

## Microeconomic Simulation Models for Public Policy Analysis

A 1978 conference sponsored by  
The Institute for Research on Poverty  
Mathematica Policy Research, Inc.  
The National Science Foundation

This is a volume in the

*Institute for Research on Poverty Monograph Series*

A complete list of titles in this series appears at the end of this volume.

# MICROECONOMIC SIMULATION MODELS FOR PUBLIC POLICY ANALYSIS

## Volume 2

## Sectoral, Regional, and General Equilibrium Models

*Edited by*

**Robert H. Haveman**

Institute for Research on Poverty  
University of Wisconsin—Madison  
Madison, Wisconsin

**Kevin Hollenbeck**

Urban Systems Research and Engineering  
Washington, D. C.



ACADEMIC PRESS

*A Subsidiary of Harcourt Brace Jovanovich, Publishers*

New York London Toronto Sydney San Francisco

This book is one of a series sponsored by the Institute for Research on Poverty of the University of Wisconsin pursuant to the provisions of the Economic Opportunity Act of 1964.

Copyright © 1980 by the Board of Regents of the University of Wisconsin System on behalf of the Institute for Research on Poverty.

All rights reserved.

No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopy, recording, or any information storage and retrieval system, without permission in writing from the publisher.

The views expressed in this book are those of the authors; they do not necessarily represent the official views of the institutions with which the authors are affiliated.

ACADEMIC PRESS, INC.  
111 Fifth Avenue, New York, New York 10003

United Kingdom Edition published by  
ACADEMIC PRESS, INC. (LONDON) LTD.  
24/28 Oval Road, London NW1 7DX

Library of Congress Cataloging in Publication Data  
Main entry under title:

Microeconomic simulation models for public policy  
analysis.

(Institute for Research on Poverty monograph series)

"A 1978 conference, sponsored by the Institute for  
Research on Poverty, Mathematica Policy Research, Inc.,  
the National Science Foundation."

Includes bibliographical references and index.

CONTENTS: v. 1. Distributional impacts.--v. 2.

Sectoral, regional, and general equilibrium models.

1. Policy sciences--Mathematical models--  
Congresses. 2. Microeconomics--Mathematical models--  
Congresses. I. Haveman, Robert H. II. Hollenbeck,  
Kevin. III. Wisconsin. University--Madison. Institute  
for Research on Poverty. IV. Mathematica Policy Research, Inc.  
V. United States. National Science Foundation. VI.  
Series: Wisconsin. University--Madison. Institute for  
Research on Poverty. Monograph series.

H22.M5 338.5'01'51 79-8866  
ISBN 0-12-333202-8 (v. 2)

PRINTED IN THE UNITED STATES OF AMERICA

80 81 82 83 9 8 7 6 5 4 3 2 1

THE JEROME LEVY  
ECONOMIC INSTITUTE

We would like to dedicate this book to David Kershaw (1942-1979),  
founder of Mathematica Policy Research,  
supporter of and participant in this conference,  
colleague and friend.



The Institute for Research on Poverty is a national center for research established at the University of Wisconsin in 1966 by a grant from the Office of Economic Opportunity. Its primary objective is to foster basic, multidisciplinary research into the nature and causes of poverty and means to combat it.

In addition to increasing the basic knowledge from which policies aimed at the elimination of poverty can be shaped, the Institute strives to carry analysis beyond the formulation and testing of fundamental generalizations to the development and assessment of relevant policy alternatives.

The Institute endeavors to bring together scholars of the highest caliber whose primary research efforts are focused on the problem of poverty, the distribution of income, and the analysis and evaluation of social policy, offering staff members wide opportunity for interchange of ideas, maximum freedom for research into basic questions about poverty and social policy, and dissemination of their findings.

**mpr**

Mathematica Policy Research, Inc. (MPR) was founded in 1968 to operate the nation's first large-scale social policy experiment, the New Jersey Negative Income Tax Experiment. Since that time, MPR has expanded considerably and now conducts social policy research, social science experiments, and large-scale evaluation research in the areas of income security and welfare, health, housing, education and training, and microsimulation modeling.

