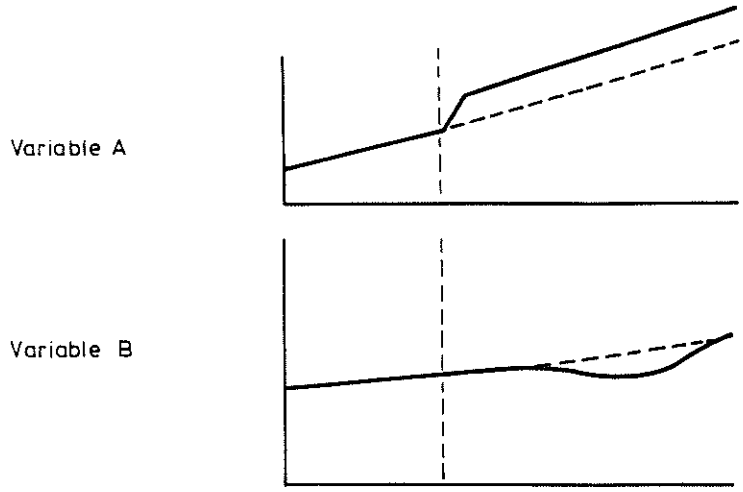


Figure 2. An impact of A on B

As defined, the D matrix is proportional to the derivative of the impacting variable. In a number of cases at the Institute for the Future, it was found that the impact ought also to vary inversely with the value of the impacting variable. An example might be a price sensitivity in which both the price change and the price level contribute to the size of the impact. Thus the price elasticity of demand for gasoline may depend on gasoline price as well as the increase in price.

To accommodate such effects, the derivative of the impacting variable in equation (1) can be selectively replaced by the derivative divided by the value of the impacting variable; i.e.,

$$\frac{1}{x_j}\frac{dx_j}{dt}$$

For computational purposes this term is in fact equal to the derivative of the natural logarithm of the impacting variable.

Typically, the structure of the model is determined by a panel of experts who decide which variables ought to be included and which variables interact with which. Then a second panel of experts assesses the nature of the interactions by drawing a graph of the response of the impacted variable to a hypothetical step in the impacting variable. This is equivalent to determining its partial cross-elasticity to the impacting variable.

Given the start up values of actual data, or the assumed starting values for the variable (when there is no actual data because, let us say, the product has not been introduced yet), the model solves the simultaneous equations describing all active interactions (some may be disabled at will) and produces forecasts of all the variables.

3. PROCEDURAL STEPS IN USING DYCIM

Cross-impact modeling is unusual because from the very beginning it incorporated joint modeling by experts. It follows then that in cross-impact successful forecasting depends as much on the procedural steps in joint modeling as on the model itself. Many of the following steps developed specifically for DYCIM would also be useful in any joint modeling effort. They were developed over a decade of experimentation and the experience gained during many forecasting efforts.

Seven distinct procedural steps are involved in using DYCIM. A brief description of each step follows.

Identifying the variables

Typically, identification of the variables starts with a discussion of the problem in which the management of the enterprise for whom the forecasting effort is being conducted are joined by the invited experts. Guided by the analyst, the panel identifies the variables that are needed to capture the essentials of the situation being modeled. The decision variables can be identified and added to the list.

Variables are then divided into exogenous and endogenous, or input and output. Normally, there would be no impact of the endogenous on exogenous variables, but the organizer could stipulate such a feed-back if needed. Decision variables are always exogenous.

Because of the time and expense of elicitation of the interrelations the number of variables should be kept small, preferably twenty or less. Of these two or three are likely to be exogenous.

Finally, the forecast for any variable can be thought to result from an interplay of influences, some positive and some negative. In real life there are few instances where all or most of the influences act one way and cause runaway behavior such as a market crash or a revolution. In most cases there happens to be a constraining variable or variables that, by their countervailing action, limit the excursion of the affected variable. In selecting the variables, the analyst should consider: What variable (input or output) would constrain the variable in question? Such constraining variable (or a proxy for such a variable) should be included in the list of variables else the model will "blow up".

Preparing preliminary forecasts

The reasons for preparing preliminary forecasts for all variables are: 1) if the variables are exogenous, their values will be "driving" the model; 2) if endogenous, their forecasts will be needed for eliciting the assessment of the interactions; and 3) the initial values of all the forecasts (typically the last value of available data) are needed to start the computation.

For each impacting variable, one has to postulate the size of a hypothetical step that will be used to elicit the interaction between the impacting and the impacted variable. The step can be

positive or negative, can be placed anywhere within the time span of the forecast and can be either a sustained change or a temporary disturbance. The size of the step has to be realistic else the participants in the workshop would find it difficult to assess its impact.

Preparing elicitation forms

Although N variables could have N^2 interactions, in practice 20 variables require no more than 50 interactions. For each interaction an elicitation form is prepared. It will be used to record the judgment and the expertise of each expert on each interaction. For 12 participants and 50 interactions 600 forms will have to be printed. A typical elicitation form is shown in Figure 3.

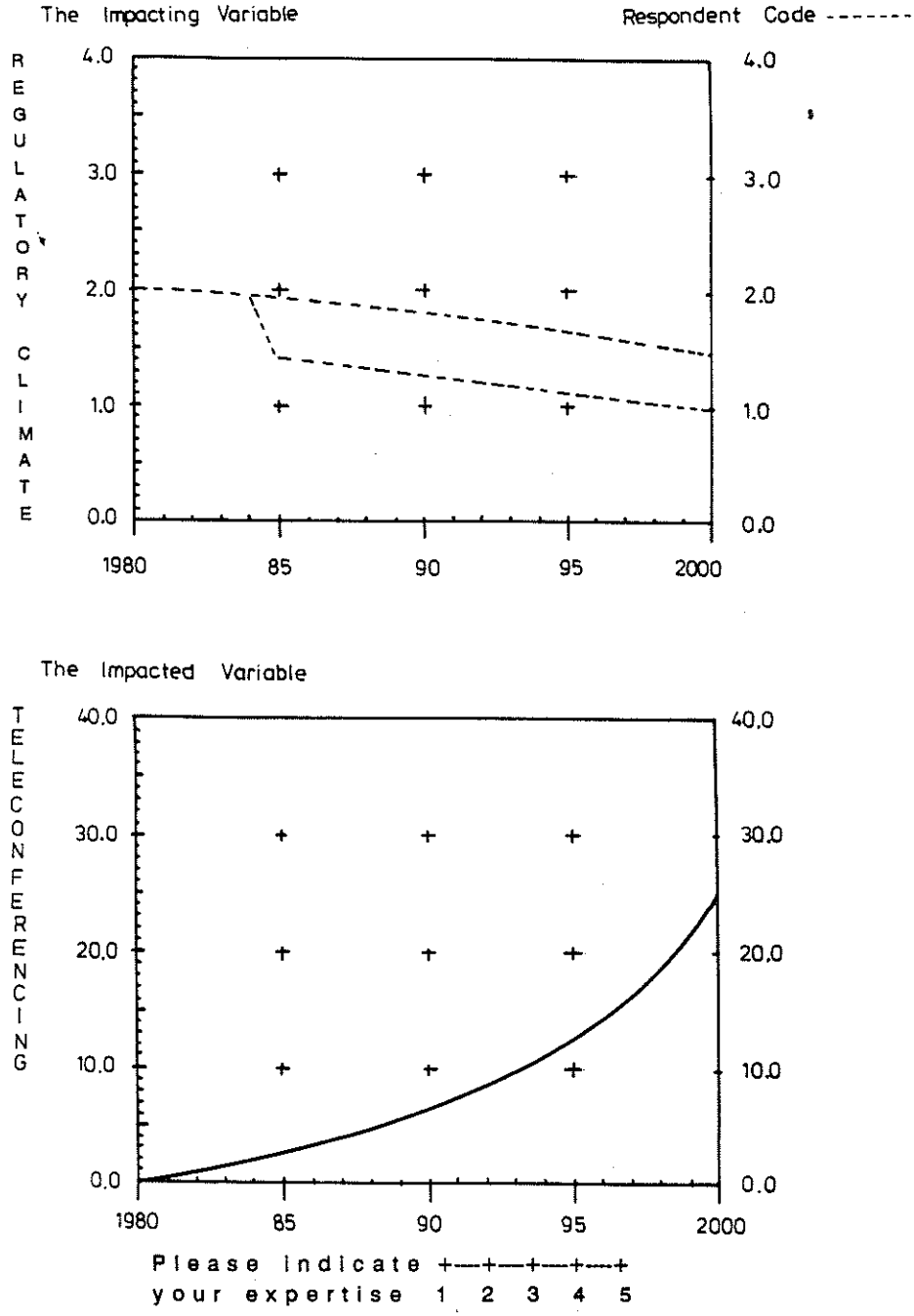


Figure 3. Typical Elicitation Form

Instructing the participants

Participants in the expert panel have been chosen because of their knowledge of the situation and their ability to communicate their experience to others. They are brought together for a workshop lasting 2-3 days.

