Volume Title: Economic Research: Retrospect and Prospect, Volume 1, The Business Cycle Today

Volume Author/Editor: Victor Zarnowitz, ed.

Volume Publisher: NBER

Volume ISBN: 0-87014-251-8

Volume URL: http://www.nber.org/books/zarn72-1

Publication Date: 1972

Chapter Title: Discussion – Afternoon Session

Chapter Author:

Chapter URL: http://www.nber.org/chapters/c4400

Chapter pages in book: (p. 311 - 322)

DISCUSSION—AFTERNOON SESSION

Includes comments by F. Thomas Juster, vice president–research of the National Bureau, who was chairman of this session; Bert Hickman, of Stanford University, and Arthur Okun, of The Brookings Institution, who acted as formal panelists; and Otto Eckstein, of Harvard University, who offered some additional remarks. Again, the oral discussion was recorded and edited with the cooperation of the speakers. The exchange of views during the open discussion period was not recorded.

Introductory Remarks by F. Thomas Juster

The only comment I'd like to add to the content of this morning's discussion refers to what many people have viewed, incorrectly, as the National Bureau's simplistic public rule which says that two quarterly declines in a row in real GNP is a recession. The outcome of applying this rule may depend, among others, on whether you look at GNP from the expenditure or the income side. If you do as I do, and look at it from the income side, real GNP did not turn up in the second quarter because the statistical discrepancy turned around the wrong way, and you have *three* quarterly declines in a row. Now, that's a curiosity; it has no substantive content. It does suggest, I think, that it would be important to have somewhat better numbers than the ones we now have, which is a useful thing to bear in mind generally.

The subject matter of this afternoon's session is quite different from the morning's discussion in that we move from talking about the indicators, which one can view as variables that reflect a consensus about where things are going, to relationships which are more specifically behavioral. Our concern is largely with the performance of econometric models—what they can or cannot tell us about the path business activity is taking. It may not be familiar to all of you that the Bureau as an institution has for some time now gone well beyond our traditional interest in tracking business cycle movements with indicators. We have taken the view that what we can appropriately do in the area of econometric models is not what some would have us do and just build one more, yet another, full-scale model—which we think would probably be a mistake on our part; but essentially to focus on evaluation of the record and the structure of the model, which we think is something that has been greatly underdone in terms of the professional work in this area, and we've been doing this now for several years. We view it as having a considerable amount of potential, and it does bear, I think, on the broad question before us today, which is: What ought to be the scope of National Bureau research in this area in future years?

* * *

Bert Hickman: I want to start with a few remarks about the Zarnowitz paper. The parts that I found the most interesting in this paper concern the econometric model forecasts. Many of the results that Zarnowitz is reporting here and the results reported earlier by Evans, Haitovsky, and Treyz were first presented at a conference jointly sponsored by the Committee on Economic Stability of the Social Science Research Council and the Conference on Income and Wealth last November [1969] at Harvard University. Those papers are presently being edited, and I hope that that volume will be out within six months to nine months from now. It has a very impressive amount of evidence on these issues. Zarnowitz has had only a chance to scratch the surface in this presentation today. I think this is a very useful research activity which is being undertaken. I might also put in that this reminds me of another conference which was recently sponsored by the SSRC Committee on Economic Stability, in April 1967, entitled "Is the Business Cycle Obsolete?" That volume of papers came out with the usual two-year lag; it's been out two months. It was edited by Martin Bronfenbrenner. I don't want to keep you in suspense; we decided the business cycle was not obsolete, particularly if you move toward a concept of something like a Japanese growth-rate cycle; so you may find some parallelism between the proceedings from that earlier conference in 1967 and the proceedings today.

Now, I want to say a few things about the issue of econometric versus judgmental forecasts. It seems to me that on a priori grounds, structural econometric models (or structural models in general) have much more to offer than do other methods of forecasting, in particular, than do any extrapolative forecasts whether they're naive models of the very simple type that Zarnowitz used here or more complex autoregressive models. There are several reasons why I think the econometric models are much more meaningful and interesting. In the first place, it seems to me that, if we are economists, we ought to try to use economic theory in our forecasting, and a structural econometric model is a deliberate attempt to use economic theory in creating hypotheses about behavior. And unless economic theory is a set of empty boxes, we ought to be able to improve on mechanical methods of forecasting by using it in some systematic way. Second, if we have a structural econometric model, it does give us entry points for the use of exogenous information, and these entry points are, I think, very important. Of course the real problem with using an econometric model, or any other kind of model, for forecasting is that structural change does occur. But the structural econometric model should give you some entry points to try to make use of exogenous information as it comes up in ex ante forecasting. And of course the primary applications here are things like adjusting tax-rate parameters and tax functions in the models, if you know that a tax change is coming, or making an adjustment in your external estimate of government spending, if you know that's coming; also, making adjustments for things like strikes that are coming up, and so forth. There are various ways of entering that information into an econometric model. I don't know of any systematic way of handling that information unless you do have some sort of structural model; it tells you what functional relationships *ought* to be affected by those extraneous changes.