MPR has a staff of over 300 persons who specialize in economic, sociological, and survey research, policy analysis, systems design and implementation, and national sample and specialized survey operations. The company has attracted a professional research and operating staff with a strong commitment to social change through policy research in the public sector. Corporate headquarters and the Research, Survey, and Information Systems divisions are located in Princeton, New Jersey. The Research Division also maintains an office in Madison, Wisconsin. MPR Denver is located in Denver, Colorado, and the Policy Studies Division is located in Washington, D.C.

# 7

## MICRODATA SIMULATION: CURRENT STATUS, PROBLEMS, PROSPECTS

*Kenneth J. Arrow*

The purpose of this conference, which is reflected in all of the papers, is to present and discuss efforts to get information at a much finer level of detail than can be obtained from macroeconomic models. Within this basic unity of purpose, two very different approaches are represented. One is to take the standard idea of complete, preferably dynamic systems and carry the disaggregation much further than has been customary. The other, quite different, approach is to construct synthetic samples of data. I know much more about the first than the second, but I will endeavor to make a contribution to the discussion of each. Before I do so, however, I would like to note that both approaches have a common core which is in line with one of the main traditions in economics—emphasis on the individual decision maker as the unit of analysis.

In the classical, textbook view, the economy is a universe composed of decision-making units. The household and the firm are the typical entities, together with the government. Each unit makes its decisions subject to outside pressures—prices and incomes—which are not completely under its control. It has certain decisions to make within a framework that is limited by these opportunities, and it makes choices which, in turn, have repercussions on the rest of the system. We factor the economic world into these decision-making units plus certain links between them, the markets, about which we have less to say.

I notice in the microsimulation research an emphasis on getting back to the individual unit; each individual unit is examined in terms of what it can do, the decisions it is free to make and not free to make, and the conditions under which it makes those decisions. We are so accustomed to this procedure that it almost seems self-evident; we tend to think that there is no other way to proceed.<sup>1</sup> This is not

<sup>1</sup> This approach is sometimes known as “methodological individualism.” Many who place particular emphasis on the methodological aspect also have strong libertarian beliefs, especially the followers of the Austrian school, though, in fact, virtually all economists tend in the same direction.

logically correct. One can imagine models in which the individual does not appear in the statement of social laws. One sociologist, George Mead (1934), used to argue that there are no individuals as such; that what we call an individual is merely the intersection of the various roles he or she plays in different social groups. And one can, for example, reinterpret Keynesian models as social laws at an aggregate level. We usually think of them as representing individual behavior, but they are stated in aggregate terms.

With all the novelty in the papers presented here, however, no model has been based on principles stated in purely social terms. Rather, in many ways even more stress than is common has been placed on the individual behavior unit, models of the whole economic system being built up from behavior at the level of the individual. The individual decision-making unit appears in the conference papers both as the mover and doer and also as the entity whose behavior we are interested in studying and in predicting from the point of view of, say, welfare implications.

### Disaggregation of Complete Systems

The first approach represented in the conference papers, then, is to try to work down to much finer levels of disaggregation than national income aggregates.

Most macroeconomic models try to get down to thinking, analytically at least, about decision-making units. The major equations—those for investment demand, consumption functions, and the like—sound like behavioral equations, although writ large. But you will always be able to find additional equations that can't be rationalized in terms of individual behavior. They seem to have no other function than to complete the system. In fact, it is hard to go through these models any more, they are so large and complex. The microdata models presented here are more coherent in that respect.

In terms of empirically implementing models of the entire economy, one conceivable line, which is represented in the conference papers, is to take the good old-fashioned general competitive equilibrium theory seriously. It certainly has all the features of emphasizing the decision-making unit. It yields, in principle, all the numbers that are involved as a basis for policy judgment. In fact, if it is carried out completely, it gives you the welfare itself because it has utility functions written in to begin with (one need no longer worry about approximations like consumer surplus).