In the opening session the participants are introduced to the DYCIM procedure. A particular emphasis is placed on the fact that none of the participants is expected to produce a forecast of the output variable(s) directly. Rather, the participants judgment will be needed of the interrelations (cross-impacts) of one variable on another. The organizer stresses that the participants have been selected for this particular exercise because of the knowledge that they possess might contribute to a better understanding of the complex interrelationship between the variables and thus allow the DYCIM model itself to produce a better forecast.

Next, the preliminary forecasts of all variables are introduced and the definitions of each variable are carefully explained, discussed by the participants and, occasionally, modified.

The structure of the problem is then explained and the most important interrelationships are identified and shown as a flow diagram. The preliminary forecasts and the flow diagram are displayed on charts and hung around the room so the participants can refer to them during the discussion.

At that time, the criteria for encoding expertise (Table 1) are introduced and discussed and the weighing scheme is explained.

Table 1
Self-Evaluation Criteria
A Guide to Ranking Expertise

Category

1. You are unfamiliar with, the subject if mention of it raises a veritable blank in your memory, or if you have heard of the subject but are unable to say anything meaningful about it.

2. You are casually acquainted with the subject if you at least know what the issue is about, have read something about it, and/or have heard a debate about it on TV or radio network.

3. You are familiar with the subject if you know most of the arguments advanced for and against some of the controversial issues surrounding this subject, have read a substantial amount about it, and have formed some opinion about it. However, of someone tried to pin you down and asked you to explain the subject in more depth, you would soon have to admit that your knowledge is inadequate.

4. You are quite familiar with the subject if you were an expert some time ago but feel somewhat rusty (even though you have kept reasonably abreast of current developments in the field); if you are in the process of becoming an expert but still have some way to go to achieve mastery; or if you concern is with integrating detailed developments in the area, thus trading breath of understanding for depth of specialization.

5. You should consider yourself an expert if you belong to that small community of people who currently study, work on, and dedicate themselves to the subject. If a professional society or conference organizers were to convene a seminar on this subject, you would expect to be invited to attend or to submit a paper. Other experts in field may disagree with your views but invariably respect your judgement.

From measurements done at the Institute, we know that most corporate decision makers tend to assign the weights given in Table 2.

Table 2
Weighing of Expert's Judgment

Expert's Substantive Expertise Ranking	Relative Weight
1	1
2	2
3	4
4	8
5	16

Next, the participants are shown a sample of a completed elicitation form such as in Figure 2. Emphasis is placed on the fact that most of the cross-impacts will be partial (that is, that more than one variable is likely to influence the behavior of the output variable) and therefore any one variable, by itself, is unlikely to have a pronounced effect. These cautionary remarks have been found necessary because most untrained respondents tend to exaggerate the partial cross-impact of one variable upon another.

A discussion of the problem's structure may identify identities, for example that A + B = C. In such cases only the effects of A on B and B on C should be considered.

If time permits, a dry-run exercise on a related subject should be conducted.

Finally, the elicitation begins. Each interrelation is introduced in turn and, after a discussion of several minutes, each participant records his/her judgment by drawing a revised forecast of the impacted variable and then records his/her expertise.

Encoding the results

For each cross-impact, we collect a judgment from each participant. Judgments indicating zero cross-impact are also collected because a weighted consensus judgment requires all judgments, including those of zero effect. For that reason, participants must enter their expertise on a form even though they do not modify the forecast of the impacted variable.

The completed forms are then placed on a digitizing tablet and, under computer direction, each participant's modified forecast of the impacted variable is traced. This procedure results in the automatic encoding of all responses for each cross-impact, which are then stored in a series of computer files.

The self-ranking of expertise is also encoded, either by marking the response encoded on an elicitation sheet, or, if there were a judgment of zero impact, by typing in the expertise directly into the computer. Throughout this procedure, the elicitation program guides the encoder along and, if necessary, requests repetition of the input data.

Processing the results

When all the forms have been encoded, the processing begins. First, by using the expertise ranking and the weights shown in Table 2, a weighted average of each cross-impact graph (i.e., the modified forecast of the impacted variable) is established. Second, the modified curve for the impacted variable is used to compute the cross-matrix elements in accordance with the basic

DYCIM formulation (Equations 1-3 above). This procedure is repeated for all cross-impact combinations.

Finally, the actual computation begins. Starting with the initial value of each variable and the appropriate cross-impact matrix coefficients, new values for successive time periods are calculated.

Displaying the results

As described, the output of the DYCIM model establishes a base forecast. Once this simulation has been run, various alternatives can be explored by varying the forecasts of the exogenous or driving variables. In the same way the sensitivity of the forecasts can be displayed by showing the effect of each impacting variable.

Uncertainty in the state variables (for example, the overall demand) is addressed by repeated runs of the model or the surface model derived from sample runs across the probabilistic domain of the state variable(s). (See section 2.6).

4. ABOUT THE MODEL

The main DYCIM model contains about 1000 lines of code. There are approximately 1500 lines of code in the five auxiliary models needed to process expert judgment, derive the interactions between the variables and selectively display the results. The main model is in FORTRAN, the auxiliary models in PASCAL.

Updating the input means reconvening the expert panel, essentially rerunning the workshop, an expensive procedure. A typical workshop requires two person months of the analyst's time, 5-10 expert days for the structure, and 30-50 expert days for the interactions.

REFERENCES

Amara, R. and A. J. Lipinski (1983). Business planning for an uncertain future: Scenarios and Strategies, New York: Pergamon.

Cearlock, D.B. (1977). "Common properties and limitations of some structural modeling techniques," Ph.D. dissertation, University of Washington, Seattle.

Kane, J. (1972). "A primer for new cross-impact language - KSIM." Technological Forecasting and Social Change 4: 192-43

Lendaris, G.G. (1980). "Structural modeling - A tutorial guide," IEEE Transactions on Systems, Man, and Cybernetics SMC-10: 807-40

Lipinski, H., and Tydeman, J. 1979. Cross-impact analysis - Extended KSIM. Futures 11, No.2: 151-54

Energy Vol. 15, No. 3/4, pp. 387–394, 1990
Printed in Great Britain

0360-5442/90 $3.00 + 0.00
Pergamon Press plc

3.7. ENERGY - ECONOMY MODELING: A SURVEY

JANUSZ COFALA

Energy Problems Department
Institute of Fundamental Technological Research
Swietokrzyska 21, 00-049 Warsaw
POLAND

THOMAS KUCZMOWSKI and JOHN WEYANT

Energy Modeling Forum, Engineering-Economic Systems Department
Stanford University
Stanford, CA 94305
U.S.A.

1. INTRODUCTION[+]

The relationship between the energy sector and the economy is central to the evaluation of energy options. Most concern with energy issues arises from the assumption that the character of future energy availability will have a major impact on the quality of life, and the level of economic activity is a primary measure of this quality. Opinion is divided sharply on the structure of the link between energy and the economy. Basic physical laws indicate that some energy is required for every activity and if adequate energy is not available the activity cannot take place. From this perspective, the historical growth of energy and the economy is cited as evidence that their future growth cannot be separated. In the short run, most would agree, for we must use the equipment and processes now in place, and their range of energy utilization is narrowly restricted. In the longer run, however, new equipment can be purchased, alternate transportation systems designed, the mix of desirable products changed, and new technology introduced. The same level of output might be obtained with a lower level of energy utilization and the quality of life maintained or even improved, some would say. This perspective is supported by the evidence of different energy utilization patterns and higher energy prices may explain the growth of energy demand in the United States.