Third, the large-scale econometric models in particular include a great deal of sectoral detail. Of course there is an argument about whether big models or small models forecast better, and I'm not going to try to get into that. But I do want to say that it seems to me that sectoral detail is very important for its own sake. If people are interested in sectoral detail, they're not interested simply in how well you can forecast total GNP; they want a breakdown of GNP, they want an industrial breakdown; if possible they want a breakdown of prices, and so forth. Also, if you have a large-scale structural model with many endogenous variables, it gives the person evaluating the forecast the possibility of examining the internal structure of that GNP forecast. This gives some better means of judging whether he thinks it's a good or bad forecast because he can examine the individual assumptions of what's happening internally in the model and need not simply accept or reject the aggregate forecast.

Another important reason for preferring structural models, particularly econometric models, is that they can be used directly for policy analysis. That is, they can be used to ask the question, If you don't like the predicted outcome, what can you do to try to achieve a better outcome? In particular, if they have been built with an eye to policy analysis, they will have included structural instruments of policy which can then be used to answer questions about what would happen if you changed tax rates or what would happen if you changed the rediscount rate, and so forth. So they permit, in other words, not only forecasting but also the analysis of the effects of alternative policies. And finally, a structural econometric model allows you to conduct the analysis of dynamic properties of the economic system represented by that model, and this in itself is a very interesting and relevant topic.

So it seems to me that it is very important for us to continue the work, which has become more prevalent recently, of systematically testing econometric models; among other things, testing their forecasting ability—both ex ante and ex post—and doing our best to improve the specification of models and hence their ability to forecast as we go along. It does seem to me that it is important to continue that work in the econometric area as well as in other areas of forecasting because the inherent promise of the econometric models is so very great. From that point of view, I very much welcome that last part of Victor Zarnowitz's paper, where he talks about plans for future research. Most of these future plans—research activities which he hopes will be undertaken at the National Bureau—deal with the analysis of the properties of econometric models, as well as other methods of forecasting. I think this is important work which was sort of opened up in that conference of last November and which I hope will be continued.

I have a few very brief remarks about the Haitovsky and Wallace paper, partly because it is a very meaty paper and hard to read. I did read it, but it will take a long time to digest. I think this paper illustrates what can be done with these models, the very interesting questions that can be asked with them. It reinforces what I have just been saying about econometric models. I think it is a very important methodological paper. It's incomplete right now, and I am sure they are going to be working further with it, but it is a very interesting pilot study. The authors clearly recognize that they have only dealt with certain particular kinds of discretionary policy rules and that their conclusions therefore relate just to those rules and not to all rules. They also recognize that they have used a very simple welfare function which involves a trade-off only between unemployment and prices. Well, there I have some problem because the policy rules deal with unemployment and prices, whereas the welfare function deals with the rates of growth of GNP and prices. Unemployment drops out of the welfare function, which contains just the growth rates of prices and GNP, whereas unemployment is in the decision function as to the policy. I think I know why this is handled this way, or one reason that it might be handled this way, but it is something that perhaps they should have a chance to talk about.

Now, as compared with previous stochastic simulations I think the major difference introduced by the authors is the following: Previous stochastic simulations have by and large dealt with shocks to the disturbance terms in the individual structural equations. In particular, that was true of all the stochastic simulations done for the conference last November. It was also true of the stochastic simulations done on the Brookings model some years ago, and so those simulations of very large-scale models generally just dealt with random disturbance terms. The earlier study by Irma and Frank Adelman also dealt with shocks to the exogenous variables, and in that sense it is similar to this current study. But the current study is the only one I am personally familiar with which has also dealt with shifts in the multiplicative coefficients in the individual functions. So they are letting really everything vary in these experiments: The exogenous variables are also subject to random variation. There is an awful lot of randomness, and that certainly turns up some interesting results. One of the most interesting ones to me was that, given the fact that they have all these kinds of shocks going on, the mean of the stochastic simulations is not close to the nonstochastic simulation, whereas in the earlier experiments, say, with the Brookings model or with the models at the November conference, where only disturbance terms were shocked, it was still generally true that the mean of the stochastic simulations was close to the nonstochastic simulation, despite the nonlinearities in the models. So that apparently what is causing this to diverge in these models now is that shocks are going on also in the coefficients and in the exogenous variables.