There have been several empirical applications of general equilibrium theory in recent years. To the best of my knowledge, the first appeared in attempts to work with developing countries. These were still highly aggregated; they were general equilibrium models wherein the word "individual" was not to be taken very literally. I am thinking of such models as those of Adelman and Robinson (1977), and, somewhat earlier, Raduchel (1971). Part of the problem in handling general equilibrium models, of course, is solving simultaneously large numbers of equations—especially accompanied by inequalities, which good general equilibrium theorists tend to emphasize. Recent breakthroughs in computer technology are making this progressively easier. Herbert Scarf (1973) developed the first successful

algorithm, and we have seen today a rich illustration of its possibilities (see Fullerton et al., chapter 3 in this volume).

To digress slightly, I might mention that there have been countrywide planning models using linear programming which could be regarded as general equilibrium models, although they are somewhat degenerate in that there is effectively only one consumer. I remember particularly the large model that Eckaus and Parikh (1968) fitted in India. The time span was five years. The temporal disaggregation was quite fine, to quarters, and there were some thirty sectors. The computation was regarded at the time as reaching the limit of feasible computational complexity. No doubt there has been a tremendous change in that respect in the last decade. A model like that could probably now be taken with equanimity.

The objections to this form of general equilibrium micromodel are, of course, the standard objections to the competitive theory. The question is not, Does the specific model represent the general theory well? It does, and in a reasonably disaggregated form. However, it does not permit testing the validity of competitive theory. (It is somewhat hard, for example, to explain unemployment in a fully competitive model.)

General equilibrium models so far have tended to be static in nature. Dynamic elements can be introduced in theory through developing a sequence of static models in which each step leads into the next. This approach will probably be implemented empirically in due course. When technological advances enable the general equilibrium solution to be computed in ten seconds, for instance, the problems of dynamizing the models are not going to look so formidable.

In dynamic models, since all relations include present and past variables the simultaneous nature of the problem, in a sense, becomes less important. It still exists though, because there are usually a number of contemporaneous variables—except in the limiting case of a dynamic model, where each equation has only one contemporaneous variable. Such a system is completely recursive: When I consider what to consume today, I may take yesterday's income as my budget constraint. The shorter the time spans involved, presumably, the more reasonable this approach becomes. From the viewpoint of appropriate statistical methodology, Herman Wold (1959) was arguing long ago for completely recursive systems in his polemics with the Cowles Commission on simultaneous equations methods. I don't think we need take a dogmatic stand on that question. But we can say that nonsimultaneous models become much more reasonable when very short time periods are considered.

Two of the models presented at the conference (Bennett and Bergmann's model of the U.S. economy and Eliasson's, of the Swedish, chapters 1 and 2 of this volume) represent, in effect, models of the complete economy, highly disaggregated in both sector and time and carried to the limit of essentially one contemporaneous variable per equation. The difficulties of simultaneous solution become trivial in these cases. They are replaced by other complications, however—a large number of variables to handle and a large number of computations to be made at each stage.

Several of the models presented at the conference are sector models, with

everything outside the sector treated exogenously. Methodologically, they are complete models, though not necessarily equilibrium models, of course, in that there may be little or no simultaneity. The model of the health care sector, by Yett et al. (chapter 7 of Volume 1), and Ozanne and Vanski's housing model (chapter 5 in this volume) fall into this category. Here again the basis is disaggregation into individual decision-making units, with market links of some kind.

All the models I have discussed so far form an emerging family of models, contiguous with but more highly developed than the macroeconomic models with which we are more familiar.

### Synthetic Sampling

Let me talk, more briefly, of my understanding of synthetic sampling. As I understand the matter, it is really a question of estimating certain properties of multivariate distributions by observing joint distributions of subsets of variables. To use the simplest case, let us take three variables: Suppose there are some observations in which variables 1 and 2 are measured; there are other observations on variables 2 and 3, and still others on 1 and 3. There are few or no observations on all three variables simultaneously, but nevertheless it is desired to estimate their joint distribution. If the variables have a joint normal distribution, then this information is adequate since the joint distribution is determined by the means and covariances. Instead of actual observations on all three pairs of variables, some other information may be used to complete the description of the joint distribution.

This procedure has been used, for example, in the studies on inequality by Christopher Jencks et al. (1972); they were fitting regression equations, but they never had a sample in which all the variables they wanted to work with appeared. They sought, among other things, to predict income from schooling, parental status, intelligence, and other factors. But they never had all these observations in the same study. Instead, they estimated the sample covariances between income and the independent variables and among the independent variables by using different pieces from different studies.