If growth in energy availability is essential to the growth of the economy, then large expenditures are indicated for programs directed at expanding long run energy supply and lowering energy costs. However, if substantial flexibility exists for adjusting energy utilization and economic output then programs which facilitate this adjustment may be employed. The best policy probably requires a careful blending of both approaches, but it is certain that these choices will be influenced heavily by the expectations of the impacts on future economic growth. It is essential to understand the links between energy and the economy.

2. THE ROLE OF MODELS

Energy models alone cannot dispose of these difficult issues. They augment our capabilities for organizing the collective understanding of a problem. Modeling does not replace careful

[+]The introductory part of this survey borrows from EMF 1: Energy and Economy.

thinking, but seeks to exploit it. A model records what we know by providing an accounting framework, organizing the data and key relationships. At the heart of most models is some simple classification scheme. For example, the model may describe all energy consumption in terms of a few sectors, aggregating the millions of households, commercial establishments, or industries.

This permits the separate but consistent analysis of the major problem components. Specifying how these components connect may simplify other complex inter-actions. The implications of a given hypothesis about the structure of the energy system can be pursued. If the model also is detailed, computer implementation may be a further aid in pursuing the needed calculation. These contributions of modeling are substantial. The accumulation of knowledge and the pursuit of the implications of that knowledge are the essence of sound analysis. Formal modeling makes the process explicit and promotes its orderly evolution.

The limitations of models are important to recognize. The basic limits of our understanding are transferred to a model. Hence, the effects of noncompetitive practices are not to be found in the study of models assuming perfect competition. A model does not create a new theory, it explores the existing theory, often in a highly simplified fashion. This simplification is essential in modeling, permitting quick analysis of many relationships and the clarification of basic causal mechanisms. Without simplification, the only model of reality is reality itself, and only one experiment is permitted. With a model, many experiments are possible. And, unlike reality, we can take them or leave them. But the model is only an approximation to reality and must be judged accordingly. Many approximations to the same reality are possible. The application of any one model must be guided by the context of the questions addressed. For problems as complex as the evaluation of energy futures, there is no single model which addresses all issues. There are many problems for which there are no models at all. The models do not replace the effort of the analyst in determining the appropriate assumptions and problem structure, but they do extended the scope of the problems that can be studied.

The benefits of simplification and the potential for detail are found in the range of available models of energy-economic interactions. Very simple aggregate analyses of the interaction can place the problem in perspective. These aggregate models provide the conceptual background for the evolution of more sophisticated systems. As the potential deficiencies of the simple models are identified and the assumptions relaxed, more detailed models are developed and the range of application expands. Therefore, a basic framework is developed for characterizing the key concepts underlying the interactions between the energy sector and the economy.

3. A BEGINNING FRAMEWORK FOR ANALYZING ENERGY AND THE ECONOMY

The issue here is the impact of a relative scarcity of energy. None of the models implies that energy is not needed for the economy. If energy prices remain stable, then an increase in economic activity should produce an increase in the demand for energy. Of course, the future growth of energy demand may be less than the historical growth because of lower projections for population increase or a trend toward a disproportionately higher growth of the less energy intensive sectors of the economy. The difficult question is, can energy demand growth be dampened further by higher energy prices without proportional reductions in economic activity?

For simplicity, we restrict attention to the long run, when energy equipment and processes can be changed substantially. In the short run, the character of the problem is different and different models are appropriate. As a further simplification for this beginning discussion, we represent the economy in terms of just two inputs - energy and all other items. That energy is only a small component of the total U.S. economy. In the U.S. as of the 1970, the value of primary energy inputs did not exceed 4% of the GNP. The analogy of an elephant-rabbit stew illustrates the implications of this low value share. If the recipe for such a stew calls for just one rabbit (the energy sector) and one elephant (the rest of the economy), will it not still taste very much like elephant stew? If energy prices had not risen after 1970, it is likely that energy

demands would have grown at about the same rate as the GNP. The 4% ratio would then continue into the future. But what is the effect when energy costs double and there is sufficient time for the economy to adapt? One estimate of the impact may be obtained by assuming a constant recipe. Suppose the rabbit is paid for with part of the stew. Then an additional 4% of the stew (GNP) must be allocated to cover the doubling in the cost of the rabbit (energy). In fact, other recipes are available that call for less rabbit and, therefore, lead to lower costs. Under these assumptions, the first doubling of energy costs would produce, at most, a 4% loss in GNP. Even in 1988, the share of primary energy inputs in the U.S. economy was only 5-6% (although it had risen to about 8% during the early 1980s), implying that the elephant-rabbit stew analogy still applies.

With a more complicated argument, it can be shown that a small decrease in energy supply leads to a decrease in economic output proportional to the value share of energy in the economy. At a 4% value share, a 1% reduction in energy input would produce a 0.04% drop in total output. By this argument, a small percentage change in energy availability produces a considerably smaller percentage impact upon the economy as a whole (Hogan, Manne, 1977).

This simple analysis provides some insight but it suffers from a major defect in failing to represent accurately the flexibility of energy utilization in the economy. The processes for future production and utilization of energy are not fixed immutably. Insulation, efficiency improvements, and changes in the mix of input factors can alter the energy requirements for a fixed level of output. Such substitution possibilities can modify the economic impact of changes in the energy system. Flexibility in energy utilization is a central factor in determining energy-economic feedback, and its treatment varies widely among the many different energy models.

4. THE ROLE OF SUBSTITUTION

The specific processes for energy substitution may be varied and intricate. Therefore, even if it is generally agreed that some substitution is feasible, it may not always be possible to identify the specific technological options available. The morass of detail may be approached gradually by expanding our first simple model and the beginning arguments based on the value share of energy. We explicitly assume now that substitution is possible between energy and nonenergy inputs to the economy. For the purpose of the present discussion, this flexibility can be summarized in economists' terms as the elasticity of substitution. Ignoring the feedback to the economy or other inputs, this parameter is the same as the elasticity of energy demand. It measures the proportional response of energy demand to a change in energy prices. Hence, if the elasticity of demand is -0.3, a 10% increase in energy prices produces a 3% decrease in energy demand.

This concept, the elasticity of substitution, provides a convenient index for summarizing the aggregate behavior of the detailed models (see Figure 1). If we assume that inputs of other factors such as capital and labor are held constant, then the elasticity of substitution virtually determines the feedback effect of the energy sector on the rest of the economy.

The implications of alternate elasticity estimates are shown in Figure 2. It is assumed that a Btu tax is imposed gradually to reduce energy consumption. Such a tax might be levied, for example, to mitigate environmental impacts or lessen import vulnerability (Hogan, Manne, 1977). This figure shows how the elasticity value affects the size of the Btu tax required to cut energy consumption in half in the US by 2010. The higher elasticity the greater the substitution possibilities, and the smaller the required tax.