Aaron Gordon had raised a point which I had also wanted to mention briefly; namely, that the procedure used in this simulation implies that you can change tax rates from one quarter to the next. And that is OK, I think, but the point really is whether those tax-rate changes will be viewed as permanent or temporary by persons who perceive them. It is the permanent income question, really. You know the response will be at least potentially different if the tax increase is viewed as a temporary one. Now, this may be partly handled in these models to the extent that, for example, they have distributed lags in the consumption functions, so that a change in disposable income has only a partial effect in the current period. In a sense that is a way of getting at the difference between a permanent and a transitory movement in the tax function. But this raises an important question. Do the estimated parameters in the models, which were not estimated under the assumption of such frequent and temporary tax changes, reflect the structure of the models that would obtain if tax changes were as frequent as quarter by quarter?

Another interesting point the authors noted was that there was an asymmetry in policy between the FRB-MIT model and the Michigan model, namely, the fiscal policy was stronger in the FRB-MIT model, which in itself is interesting, since that model was really constructed to try to emphasize the effects of monetary policy, and the reverse was true in the Michigan model. This may well be due to the different structures of the models, but I also wonder if it is possibly due to the asymmetry of monetary policy itself. At least, it used to be argued that monetary policy is much more efficient as a restraining device on economic activity than as a stimulating device; that is, when you had excess capacity and were operating at a high level of unemployment, increasing the money supply might have a smaller effect than if you were trying to restrict the growth rate of activity. Whether that is the case with these models I don't know. It partly involves a question of what the structures of the models are like and whether they differ in respect to possible asymmetries in monetary response. But it is one possible thing that could account for this asymmetry, it seems to me, because, in the policy simulations that Haitovsky and Wallace present, the Michigan model leads to restrictive policies and the other model to expansionary policies. So that is a difference between them in addition to the difference in their structure.

Arthur Okun: The most constructive and most important contribution of the Haitovsky-Wallace paper in my view is its underlining of the problems of uncertainty, the stochastic elements, in judging policy and models, and applying them to the world. We have been aware for a considerable period of time that the effectiveness of policy is not primarily a matter of a bang for a buck. It isn't a question of how much GNP we get by moving the money supply or changing tax rates a given amount. As the authors point out, if the effects are small, then larger shifts in policy are required. If there were no problem of uncertainty about the effects, larger shifts in policy would be justified, and they would not come through as larger shifts in the economy. Similarly, it is clear that uncertainty and lags interact; in making a decision to take an expansionary action now on the basis of the current economic situation, you are implicitly tying yourself to a forecast that the expansionary effects of that policy will still be appropriate during the period in which it is going to be stimulating the economy.

Thus, we have become aware of the importance of questions about how reliable a policy tool is in generating some extra GNP and about how reliable our forecasts are in guiding us to want some extra GNP. But, until now, I believe we have always assumed that the *average* expected impact of an instrument can be derived from multipliers of a model without worrying about the stochastic element. Haitovsky and Wallace come up with a *dramatic* discovery as applied to the FRB-MIT model, that the average effect expected from a policy move when uncertainty is ignored is just a different world from the average result of the stochastic process, which does take randomness and uncertainty into account. This result comes through so clearly that it dominates everything else in the paper. And that result makes it hard to interpret the findings for the stochastic world.

The stochastic world, particularly in the FRB-MIT simulation, is a disaster area. Uncertainty keeps putting a deflationary bias into the results—and it isn't at all clear why the process of generating uncertainty always pulls *down* on real output. The difference between the stochastic and nonstochastic answers is enormous by 1972. Perhaps the random number generator has a built-in dummy variable for Republican occupancy of the White House. Perhaps the Alvin Hansen of the 1930's is generating the residuals and creating a tremendously stagnationist bias in the stochastic version. Starting with the initial boom conditions of the fourth quarter in 1968, the problem becomes that of fighting deflation rather than inflation in no time flat. Although policy seems to be doing reasonable things and is interpreted through a model which has turned out to be reasonable in other exercises, it just can't make this economy go up.