Among the papers at this conference, Minarik's MERGE file (chapter 1 of Volume 1) is the easiest one for me to understand. A tax law is a mapping from a vector whose elements are the income characteristics of the individual (wage income, dividends, capital gains, and all the other items in the income tax form) to tax liabilities. It is supposed to be a well-defined function; no economic analysis is needed. (Perhaps, realistically, the variables should be extended to include the ability of the taxpayer's attorney.) In fact, to use this information one wants to know the distribution of the burden by some classification of lower dimensionality than that used in the tax law. To go one step further in the simplification, let us suppose that tax liability was completely determined by capital gains and income other than capital gains. Then, the tax law would be a mapping from those two variables to tax liabilities. But if what we really want are tax liabilities by income class, what we need is the conditional distribution of capital gains for any given

income. For a given income, tax liability is a function of capital gains; hence, average tax liability can be computed for each income level.

One of the most powerful applications of this technique was estimation of underreporting; that is, with some kinds of information, it is possible to correct the data. Presumably, hypothetically there is a joint distribution of true income and reported income, and therefore a conditional distribution of true income given reported income; the logic of this correction is the same as that just given. The results are a bit discouraging for those who fit regressions to published data—as in the very large volume of empirical work done on the basis of looking up a few columns in the *Statistical Abstract*. After listening to Minarik's story, one feels very inhibited, for at least a few months!

The joint distributions are not only contemporaneous, they can also be between variables at different points of time. Much emphasis has been placed, in fact, on the dynamic aspects of the simulation where individuals have characteristics that evolve over time according to some stochastic process. Myself today and myself tomorrow are two individuals whose characteristics differ, though their values are related. Some changes are simple; my age next year will be one greater than my age this year, although even that simple proposition has to be adjusted for the probability of survival. More meaningful and complex correlations among characteristics at different times will usually need to be estimated. For example, income transformations over time may be estimatable, from one set of data, such as the Michigan Panel Study of Income Dynamics. One might then estimate an auto-regressive relationship or some other Markov process representation of income changes. Estimates of geographic mobility come from yet a different source. Although the context is now dynamic, the problem of putting together a number of separately estimated joint marginal distributions to find a joint distribution is logically the same as in the static context.

To return to our example of the tax law situation, synthetic sampling can be used to derive tax liabilities as a function of income. But, in fact, nobody is really satisfied to stop with such a relation. Everybody knows some economics; the economists have their professional pride at stake; so they all say, "The tax law is going to change things." The next step is to consider the rather old-fashioned concept of the incidence of the burden. The analyst starts to estimate the shifting of the taxes; and then—particularly if the taxes being shifted are considered significant—somebody brings up the question of implicit taxes: "After all, you've got to predict changes in municipal bond returns; they are going to adjust because of the corporate income tax (because otherwise there would be disequilibrium in the market)."<sup>2</sup>

Once the model builder starts considering the incidence of the initial, legal relation, he or she is moving toward a general equilibrium model, or at least toward

<sup>2</sup>On that particular point, it should be noted that among the alternative incidence assumptions in Pechman and Okner (1974) was the so-called "new view," where the corporate income tax falls on all property. This assumption automatically takes care of the implicit-tax argument.

a complete model if disequilibrium is admitted. The synthetic sample yields one relation—a very complicated relation, in general—but still one that, of itself, becomes an input into the kind of model I discussed in my first section. It is in this sense that the two topics of this conference are complementary. The MATH model (see Beebout, chapter 2 in Volume 1) illustrates the first step in going from a technical relation derived from a synthetic sample to the most immediately adjacent behavioral and feedback relations. Sometimes, especially when talking about an impulse starting in a relatively small sector, it may be perfectly reasonable to say that certain feedback relations are negligible. This principle applies to relatively specialized income maintenance programs, for example. The number of individuals on a particular income maintenance roll who will be employed as a result of the purchasing power generated by an increase in the scale or generosity of the program and therefore removed from the roll may reasonably be neglected. But ignoring feedback may be an important error for relatively large programs.