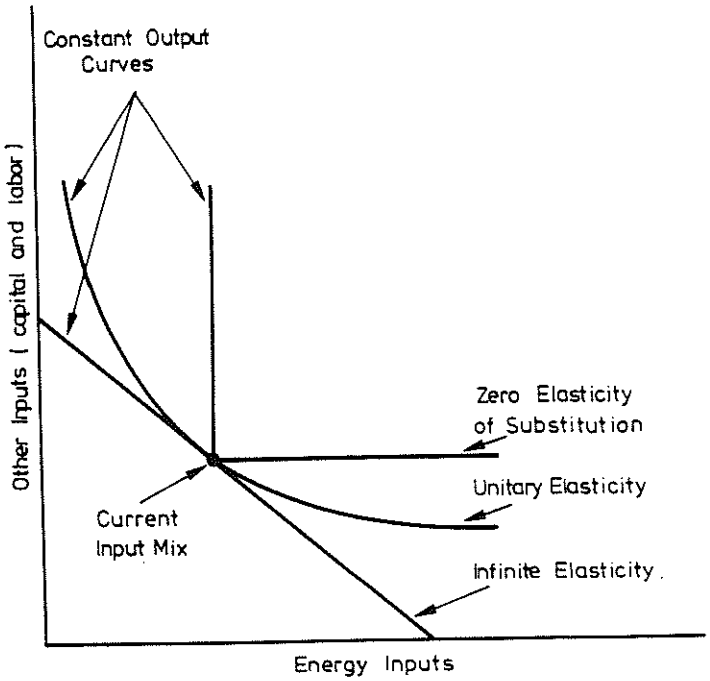


Figure 1. Concept of the elasticity of Substitution

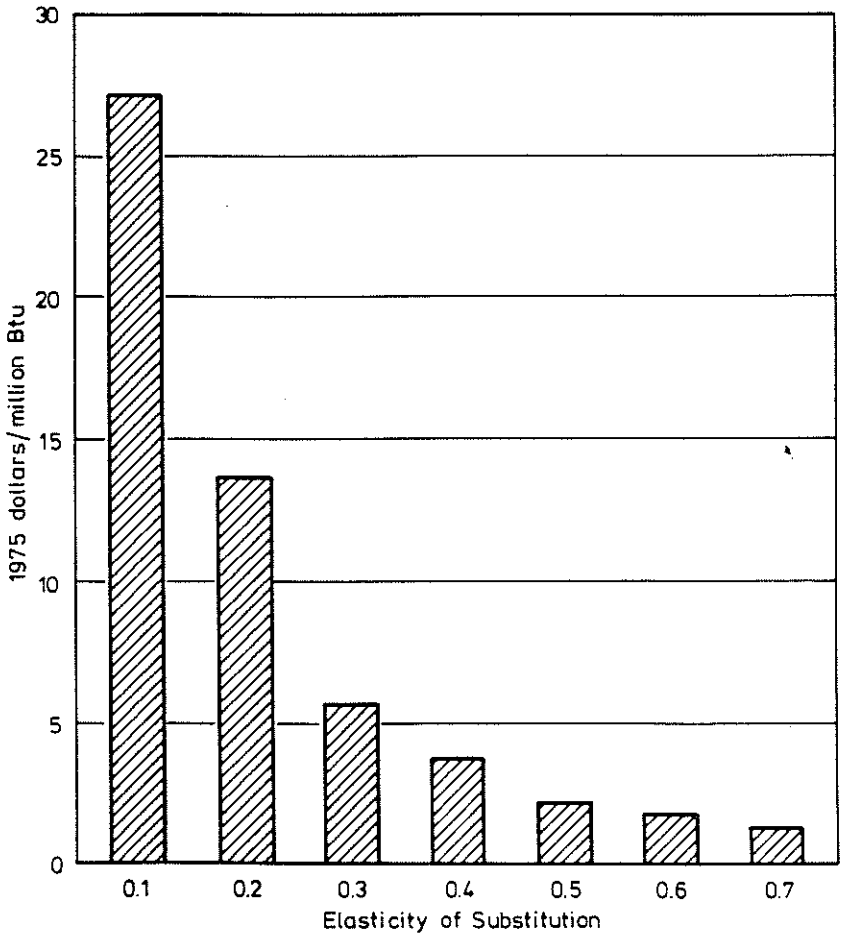


Figure 2. Implicit Tax, Cutting Energy Use in the U.S. in Half by Year 2010. Source: EMF 1

The analysis is based on a hypothetical refèrence forecast, but the qualitative conclusions are not sensitive to this reference. A small change in energy availability has almost no effect on GNP. The loss in output is exactly balanced by the savings from the reduced expense of energy. This is what the price represents, the value of the product at the margin. This value will change as the quantity of energy input changes but the output does not decrease in proportion to the decrease in energy input. Substitution of other input factors compensates for the reduction in energy input.

The importance of the long run elasticity of substitution is startling in the context of this analysis. A 50% reduction in energy availability produces a 28% reduction in GNP if the elasticity is as low as 0.1, but only a 1% reduction in GNP if the elasticity is as high as 0.7. Seemingly small changes in the substitution potential produce major changes in economic impact. Even the smaller GNP reductions have a large value, however. If the economy is growing at 3% in real terms and we discount future consumption at 6%, a 1% reduction in annual GNP corresponds to a present value of nearly half a trillion dollars for the U.S. economy. This is only 1% of the present value of future output, but it would justify a substantial research investment aimed at developing low cost technologies which can expand energy supply or improve the efficiency of energy utilization.

5. THE REAL MODELS

Several immediate difficulties can be found in the aggregate analysis of the preceding section. First, the aggregation itself may disguise distinctly different behavior in component economic sectors. This behavior is of interest in itself and can be captured only by more detailed models. Second, the aggregate substitution parameter does not provide a description of the new processes and technologies that must be adopted. Again, more disaggregated analysis is necessary to provide the detail to support the credibility of the simple analysis. Third, investment, savings, and capital accumulation might be affected by the price and availability of energy providing another reason for the construction of more sophisticated real models.

6. LIMITATIONS IN ENERGY - ECONOMY MODELS.

There are several tools to model the energy-economy interactions (as presented before) starting from small "compact" models to large scale planning models or set of models. In the later case, as usual, the energy system is described by a process model which is driven by a macro-model of the economy.

There are some limitations for either approach. As a rule single "compact" models employ many assumptions about the energy-economic relationship in order to keep the models small. For instance, in some of them the elasticity of substitution between energy and other production factors is exogenous. This assumption can have significant impacts on model outputs. Keeping the model small enables easy communication with the model user who has only limited time to get acquainted with the model documentation. This documentation written by modelers is often incomprehensible to users.

The number of models of this type is increasing due to growing interests in personal computer based modeling. (See Buehring, "Energy and Economic Modeling on the Microcomputer," Chapter 6).

The large-scale models are at the other extreme. Their general goal is to aid in planning. Disaggregated models require a great deal of data and are generally run on mainframe computers. Despite the shortcomings, such as time delays involved in data collection and the expensive model construction process, large models are attractive because they provide an internally consistent framework to analyze the interaction between energy and the economy.

There are significant differences in modeling techniques for market and non-market economies. The former employ the concept of general economic equilibrium as a unifying concept. Even for short-term planning, when temporary disequilibria are represented, the modeling methodology is based on the equilibrium prices as stressed in Chapter "Short and Medium-Term Modeling and Problems of models Linkage."

Partial equilibrium analysis is typically based upon models such as that illustrated in Figure 3. This figure refers to a single energy form--quads of energy--in a single time period. Consumers' willingness to pay is shown as a decreasing function of the amount of energy available to them, and producers' incremental costs are shown as a increasing step-function of the amount to be supplied.

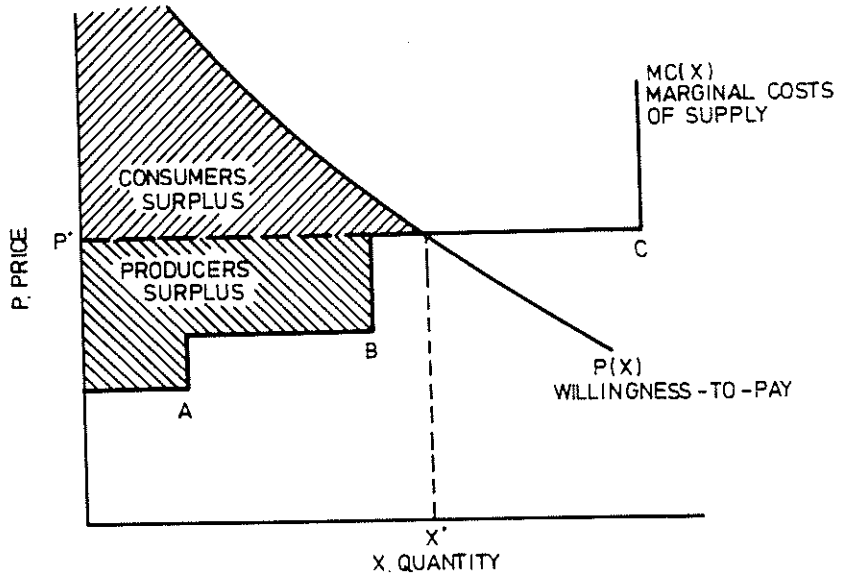


Figure 3. Market Mechanisms and Maximization

Supplies and demands are balanced through an equilibrium price--one that equalizes the consumers' willingness to pay against the producers' incremental costs. It is as if the economy were attempting to maximize the size of the shaded area (net economic benefits) in this figure. This is what leads to the similarity between market mechanisms and optimization models--both for this simple example and also for the more complex cases involving more than one form of energy and more than one time period.