I am not going to be happy until I know what makes that deflationary result come out of the black box, and the authors don't tell me. Moreover, once the economy is in a disastrous situation, the comparisons between the policies merely reveal ones that are a little less disastrous versus ones that are a little bit more disastrous. Nothing is any good. I can't take any satisfaction in the less disastrous results that emerge for some fiscal policy instruments. Some of the conclusions on the monetary instruments need modification. If for some reason a given growth of unborrowed reserves is translated into a smaller growth of the money supply, then a good monetary rule would simply raise the growth target of unborrowed reserves. Before Milton Friedman or some other monetarist tears this test apart, let me go on record saying that I don't think it is a fair test.

It is apparently a property of the FRB-MIT model that a growth rate of the money supply of 4 per cent a year is inadequate to produce healthy growth. Thus, a 4 per cent rule in itself would produce a deflationary bias, although that would apply to the nonstochastic as well as the stochastic version. In his research for the Brookings Papers on Economic Activity, William Poole tried to get the economy on a full employment path by 1973, and estimated the monetary growth requirements with a given fiscal policy. It took a thumping 8 per cent growth rate of M_1 to do the job. Since money GNP had to rise at a rate of even more than 8 per cent, it shouldn't be surprising that it took an 8 per cent growth of M_1, since there is no secular uptrend in velocity in the FRB-MIT model.

Haitovsky and Wallace raise many other interesting issues I could discuss. I do feel, however, that their technique tends to raise too many issues at once in a way that makes it hard to sort out the key factors in operation. For example, if they took a single move of policy and traced out its implications compared with the nondiscretionary model through sixteen quarters, that would give a valuable benchmark of the size and time-shape of the effects. The results of implementing this rather complicated dynamic set of rules for decision-making tell us whether the package works as a package, but does not reveal much about the pieces.

Let me turn to the Zarnowitz paper. I have just a few comments on the earlier part. One, in looking at the CEA [Council of Economic Advisers] comparison table—which I have an historical attachment to—I was particularly struck by the fact that the quantitative errors recorded year by year really don't always match my ex post feeling about the adequacy of the various forecasts. The numbers in the table say that the forecast for 1964 was great while that for 1962 was poor, and my ex post feelings agree; but 1969 was a very good year and 1967 only a little bit better than an average year according to the table. I feel that 1969 was a poor year in terms of the policy implications of the forecast, largely because the 1969 expectation was for slower growth during the first half, and a speedup in the second half, and we got just the reverse. The profile during the year was quite different from what was contemplated at the beginning. Also, the real and price parts of the forecast were off the track, but balanced out. On the other hand, 1967 was a jewel in terms of its policy implications. It really did catch the profile of the year: a flat first half and the danger of an upsurge in the second half. Although the GNP number did not come out exactly on the nose, the accuracy of the profile dominated the slightly larger error in the annual forecast.

I am not suggesting that Zarnowitz can do anything to adjust for such matters, particularly since the data come to him in annual form. But the illustrations suggest, more generally, that a full evaluation of forecasts has to ask other questions: What is the forecast used for? Is it doing the job it is supposed to do? And sometimes the magnitude of the GNP error will not be a good indication.

Like Bert Hickman, I was most interested in the part that Zarnowitz did not get time to summarize, namely: Where do we go from here? I'm concerned about the possibility of answering the question: Do econometric models do well? The sample of econometric models remains small, even with the additional ones that Zarnowitz is planning to include in the future. And they are special in ways other than being econometric. Most of them were developed by very proficient, expert economists, who were able to finance a large investment in a forecasting system. Novices in the field don't get that opportunity. Whether econometric models work better may be a little like the question of whether baseball players from Oklahoma have better batting averages than all baseball players as a group. I don't find the latter question very interesting, and I don't see any use to the answer. Similarly, I have no particular reason to believe that econometric models will be better or worse as a group than equally serious judgmental forecasts. And I would conjecture that forecasting differences *among* econometric models will be as large as the differences between them and judgmental forecasts. I doubt that the forecasts that come off a computer and are very formalized are going to have anything in common. Some of the reasons for my skepticism are illustrated by the wonderful paradox cited in the paper: If perfect foresight on exogenous variables is plugged, ex post, into some of the econometric models, they come out worse than they actually did ex ante without that foresight. I'd throw another possible explanation for the paradox into the hat: the possibility that an econometrician who feels a little more bullish than his model may let that bullishness show up in higher projections of exogenous variables. And when he feels more bearish than his model, he may unconsciously hold down the exogenous variables for the next several quarters rather than make a specific residual adjustment. I wonder how the model builders feel about this paradox. It is a great personal tribute to their ability to make adjustments. But it is capable of an interpretation they wouldn't like: "If this smart fellow wants a toy to play with, it's all right because he's smart enough to correct it anyway." I don't believe that that is the proper interpretation. I think that the model is a tool for developing forecasting techniques, and it is a constructive one, provided it is not used as a substitute for good judgment and provided it is not so structural that it can't absorb barometric and anticipatory data that don't fit into nice causal relationships.