There is one problem that I haven't heard mentioned with regard to synthetic samples, or indeed in any use of statistical relations to map a distribution into fewer dimensions. This is the fact that the information used is essentially a set of correlations observed in historical data. But if it is a policy change that is being analyzed, it should be admitted that the change may affect the correlations being used. The synthetic sample is drawn from a particular set of conditions, let's say a particular legal structure, and its statistical properties will change with a change in conditions. For example, the correlation between capital gains and income could easily be changed by economic reactions to changes in the tax law.

### Why Micromodels?

There has been an increasing demand for detailed micromodels stemming, in my judgment, from several entirely distinct sources.

The really big impulse has not been scientific curiosity or the demand for better models but the need for a kind of answer that simply didn't come out of the existing macromodels. Input-output analysis was supported at various stages, if not originated, for similar reasons. It was developed to answer the question: What happens to any specific industry? Similarly, the current work on microsimulation is not merely designed to yield better answers for aggregate questions but to yield answers to questions at the microlevel. The form of the models reflects the form of the answers sought. When discussing income maintenance, for instance, we want to know what happens to particular income groups, because the purpose of the whole exercise is to improve the distribution of income—maybe the size distribution of income; maybe the distribution of income by region or by other criteria (which, as an economist, I find less interesting than the size distribution). But, whatever it is, there is some socially compelling reason for a distributive question, the answer to which must be stated in distributive terms. It is no use knowing that a large transfer program presumably doesn't affect aggregates at all to a first approximation and

that, if it does have an effect, it is probably negative. The question is: Is the program being targeted to an appropriate group? The need for answering this question was recognized by the President's Commission on Income Maintenance, back in the late 1960s.

Closely related to the need for disaggregated answers is the fact that appropriate data are now available from various surveys—particularly special-purpose surveys like the Survey of Income and Education, or the Survey of Economic Opportunity—that didn't exist before. The timing of the development of data appropriate for micromodel building is not a coincidence but came from the same perceived need for detailed answers.

A second motive for microsimulation modeling is the belief that the technique results in better models. One reason for this belief is implicit in what has already been said; disaggregated relations are simpler and more transparent than aggregated relations. This may have been one of the reasons, years ago, behind Orcutt's work (chapter 3 in Volume 1). Our aggregated relations conceal distributional assumptions of some kind. Therefore, there is a need for variables that are not decision-making variables but reflect, in some way, the aggregation procedure. If these variables are not introduced, the aggregation implicitly and illegitimately assumes a vast mass of identical individuals. The error is especially striking if the individual relations are nonlinear. Even if the individual relations are linear, there are aggregation problems; each one may be linear but in a variable specific to that individual. Thus, my consumption depends on my income, and somebody else's consumption depends on his income; even if each individual relationship is linear, there is an aggregation problem if the propensities to consume differ among individuals. If the model deals with the individual level, or some relatively fine level of disaggregation, the relationships should be much simpler and our understanding much better than at the aggregate level.

A third motive, less completely accepted, is the view that values of some of the parameters can be deduced directly from the detailed observations. In other words, by disaggregating to small parts of the economy which we understand better, we can get—by direct questionnaires or by relatively simple observations—estimates of parameters whose meanings are murky in the system as a whole. There can be differences of opinion on this question. One problem is that observations in the individual sectors, though easily obtained, may not aggregate well. There are also comparability questions. On the aggregation question, it seems to have been taken as axiomatic that a complete model of the entire economy should be consistent with national income accounting. That may not, however, be right. Since we happen to have a large number of national income statistics, there are obvious manifestations at the individual level, the national concepts as we have them now are not in fact useful.

All these reasons behind the recently increasing demand for detailed micromodels seem to lead to the conclusion that models built up from individual be-

havior should yield better predictions even of aggregates. More information is being poured into the system because it is capable of absorbing individual information that conventional models cannot absorb.

### Methodological Issues

The development of microanalytic simulation models also raises methodological issues, which may not be totally new but certainly appear in new versions. The first one I want to mention is the problem of estimation. As I have already mentioned, one key methodological hope for microanalytic models is that parameters can be estimated "directly," in some sense. But strictly speaking all parameters are parameters of statistical relations, and the general principles of statistical inference should apply. No matter how detailed the observations in a particular sector are, there is still a statistical inference of some kind to arrive at the estimate of a parameter. It is very important, in particular, to know that parameters are estimated with uncertainty. Unfortunately, as far as I can see, in all uses of models for policy purposes (including those at this conference) there is no confidence or error band.