In contrast in non-market economies (both centrally planned and semi-socialized[1] do not have the market prices for all goods. The prices for basic commodities are set by government agencies. As a rule these prices are below market equilibrium in order to support domestic consumption and/or export. In some cases, the prices are fixed for a period of time, while others are related to the international market prices with a time delay.

Given this environment it is necessary to employ a modeling technique that in the absence of market prices for allocating resources, can still provide the information about the scarcity of resource. Hence, the input-output model is the basic approach. It allows the calculation of shadow prices for scarcity values.

[1]Semisocialized economies are hybrid from between free market and socialized economy. SSE are composed from both large government and privately owned sectors. For basic commodities, prices are set by the government.

The analysis is based on a hypothetical reference forecast, but the qualitative conclusions are not sensitive to this reference. A small change in energy availability has almost no effect on GNP. The loss in output is exactly balanced by the savings from the reduced expense of energy. This is what the price represents, the value of the product at the margin. This value will change as the quantity of energy input changes but the output does not decrease in proportion to the decrease in energy input. Substitution of other input factors compensates for the reduction in energy input.

The importance of the long run elasticity of substitution is startling in the context of this analysis. A 50% reduction in energy availability produces a 28% reduction in GNP if the elasticity is as low as 0.1, but only a 1% reduction in GNP if the elasticity is as high as 0.7. Seemingly small changes in the substitution potential produce major changes in economic impact. Even the smaller GNP reductions have a large value, however. If the economy is growing at 3% in real terms and we discount future consumption at 6%, a 1% reduction in annual GNP corresponds to a present value of nearly half a trillion dollars for the U.S. economy. This is only 1% of the present value of future output, but it would justify a substantial research investment aimed at developing low cost technologies which can expand energy supply or improve the efficiency of energy utilization.

5. THE REAL MODELS

Several immediate difficulties can be found in the aggregate analysis of the preceding section. First, the aggregation itself may disguise distinctly different behavior in component economic sectors. This behavior is of interest in itself and can be captured only by more detailed models. Second, the aggregate substitution parameter does not provide a description of the new processes and technologies that must be adopted. Again, more disaggregated analysis is necessary to provide the detail to support the credibility of the simple analysis. Third, investment, savings, and capital accumulation might be affected by the price and availability of energy providing another reason for the construction of more sophisticated real models.

6. LIMITATIONS IN ENERGY - ECONOMY MODELS.

There are several tools to model the energy-economy interactions (as presented before) starting from small "compact" models to large scale planning models or set of models. In the later case, as usual, the energy system is described by a process model which is driven by a macro-model of the economy.

There are some limitations for either approach. As a rule single "compact" models employ many assumptions about the energy-economic relationship in order to keep the models small. For instance, in some of them the elasticity of substitution between energy and other production factors is exogenous. This assumption used to have a significant impact of model output. Keeping the model small enables easy communication with the model user who has only limited time to get acquainted with the model documentation. This documentation written by modelers is often incomprehensible to users.

The way to achieve compactness of models is sometimes open to objection. The number of models of this type is increasing due to growing interests in personal computer based modeling. (See Buehring, "Energy and Economic Modeling on the Microcomputer," Chapter 6).

The large-scale models are at the other extreme. Their general goal is to aid in planning. Disaggregated models require a great deal of data and are generally run on mainframe computers. Despite the shortcomings, such as time delays involved in data collection and the expensive model construction process, large models are attractive because they provide an internally consistent framework to analyze the interaction between energy and the economy.

The PILOT (Planning Investment Levels Over Time) modeling technology (based on dynamic input-output model where flows are in physical units) seems to be appropriate for both market and non-market economies. It was developed for U.S. economy as described by Dantzig et al., in Chapter 3.7. This methodology seems to be useful in studying planned economies in which enterprises are to pursue profit maximization. The PILOT modeling technology has been applied to China by Xia Shi (1987). A similar approach employed in the GNOM model, which has been constructed in Poland; Kuczmowski, Rojewski (1984).

The best known and documented research with sets of formally integrated models has been performed by IIASA. The IIASA researchers make clear the distinction between formal mathematical models and informal procedures (assumption, judgments and manual calculations).

The IIASA model set was designed to operate at a global level (Basile, 1980). For this purpose, the world was divided in seven geopolitical regions and individual scenarios were developed for each region. These were then aggregated to yield global scenarios (Energy in a Finite World, 1981). The set is flexible enough to operate on the national level and was used for national projects.

One should not treat sets of models as giving one and only one correct answer. They should be used rather to back-up the hypotheses of experts and specialists from individual areas of social, economic and energy systems. In such a case the iterations shouldn't be treated in a strictly formal way. In this approach it is rather a process of step - by step adjustments in scenario parameters. A description of this scenario development process might be found in Sassin et al. (1983) and Cofala (1986).

One feedback between the energy system and economy is through capital investments in energy supply technologies. For energy-intensive economies these investments are as high as 30 - 40% of total investments in industry. Such a feedback may be satisfactory treated in a dynamic input-output model (e.g. Cofala, Ostromecki (1988)). Input-output models help to identify and - if needed - to modify main macroeconomic proportions taking into account implications from the energy system. However, a big role in assessing the results is left to the experts because of high uncertainty regarding values for model set parameters.

Because of the problems with the closing of the feed-back loop (in energy-economy sets of models), the tendency recently observed is to diminish the role of formal model in favor of expert judgements. These judgements are enhanced with requirements for capital, materials and other resources, which are calculated by output models. As a case of this type of approach the WELMM code (Water, Energy, Labor, Material, Manpower), which was developed by Grennon and Lapillone (1976) may be cited. Such model may be used to calculate the resource requirements at any level of the economy i.e. national, regional or local. A description of this scenario development process can be a found e.g. in Sassin et al. (1983), Cofala (1986).

Taking into consideration all these experiences, the preferred design for a set of energy economy models might be a linked set of separate energy and economy models. Thus, the crucial problem is adequate linkage of models. This problem was discussed by Capros et al., (Section 3.1). However, informal judgements used in the construction of model sets is essentially inaccessible to external review (von Hippel, 1981).

REFERENCES

Baugham, M.L., and P.L. Joskow (1975). Energy Consumption and Fuel Choice by Residential and Commercial Consumers in the United States, MIT Energy Laboratory.

Basile P. et al. (1980). The IIASA set of Energy Models: Its Design and Application IIASA , Luxenburg.

Energy Vol. 15, pp. I–III, 1990
Pergamon Press plc. Printed in Great Britain

EDITORIAL:
SOFTWARE SURVEY SECTION

The purpose of the Software Survey Section in ENERGY, THE INTERNATIONAL JOURNAL is to encourage the open exchange of information on software programs unique to our professional field. With the rapid penetration of computers into academic and industrial institutions has come a parallel increase in the number of scientists and researchers designing their own software. The existence of much of this software remains unknown even to those of us who could most benefit from its use. We believe that it is of vital importance to our readers that such information be made available. We believe also that a professional journal is the best place to share such information. Your contribution would be most welcome.

The questionnaire on the following pages is designed to assist you in reporting on software that you may have developed or be in the process of developing. By completing this form, your information will reach thousands of your colleagues who may benefit from your work and may possibly offer suggestions for further enhancements to your software. Please complete the enclosed form and return it to:

Professor A. V. Sebald, Software Section Editor
ENERGY, THE INTERNATIONAL JOURNAL
Department of Applied Mechanics and Engineering Sciences
University of California, San Diego
La Jolla, CA 92093

We do not intend to review or comment on the contents of the questionnaire. It will be published as is to expedite the information cycle process. We would welcome any comments you may have.

THE EDITORS

JOURNAL NAME ENERGY

PERGAMON PRESS
SOFTWARE DESCRIPTION FORM

Title of software program: _____

_____

Type of program: [ ] Application [ ] Utility [ ] Other_____

Category: _____ (ie. Psychological assessment, statistics, thermodynamics, etc.)