I don't know whether I may be disagreeing with Bert Hickman in that last proviso. I worry about forecasting and analytical objectives getting confused, and thereby compromising the effort to forecast by insisting on structural explanations and nice causal relationships which exclude indexes and surveys. The barometers help forecast, and if we don't take advantage of them, we pay a considerable price in our forecasting ability.

In looking ahead to future research, I would underline the set of questions in Zarnowitz's paper, that ask the why's about the mistakes in the models. The models do provide, as Bert Hickman noted, a unique opportunity to trace the errors and nail them down and determine why mistakes were made. That should help reveal the relationships we're most weak on, what we need to know, whether particular relationships in particular models are outstandingly good (quite apart from whether the model does well as a whole). What kinds of situations seem to lead *all* economists astray? There is a high payoff in focusing on some key surprises; for example, why plant and equipment turned up so strongly at the end of 1968; why the consumer keeps changing his mind; why prices accelerated in 1969. What do these surprises tell us about the world?

It seems to me that there are opportunities to explore the why's of the results on the judgment forecasts in the ASA [American Statistical Association] survey. Could you get a "postmortem" reinterview with some of the people who answer the questionnaires? Give it back to them after a year and say: "Here's what you were predicting. Here's where the economy came out. How do you feel about that forecast? Did it meet your needs? How would you grade yourself? Where do you feel you went wrong? What have you learned from this experience that might help you in the future? Where does the result point to in your own research?"

We do want to be sure that we're not just collecting batting averages of baseball players arranged by state. It's fine to collect averages of left-handed hitters against left-handed and right-handed pitchers; that might suggest whether platooning is a good idea. Similarly, the research should help guide decisions on how to improve forecasting. The work that Zarnowitz has done already has answered several important questions. It demonstrates beyond any shadow of doubt that economists are better forecasters than any naive models, and yet that we have a long way to go before we achieve the predictive accuracy we'd like.

Incidentally, forecasting accuracy interacts with policy issues. Suppose our ability to forecast improved greatly and was translated into economic policy, because the politicians finally accepted our forecasts and advice. In that world, presumably, economic activity would move pretty smoothly over time, and extrapolative methods of forecasting would do very well, just as well as professional economic forecasts. So, in some ultimate sense, the economic forecaster can't win both games of being right and of convincing the politicians to take his advice. In fact, the time the economic forecaster will look best is when the profession is very smart and the politicians are very stubborn. If there were a tremendous gap between what we know policy ought to be doing and what it is doing, then we could confidently predict the major fluctuations. Maybe that would maximize the economic welfare of the profession but it would be sad for the nation.

* * *

Otto Eckstein: Let me try to clarify the puzzle about ex ante and ex post forecasting. There are three essential elements in all forecasting of this type: (1) the information that is used; (2) the extrapolation methods based on past relationships; and (3) judgments about policies and other key elements.

The models contain a lot of information, but they don't reflect all available, useful information. If you run a model in the ex post method, you use only that information which happens to be for variables incorporated in the model and throw away all the rest. The use of constant adjustments or add factors is a method to bring other known information, including leading indicators, expectational swings, industry information, and many other things into the forecasting process.

Extrapolation, which every forecaster employs, is essentially a process of taking past relationships, some primitive, some sophisticated, and projecting them into the future. The computers do that much better than the pencil. Models assure consistency and accuracy.

The third element is judgment. The human brain filters information and analyzes it; the informal forecaster has an informal model in his head. The extent of judgment used depends very much on the personality, interest, and ability of the forecaster. Whether models are used or not has little impact on the extent of judgment exercised.

The superiority of ex ante forecasts is explained, I think, essentially in this way. Outside information and judgment help correct for errors that would otherwise occur in the model. You don't really start out on the next quarter (or the current quarter, which usually is forecast) knowing nothing about the error terms. The first error term in large part is observed, and even the second and third error terms are observed in part, and can be corrected.

Let me make another point regarding the reported FRB-MIT model simulations. The FRB-MIT model is a delicate race horse, more sensitive than the Michigan or Wharton models. Its intricate financial structure and the role of the stock market heighten the possibility of unstable model runs. In the hands of an experienced forecaster, the FRB-MIT model would not be allowed to explode. In the hands of amateurs, the model produces results that are not totally serious. The enormous stochastic variation in the reported runs shows nothing about the economy and only a little about the model. It only proves something about the interactions of the model and the men using it on that particular occasion.