What is done is to predict on the basis of many different alternative policy assumptions (the purpose, of course, of the exercise). Sometimes model builders do get so far as to present the consequences of alternative assumptions about the model. But rarely have I seen statistical theory used to generate a confidence band. And yet the statistical theory itself—the very statistical calculations performed—will provide, as a by-product, confidence bands on forecasts. Now, I suppose somebody will tell me that the fault lies with the consumers; they do not understand error. Of course, they know the analyst is very probably wrong, but they don't want to see a standard error. Cochrane's recent article (1975, p. 203) discussed the famous guidepost formulation in the middle 1960s of 3.2% per annum wage increase. The Council of Economic Advisors was aware that the basis of that figure was shaky. But they said it was impossible to present this uncertainty to President Johnson. To him a "range" was a place where you kept cattle.

It may be a common characteristic of decision makers, of people who actually have to do something, to dislike a reminder of uncertainty. But since uncertainty is real, it does seem that we should recognize and quantify it, however crudely. In some cases, as in input-output models such as that of Haveman et al. (chapter 4 in this volume) and that of Dresch and Updegrove (chapter 6 in this volume) or in the model presented by Fullerton et al. (chapter 3 in this volume), there is only one observation. In those cases it is impossible to get any estimate of variance at all. There is no way of deriving a confidence band from a single measurement. But we know that no model is correct. It couldn't be; there are so many factors omitted, even under ideal circumstances. Estimates based on a single observation, therefore, must also be unreliable, even though we cannot obtain measures of their unreliability. What is needed is replication, repeated observations within a time series or a cross-section context (although the latter has other difficulties). So it has to be

understood that even direct observation should be tested by repeated observations, at several points in time or for several individuals.

There is one example I know of in which system parameters were obtained with very low error by direct questioning. It concerned pricing rules. The idea of replacing maximization models by assuming that economic agents follow simple rules of thumb in matters like pricing is probably very old, but it was certainly pushed hard by Herbert Simon and his disciples at Carnegie-Mellon University. In particular, Richard Cyert and James March (1963) went to a department store and learned from the management what pricing rules were being followed. A department store, assuming it has any monopoly power at all, should be worried about the cross-elasticities of demand between any two of its products and the demand for them. But instead of using a complex set of rules incorporating all this information, it turned out the store classified all items into three categories with the same markup on all items in the same category; for example, the markup on an item in the "standard" category was two-thirds over purchase price (subject to a round-off). There were also rules for markdowns when an item didn't sell. (Remember that these operating rules were found by simply asking the management.) They then priced the commodities, and the fit was fantastic. For 188 out of 197 cases, the error was zero. This is not ordinary economics.

But what is more striking is that a number of years later, William Baumol and Maco Stewart (1971) went to another department store and checked the Cyert-March rules. Of course, they didn't fit perfectly. The markups tended to be higher and less uniform. But in virtually all cases, the markup was between 67% (the Cyert-March rule) and 82%—indeed, 60% of the items had exactly an 82% markup. Now, that's the kind of prediction economists don't find very often. But somehow, nobody paid any attention to these results, as far as I can make out. One research problem, of course, is cost. In this case, for instance, one would have to go to every department store to find the rules. It might be hoped that, after a while, some pattern would emerge. But this line of research has not been much pursued. I think the main reason is that the theory and the observation are not far enough apart to be interesting. These studies do, in any case, exemplify the argument that information about system parameters can be obtained by direct questioning if the system is sufficiently disaggregated. It is hard to know how to assess the parameter values statistically, although the fit is so good that one might not worry about that problem.