Developed for (name of computer/s): _____

in (language/s):_____

to run under (operating system): _____

available on: [ ] Floppy disk/diskette. Specify:

Size _____ Density _____ [ ] Single-sided [ ] Dual-sided

[ ] Magnetic tape. Specify:

Size _____ Density _____ Character set _____

Hardware required: _____

Memory required: _____ User training required: [ ] Yes [ ] No

Documentation: [ ] None [ ] Minimal [ ] Self-documenting
[ ]Extensive external documentation

Source code available: [ ] Yes [ ] No

Stage of development: [ ] Design complete [ ] Coding complete
[ ] Fully operational [ ] Collaboration welcomed

Is program in use? [ ] Yes How long? _____ How many sites? _____
[ ] No

Is the contributor available for user inquiries: [ ] Yes [ ] No

Distributed by: _____

Cost of program: _____

Demonstration disk available? [ ] Yes Cost: _____
[ ] No

(continued)

RETURN COMPLETED FORM TO:

Professor A.V. Sebald, Software Section Editor
ENERGY, THE INTERNATIONAL JOURNAL
Department of Applied Mechanics and Engineering Sciences
University of California, San Diego
La Jolla, CA 92093

[This Software Description Form may be photocopied without permission]

Description of what software does [maximum: 200 words]:

Potential users: _____

Field/s of interest: _____

# # # # # # #

Name of contributor: _____

Institution: _____

Address: _____

_____

Telephone number: _____

# # # # # # #

Reference No. [Assigned by Journal Editor] _____

- -

[The information below is not for publication.]

Would you like to have your program:

Reviewed? [ ] Yes [ ] No [ ] Not at this time

Marketed and distributed? [ ] Yes [ ] No [ ] Not at this time

[This Software Description Form may be photocopied without permission]

ENERGY - The International Journal

Guidelines for the Preparation of Typescripts for Direct Reproduction

The Journal now offers authors a medium for rapid publication, without rigid space limitations or page charge. Papers accepted for publication may now be printed by ***direct reproduction*** from the authors typescript. Such manuscripts should be typed in accordance with these guidelines **on the specially prepared laysheets which can be obtained from Pergamon Press offices at Elmsford, USA, or Oxford, UK (addresses overleaf).**

These notes and those on the laysheets are intended to supplement the general guidance on manuscript preparation contained in the Instructions to Authors published inside each issue of the Journal. Authors must continue to follow the normal stylistic requirements of the journal.

The introduction of this method of production will allow the publication of accepted typescripts which are prepared satisfactorily without the need for modification, ***within 3 months*** of their acceptance. Where changes have been recommended by the referees, publication will occur within 3 months of the receipt of the amended typescript. All amendments must be made by the author - neither Pergamon Press nor the Editors can undertake to retype manuscripts.

Manuscripts should be in English and normally be 4 to 12 pages long, with an even number of pages to avoid wasted space. **Authors must follow the printed *Information to Contributors*.**

Instructions for Typing

1. The sample pages which follow show the general layout of a typical typescript which has been typed on a special laysheet and reproduced by photo-offset to 80% of the original size. Any good-quality electric typewriter can be used provided the characters are properly cleaned and a carbon ribbon or new black ribbon is fitted. The sample used a "golf-ball" typewriter such as is produced by all the major manufacturers. Other suitable machines include IBM composers, word processors or Compugraphic typesetters.

2. The use of laysheets will assist in assuring a uniform and consistent format, but good-quality white bond paper can be used provided the layout and dimensions are followed exactly. The typing area of each sheet is 181 mm wide x 287 mm deep. Each page should, where possible, be filled with text and/or figures, leaving minimal waste space.

3. The first page carries the Article title, Authors' names, Addresses and Abstract all of which are centred on the page with a maximum line width of 15 cm.

   The title should be typed in capitals 2 cms from the top of the typing area. Any further lines are double-spaced.

   Leave 4 line spaces between the article title and authors' names. The authors' names should also be in capitals. Any further lines are double-spaced.

   Leave a 2-line space between the authors' names and their addresses and a 6-line space between the addresses and the commencement of the Abstract. The Abstract should be typed in single-spacing.

   The main text commences 2 lines under the Abstract and should be typed single-spaced throughout.

Headings

Authors should restrict headings to 3 levels if possible. The headings should be presented as detailed below and as shown on the specimen page. All headings should be *preceded* by a 2-line space.

(1st level) MAIN HEADINGS
Should be typed in capitals (but not underlined) centred on the page.

(2nd level) Sub-headings
Should be typed in upper and lower case with an initial capital for the first word and proper nouns only. These headings should be ranged left on the solid blue guideline and underlined.

(3rd level) Sub-sub-headings. Should be typed as sub-headings and underlined but text will follow on after a full stop.

Footnotes

Necessary footnotes should be typed single spaced, at the foot of the appropriate page, separated from the main text by a line 7 cm long, starting at the left hand margin.

Equations

Displayed equations should be numbered consecutively throughout the text with equation numbers in parentheses flush with the right hand margin. Extra space should be left above and below the equation, any symbols which cannot be typed should be drawn in carefully in black ink. Lengthy equations should be broken so as to fit within the page width.

Illustrations

Figures should be placed in the text as close as possible to the first reference to them. When preparing artwork for diagrams, authors should bear in mind that *the final camera ready pages will be reduced by 20%,* allowance for this should be made in the size of the lettering and the thickness of lines. Only original half-tones and line drawings should be supplied (not photocopies).

Figure captions should be typed single spaced under the figures and centred between the two outer broken guidelines. The figure number should be given as Fig. 1 etc. and the caption should run on after a full stop.

Tables

Tables should be typed within the type area, leave 2-line space between preceding text and the table heading, which should be centred between the two outer broken guidelines. The table should then be typed and only horizontal lines typed or drawn in above and below the column headings and at the foot of the table.

Where possible, Figures and Tables should be positioned side by side, in order to save space.

References

Should be indicated within the text by consecutive numbers raised above the line and following the punctuation i.e. [1,2] The full reference should be cited in a numbered list at the end of the text in single spacing. Please see the *Information to Contributors* for details.

Corrections and amendments

Corrections to literal errors may be made using correcting tape. If extensive alterations are necessary, these should be typed on a separate sheet and pasted over the errors.

Submission of manuscript

The original and 2 photocopies should be sent to the Editors (see inside front cover of the journal for address). The typed sheets should be carefully packed flat and not folded.

Typing sheets for the Journal can be obtained from the Publisher at either of the following addresses:

Pergamon Press plc
Production Department
Headington Hill Hall
Oxford
OX3 0BW
UK

Telex: 83177 Tel: (0865) 64881

Pergamon Press, Inc.
Production Department
Fairview Park
Elmsford
New York 10523
USA

Telex: 137328 Tel: 010 1 914/592 7700

JGJ 12.87

SPECIMEN PAGE

HIGH-VOLTAGE MULTIPLE-RESISTIVITY DRIFT-REGION LDMOS

JOSE G MENA† and C A T SALAMA

Department of Electrical Engineering, University of Toronto,
Toronto, Ontario M5S 1A4, Canada

Abstract - A new multiple-resistivity drift-region structure, which leads to an improvement of the on-resistance of high voltage LDMOS, devices is proposed. The distinctive feature of the novel structure is the low sensitivity of breakdown voltage to process variations. The important parameters of the structure such as on-resistance, threshold voltage and punch-through voltage in the channel are discussed and related to breakdown voltage. Experimental structures are fabricated and characterized to illustrate the advantages of the structure.

1. INTRODUCTION

High-voltage integrated circuits are becoming important in various display and switching applications. Several technologies have been used to combine high-performance logic circuits with high voltage MOS devices.[1-3] One of these involves the use of a lightly doped (LD) ion-implanted n^- drift region between the active device channel and the n^+ drain contact. The gate electrode extends over this n^- region on top of a thick field oxide, helping to shape the surface fields in the structure. The advantage of such a structure is that breakdown voltage can be controlled by adjusting the ion dose.

2. MULTIPLE RESISTIVITY DRIFT-REGION PRINCIPLE

The cross-section of a multiple-resistivity drift-region metal gate LDMOS (MR-LDMOS) structure is shown in Fig. 1.

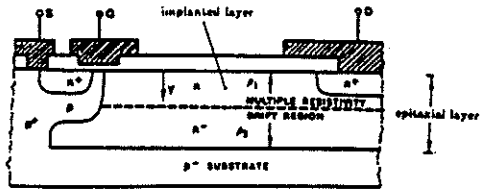


Fig. 1. Multiple resistivity drift region metal gate LDMOS structure (MR-LDMOS) fabricated on n^-/p^- single epitaxial substrate.

The resistivity of the top n layer ρ_1 (obtained by ion implantation) is lower than the resistivity of the bottom layer ρ_2 (obtained by epitaxial deposition); however, the total doping integral across the drift region satisfies the resurf condition.[4] Conformal transformation of the multiple-layer drift region, without neglecting the junction depth of the n^+ drain diffusion, shows that most of the current in the structure flows in the low-resistivity region near the surface, reducing in this way the effective drift-region resistance and thus the on-resistance. The on-resistance reduction helps to decrease the effect of pinch off of the drift region on device transconductance and current-handling capability.

Punch-through voltage in the channel

The critical punch-through breakdown in MR-LDMOS devices occurs in the channel

† To whom all correspondence should be addressed.

under the gate. Although the depletion layer of the channel-drain junction spreads primarily into the more lightly doped drain region, the short channel length raises the possibility of punch through, due to the spreading of the depletion layer into the heavily doped channel region. The punch-through condition neglecting the voltage required to deplete the long epitaxial p⁻-region fully is given by

$$W_D + W_S = L. \qquad (1)$$

Threshold voltage

To evaluate the threshold and punchthrough voltages, the non-uniform channel impurity distribution is approximated by a linear function, as shown in Fig. 3. The linear profile is selected so that the total number of impurities in the channel is conserved.

3. FABRICATION AND EXPERIMENTAL RESULTS

The experimental work described here has two major objectives. The first is to investigate the effects of the multiple resistivity drift region and channel impurity profile on MR-LDMOS characteristics. The second is to verify the accuracy of the analytical models used for the device parameter.

Comparisons between the calculated breakdown voltage, threshold voltage, R_{on} X Z, and the values actually measured for the two series of runs are shown in Table 1. Good agreement is obtained between theory and experiment.

Table 1. Comparison of calculated and experimentally measured MR-LDMOS electrical parameters.

Parameters	Run 1 (optimized)		Run 2	
	Experiment	Theory	Experiment	Theory
BV(V)	460 + 10	440	425 + 5	430
V_r(V)	0.9 + 0.3	1.2	0.8 + 0.2	1.1
R_{on}XZ (n-cm)	17.2 + 0.6	16.4	17.7 + 0.7	16.6

If the device layout is optimized, using for instance a linear geometry with the drain pad placed outside the device active area, R_{on} X Z of the enclosed structure, can be reduced to 8.5 n-cm.[5]

4. CONCLUSION

A new, high-voltage, multiple resistivity drift region LDMOS structure obtained by a combination of epitaxy and ion implantation has been demonstrated. Features of the new structure are: (1) flexibility to minimize the on-resistance for a given breakdown voltage by the proper choice of implant and epitaxial layers, (2) low sensitivity of breakdown voltage and on-resistance to processing parameters variation, and (3) short channel transistors.

Analytical models relating threshold voltage, punch-through voltage in the channel, on-resistance and breakdown voltage of the structure were developed and used to minimize the on-resistance for a 450V breakdown voltage device.

REFERENCES

1. J. D. Plummer and J. D. Meindle, IEEE J Solid-St Circuits SC-11, 809 (1976).
2. H. Sakuma, T. Kuriyama and T. Suzuki, IEDM Dig 594 (1979).
3. M. Pomper, L. Leipold, J. Tihanyi and H. Longo, IEEE J Solid-St Circuits SC-15, 328 (1980).
4. J. A. Appels and H. M. J. Vaes, IEDM Dig 238 (1979).
5. J. G. Mena, PhD thesis, Integrated Circuits Lab., University of Toronto, Toronto, Canada (1985).

· BACK ISSUES OF ·
PERGAMON INTERNATIONAL RESEARCH JOURNALS

Back issues of Pergamon journals are available in hard copy. New subscribers to a journal may purchase back issues of that publication in hard copy edition at 25% discount on the standard price. Stocks of back issues are maintained at all of Pergamon's numerous worldwide offices and orders may be placed with confidence at the location most convenient to you. Should any issue of a volume be temporarily out of stock at the time of ordering, a high quality photoduplicated copy will be supplied at no extra charge to complete your order.

SAVE UP TO 25% BY PURCHASING COMPLETE SETS
Customers wishing to purchase complete sets can do so at a saving of 25% less than the individual volume price.

MICROFORM EDITIONS
Back issues in microform of Pergamon research journals are also available.

BACK ISSUES PRICE LIST
Full details of the rates of back issues of all Pergamon journals can be found in our Back Issues Price List. Please contact your nearest Pergamon office for a copy.

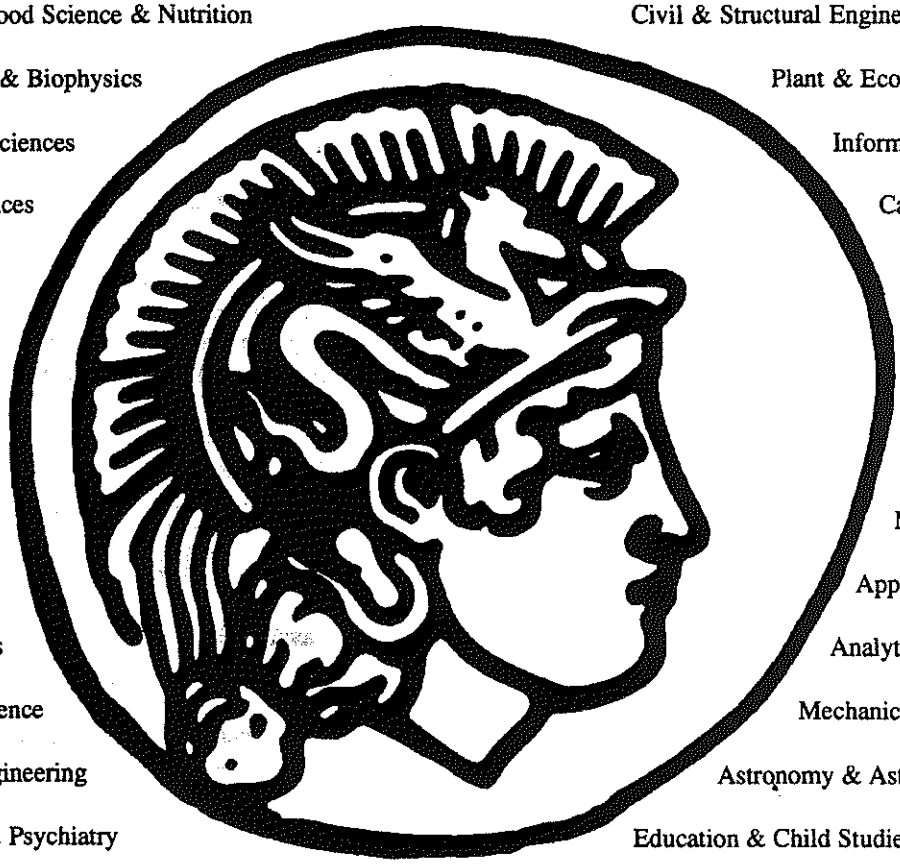
PERGAMON PRESS

Member of Maxwell Macmillan Pergamon Publishing Corporation

Pergamon Press plc, Headington Hill Hall, Oxford OX3 0BW, UK
Pergamon Press Inc., Maxwell House, Fairview Park, Elmsford NY 10523, USA