An alternative approach to estimating in these models, and one that was suggested by Bennett and Bergmann (chapter 1 of this volume) among others, is to guess the parameters and then fit the model; if the fit is poor, then change the estimates. To some degree, this method is logically not that different from statistical estimation. The statistical properties of the Bennett-Bergmann estimates cannot be derived from statistical theory. They can, however, be derived by simulation. For a given set of parameter values, the model can be simulated with random shocks. By repeating this operation often enough all the desired estimates of statistical reliability, whether confidence intervals or tests of hypotheses, can be obtained. Deriving statistical properties this way is not cheap computationally; but in view of cur-

rent prices, that may not be a very serious obstacle. That the usual apparatus of normal distributions cannot be used merely illustrates the perfectly good argument that so many have referred to today, that you can use computers instead of mathematics. Computers are, after all, a form of mathematical calculation; and, ever since it has been established that four colors are enough for coloring a map, even relatively pure mathematics has been invaded by the computer.

Let me turn to one concern about large models of any kind (and models may get to be much larger than the largest macroeconomic models used today), the danger of rigidity. It becomes harder and harder to see through them.

Suppose a model is not doing very well at prediction or plausibility. It may become difficult to see what is wrong. It is possible to spot a particular relation that isn't fitting very well. But, more often, what is needed for improved prediction is not just putting one new variable in and taking one away but rethinking the model more comprehensively. If what is needed is a change in concepts, for example, definitional changes will be needed everywhere in the model. In effect, scrapping a large intellectual investment is called for, and this gets harder and harder the larger and more complex the model.

In this context, the demands for consistency of data voiced at the conference worry me. It is a very attractive idea. It is frustrating, on the one hand, to find that data from different sources cannot be amalgamated because the definitions are not identical. On the other hand, the analytic community can get locked into a consistency which becomes harder and harder to change. This is the old argument about individualism and *laissez-faire*. There is a loss through lack of coordination; and the argument for coordination, for direct interference, is pretty strong. But the fear that innovation will be inhibited has got to be taken care of.

It might be argued that the supply of innovative ideas will be insured by the fact that nobody can make a reputation except by being different. But it must be kept in mind that these are rather costly new ideas. As I mentioned earlier, the demand for microanalytic simulation models was not generated within the field of scholars but outside. The interest in innovation must be financed by the consumer.

I would like to raise still another question. This conference had as its topic microsimulation for policy analysis. But, of course, someone like myself is interested in the fact that new knowledge should have scientific usefulness. "Scientific" signifies, if you like, the generalizability of the knowledge as a basis for a great many different policy analyses. Perhaps it is also relevant that it be interesting for its own sake. Finally, scientific study explains the policy analysis to theorists, which is important because that's where the next generation of policy research is going to come from anyway. But regardless of how I try to defend my curiosity, I just want to understand the relations better. Now, many of the hypotheses used here are basically familiar, and the emphasis has been rather on their empirical implementation. But, a few of the analytic structures implied somewhat unconventional approaches—most strikingly in the models of the entire economy where markup pricing was used. Markup pricing still represents a challenge to orthodox thinking;

determination of whether a market-equilibrating price is a better or worse fit to the data than a markup price would be most important. Of course, the markup price theory in itself conceals a rather wide variety of possibilities; two distinctly different ones have been presented in the conference papers.

With respect to this point, one trouble is that the large macroeconomic models have, I think, gotten beyond the point of being scientifically useful. They are very important for prediction and policy purposes; but if we want a new idea about a consumption function or about investment demand, the context of a large macroeconomic model is probably a poor environment for developing it. Most of the equations in the macromodels did not originate, in my judgment, from earlier models but from somebody just studying that particular relationship.

Whether or not these remarks are fair, they lead to the questions of validation. How do we validate these relationships?

One possibility is to use the model for forecasting for a different period. This criterion was mentioned in the discussion of the Ozanne and Vanski housing model (chapter 5 of this volume). One of the comments was that conditions had changed entirely, which, unfortunately, is true. The model is complete enough; the factors that have changed so much are in the model—if only as exogenous variables. Presumably the model should have been capable of generating conditional forecasts. One could calibrate the world as of 1970 and take the exogenous variables as they were in 1960. The model would generate a prediction, which might be termed a "hindcast." Hindcasting is an alternative technique which at least has the virtue that it can be done rapidly. You don't have to wait a year for verification. I think some of the macroeconomic models proposed in this conference could be profitably run back in time, to see how well they would have performed. That provides some kind of validation.