PERGAMON
SCIENTIFIC PUBLISHING
AT ITS BEST

Clinical & Medical Research
Electrical & Electronic Engineering
Industrial Technology
Agriculture, Food Science & Nutrition
Civil & Structural Engineering
Physiology & Biophysics
Plant & Ecological Science
Biological Sciences
Information Science
Visual Sciences
Cancer Research
Pharmacology
Social Studies
Chemistry
Entomology
Energy
EFL/ELT
Physics
Biochemistry
Mathematics
Metallurgy
Immunology
Applied Chemistry
Neurosciences
Analytical Chemistry
Computer Science
Mechanical Engineering
Chemical Engineering
Astronomy & Astrophysics
Psychology & Psychiatry
Education & Child Studies
Systems Engineering, Automation & Control Engineering
Geosciences & Oceanography
Materials Science & Technology
Economics & Social Planning
Database Publications
Management Sciences & Business Studies
Environmental Science & Water Technology



Pergamon Press plc
Headington Hill Hall
Oxford OX3 0BW
UK

Pergamon Press Inc
Maxwell House
Fairview Park
Elmsford NY 10523 USA

Member of Maxwell Macmillan Pergamon Publishing Corporation

ENERGY—THE INTERNATIONAL JOURNAL

A STATEMENT OF PURPOSE AND GOALS

The all-inclusive description of this journal indicates that it is intended to serve as a multidisciplinary focus for activities relating to the development, assessment, and management of energy-related programs. The Editors are aware both of the unusual demands and unusual opportunities for readers and contributors of this journal: classical disciplinary boundaries cannot be observed in significant discussions on energy issues. We ask the indulgence of our authors in meeting exceptional requirements for clarity and careful documentation for the benefit of readers with widely different specialties. We ask the indulgence of our readers when they are confronted by topical contributions that represent significant advances in esoteric fields of which they will appreciate the importance without necessarily understanding all of the intermediate analyses. It is our hope and expectation that this publication will prove to be an important factor in raising the standards of discussions, analyses, and evaluations relating to energy programs.

Topical coverage should be modified according to subject matter. For example, a detailed experimental study or analysis on the physico-chemical aspects of a particular improved fuel-cell may belong in a specialty journal; on the other hand, an up-to-date account of fuel-cell efficiencies, designs, applications, costs, economic and environmental impacts, large-scale introduction into the energy industries, comparisons with alternative uses for the reactive chemicals, etc. should find a preferred publication outlet in this journal. Input–output analyses relating to energy-consuming systems, careful resource or reserve assessments of all types, energy-conservation measures and their implementations, incisive evaluation of energy-systems managements, environmental-impact assessments, and policy alternatives stressing economic implications are among the topical areas on which important contributions are solicited.

INFORMATION TO CONTRIBUTORS

1. Manuscripts should be sent for publication to Professor S. S. Penner, *Editor*, ENERGY, Energy Center, University of California/San Diego, La Jolla, CA 92093, U.S.A. Submission of a manuscript will be taken as a representation that the manuscript has not been copyrighted, published, or submitted for publication elsewhere. Prior endorsement of the manuscript by one of the Associate Editors will accelerate acceptance and printing.

2. All manuscripts must be submitted in concise, idiomatic English. Manuscripts should be submitted in duplicate, and should be typewritten (except possibly for mathematical expressions), double-spaced with wide margins on substantial white paper. Only one side of each sheet should be used. Contributors are advised to retain an additional copy for themselves as the Editors cannot accept responsibility for damage or loss of papers submitted. (The original manuscript and diagrams will, in any case, be discarded one month after publication unless the publisher is requested to return original material to the author.) Authors are requested to notify the publisher of any change of address which occurs whilst their paper is in the process of publication. Authors should underline only commercial book titles and journal titles and not company reports.

3. **The authors are requested to mark those passages which may be printed in small type.**

4. All articles and Notes should be prefaced by a brief Abstract. Authors should *not* include references in Abstracts.

5. Footnotes should be indicated in the text by the following symbols in order, commencing afresh on each page: † ‡ § ¶ ‖. The footnotes should be typed on the pages where they are referred to.

6. References must be numbered in the order in which they are referred to in the text. Reference numbers *follow* the punctuation as superscript Arabic numerals without parentheses: e.g. Coutant,[1] Hirst and Carney,[2] Tukey.[3] Reference numbers may be used without names. In the text, such abbreviations as Ref. 5 or Refs. 7–10 may be used. Authors may also use on-line displays of reference numbers, e.g., Coutant (Ref. 1), etc. *before* the punctuation.

7. The full references should be listed in the numerical order of citation in the text at the end of the article and should not be given as footnotes. They should include the names of all of the authors with their initials. The full title of a book or the standard abbreviated name of the journal should be given (see: *World List of Scientific Periodicals*, 4th edition, Butterworths, London), followed by the volume number and first page for articles; also, for books the publisher and city of publication should be listed. Book titles should be underlined. For company reports, give the title in quotations and sufficient information (e.g. company and *complete* address) to allow the reader to locate the material. Only *commercial* book titles (not company reports) and journal titles should be underlined. Examples:

1. C. C. Coutant, *Environ. Sci. Tech.* **10,** 868 (1976).
2. E. Hirst and J. Carney, *Science* **199,** 845 (1978).
3. J. W. Tukey, *Exploratory Data Analysis*, Addison–Wesley, Reading, MA (1977).
4. LASL HDR Project Staff, "Hot Dry Rock Geothermal Energy Development Project," *Annual Report*, Fiscal Year 1977, Los Alamos Scientific Laboratory report LA-7109-PR (Feb. 1978), Los Alamos, NM.
5. W. M. Shaub and W. Tsang, in *Chlorinated Dioxins and Dibenzofurans in the Total Environment* II, pp. 469–487, L. H. Keith, C. Rappe, and G. Choudhary eds., Butterworths, Stoneham, MA (1985).

8. Tables and figures should be sent with the paper but should not be affixed in place. The proper places for insertion should, however, be indicated in the text. All photographs, graphs, and diagrams should be referred to as "Fig.", or "Figs." and should be numbered consecutively. The legends should be typed on a separate page at the end of the article and should constitute complete, *self-explanatory* English statements. The author's name, the figure number, and an indication of its proper orientation should be written lightly on the back of each figure. The words Figure or Figures appear in full at the beginning of a sentence and as Fig. or Figs. elsewhere. The words Section or Sections appear in full at the beginning of a sentence and as Sec. and Secs. elsewhere.

9. Photographs will not be accepted.

Line drawings must be submitted in camera-ready form and should be drawn in India ink on stiff white paper or on tracing cloth or vellum and should be approximately twice their final size. The letters and symbols should be drawn so as to be clearly legible when reduced.

10. In the interest of economy and in order to avoid the introduction of errors, tables (as well as figures) will be reproduced by photo-offset means directly from the author's typed manuscripts. The guidelines detailed below should be followed:

1. Original typed tables are required; they must be carefully lined in India ink in a form that is suitable for direct reproduction (i.e. with sparse, carefully set margins).
2. Page size should be 8.5 × 11 in. (or 21 × 30 cm).
3. Large or long tables should be typed on continuing sheets.
4. Table and figure headings should be typed on separate sheets but identifying numbers should be placed on the upper right-hand corner of each sheet of tabular material. The headings should constitute complete, *self-explanatory* English statements.

Please note that oversize (i.e. larger than 8.5 × 11 *in.) figures and tables cannot be accepted.*

11. Particular care should be exercised in identifying mathematical symbols used in equations. When equations extend beyond one line, authors should show appropriate symbols (+, −, ÷, ×, etc.) at the beginning of the next line. Subscripts and superscripts should be set off clearly; Greek symbols should be identified in the margins. All equations should be very carefully set; one-line displays should be used provided they are clear; please use the following sequence of parentheses:)]}. Please use running punctuation with the equations. Equation numbers should appear in parentheses. The words Equation or Equations appear in full at the beginnings of sentences but should be abbreviated to Eq. or Eqs. elsewhere. Zero should be given before a decimal point. Please note that all equation numbers must appear on the right-hand side of the page. Authors should use the abbreviation vs (not vs.) for versus and et al rather than et al..

12. ***Authors should continue to show the original date of receipt (as indicated by the Editor) on the title page of revised manuscripts.***

13. **Corrections to the proofs must be restricted to printer's errors.** Other than these, any substantial alterations may be charged to the author.

14. **Reprints are available if ordered on the order form which will accompany first proofs.**