Forecasting or hindcasting is a way of validating a whole system. One would like, really, to be able to validate individual relations as well because, if the whole system doesn't work well, it is necessary to know where the repair job is needed. This problem requires a methodological discussion which I have not yet seen.

As a small side remark on the relation between scientific as against policy or predictive uses of models, I wonder about the treatment of consumption. The hypotheses about consumption that have had the most success have been based on dynamic considerations, the permanent income or life-cycle hypotheses. A number of the models deal with processes extended in time. But it seems to me, unless I missed something, that none of the consumption hypotheses used in any of the models had any permanent income component. I wonder why that was the case when, presumably, a dynamic model would enable one to use the permanent income hypothesis more fruitfully than is usually done.

Let me conclude with a few remarks on the policy applications. The presentation of results is, I suppose, to some extent imposed by the user. Nevertheless, some more extensive use of welfare-economic concepts should be made. As was pointed out in discussion, we are interested in two aspects, equity and efficiency. We want



some idea of aggregate benefits and losses, and we want some idea of the distributional aspects. But the equity aspects should take account of all effects, not merely the financial.

Consider, for example, the study of the effects of alternative energy policies. There is certainly going to be some restrictive policy on gasoline—say in the form of a tax on gasoline consumption. To measure the welfare effects, it is not enough merely to look at individuals' taxes. One pattern seems to be first to look at the financial burden; the tax has behavioral effects, in lowering of the consumption of energy, which will partly offset the initial tax. Then the tax burden or the after-tax energy consumption is measured. But that is an underestimate of the welfare loss; it neglects the fact that the individual is forced to reduce his or her consumption. Something parallel has occurred in a number of case studies.

Now, it is true that the alternative of calculating the tax on the initial consumption, not correcting for behavioral effects, would overestimate the welfare loss. The truth lies between the two.

The trouble here is that, instead of looking at equity and efficiency as the economist's measure, there is a tendency to look at budgetary costs. Now, theoretically, the only justification for even considering budgetary costs as such is that they are financed by taxes, which have their own distortionary significance. The financial cost may be a rough way to at least remind the user of the problems. But there is no harm in supplying, in addition to the asked-for information, some measures of benefits and costs calculated along economically more interesting lines. Typically, the informational raw material for the welfare measures consists of calculations already made.

Reference was made, in discussion, to the inherent difficulty introduced by the fact that really big changes are unpredictable. Big shifts in fertility are the classic examples given. Although there are various economic hypotheses about such changes, I do not think they have stood up very well. We do have changes; and, I suppose, detailed consumption categories, rather than just total consumption, are subject to major shifts. Even aggregate consumption has shown unpredicted changes, such as the rise in the savings ratio in the last few years. This shift may be explicable somehow, but I don't think anyone actually predicted it. A cynic might thus ask, "What's the use of talking about policy at all, when the conditions on which you base your forecast can change abruptly?" The answer is that these changes typically are not all that rapid. These models and the policy responses based on them must be thought of as, to some extent, adaptive mechanisms that respond to these changes. The really long-run consequences of a policy may not be that important, because we can hope that remedial steps will be undertaken if in fact the policies being implemented turn out badly.

I heard at the conference the familiar complaint against all quantitative analysis: that the nonmeasurable magnitudes have been omitted; that there is a tendency to exaggerate the importance of measurable as opposed to nonmeasurable magnitudes.

The charge is true. But it is also true that the situation is not in any way im-

proved by not doing measurements. Many examples can be given from, say, the field of water resources—where exactly the same cry is heard all the time—to suggest that, on the whole, the analysis based on the quantitative work is an improvement and does not necessarily lead to biased results. It is of course incumbent upon the analyst to make sure that the relevant factors are accounted for, to the extent that this is possible. And as a matter of fact, the general approach represented at this conference does introduce many factors not brought in before. Those are just the repercussions that one can take account of.

Finally, I noticed severe complaints about the quality of the data, the workmanship, and so forth. This is a very encouraging sign. If one thing is clear in any dynamic branch of scientific activity, it is that the ratio of complaints to accomplishment is roughly a constant. Therefore, you may even use the volume of complaints as some measure of the amount of accomplishment.