DISCUSSION

Lindsey Churchill, Russell Sage Foundation

Introduction

I came to the field of computer simulation very recently, in the last two years. I am a sociologist by training, and as you undoubtedly know, sociology is just beginning to use computer simulation in its problems. Therefore, I am a novice in this area. However, I have had a good deal of experience with mathematical models and statistics in sociology, and I discovered that if I thought of a simulation man as a mathematical modeler with a powerful friend, the computer, I didn't go far wrong. I found that I could bring an evaluation scheme that I had used on mathematical models over to the study of computer simulation, and I will evaluate the present papers in terms of it.

My scheme consists of evaluating a computer simulation in three areas:

a. Degree of understanding of the substantive problem;

b. Amount and sophistication of the statistical testing of the simulation; and

c. Amount and sophistication of the mathematics used in the simulation. (This scheme is not original with me, I am sure.)

I turn first to the paper by Manser, Naylor, and Wertz[1], then to Schulz' paper [2], and finally to a comparison of the two. In advance, let me say that I like both papers. I will make a number of critical comments, but I intend them only to place the papers in perspective. I am sure that the authors are aware of my criticisms, but simply couldn't do everything at once.

Evaluation of Manser, Naylor, and Wertz

The first question to ask is this: Exactly what are the authors trying to do in this paper? They are trying to invent a strategy by which Federal policy makers can disburse aid to education at the State level so as to "stimulate" the States to pay out the greatest sums possible for education. That is, they want to find a way to manipulate the purse strings for education in each State. In particular, they would like to prevent States from cutting down on their own expenditures by using Federal funds to make up the difference.

How do they proceed to find such a strategy? Very sensibly, they start from the point of view of the State budget-maker for education. They argue that he takes three major variables into consideration: the size of last year's education budget; the sizes of last year's education budgets in neighboring States; and the size of last year's Federal grant to the State for education. Further, the same three factors are assumed to be the central ones in each State.

They show that a linear weighting of the three variables obtained through a regression method rather accurately reproduces the budgets for each State for the years 1951 through 1965. Since the regression function is a good predictor of budget size, the authors argue that they can study the effects of possible Federal strategies in giving education money by working with the function rather than by trying the strategies out in practice.

The authors then try out six different strategies that Federal people might use to induce State budget-makers to increase their education budgets. They project budget sizes for each strategy for the years 1966 through 1972, and use their statistic Z as the criterion for the relative success of the strategies. Their conclusion is that the two strategies that penalize States for using Federal money in place of State money and that reward States for using Federal money in addition to State money maximize the total amount spent collectively by the States for education. (The penalty used was a smaller amount of Federal aid than usual, while the reward was a larger amount of Federal aid than usual.)

Somehow, I don't find this result surprising. I can readily understand that State budget-makers would be responsive to a "carrot-and-stick" strategy. On the other hand, I am surprised that the authors found this classic strategy with their method. I don't see how the "static" model underlying the authors' regression scheme allows a "dynamic" conclusion like this one.

Exactly how do the authors reach this conclusion? I think they make an inferential leap from their statistics to a theory of budget-making that is unwarranted. (Actually, the authors agree about the inferential leap. They say that their results only suggest the conclusion. However, the leap is vital to their paper since they are after a strategy to loosen State strings, and it is important to see how they did it.)

Using their regression method, they found the best set of coefficients to typify each State over the entire period 1951-1965. They interpreted the coefficients c^i of the Federal aid variable as indicators of the "stimulation" or "constriction" effect of Federal aid. Coefficients less than one indicate the "constriction" effect, or substitution of Federal aid for State money; coefficients greater than one indicate the "stimulation" effect. Thus, States are fixed in advance as "stimulated" or "constricted" by the nature of the regression method; it only remains to find out which state a State is in.

It is not obvious that a "carrot-and-stick" strategy should emerge as superior because there

is no provision for the States to react to the strategies. The regression method here does not contain a simulated way by which States can respond to Federal policy over a period of time. Why, then, do Plans 3 and 4, the "carrot-and-stick" strategies, emerge as the best? The answer seems to be that the arithmetic of the calculation of the statistic Z just comes out that way.

The statistic Z can be expressed as follows:

$$\Sigma\Sigma S_t^i P_t^i \;/\; \Sigma\Sigma G_t^i P_t^i \;=$$

$$\Sigma\Sigma [a^i S_{t-1}^i + b^i N_{t-1}^i + c^i G_{t-1}^i] P_t^i \;/ \Sigma\Sigma G_t^i P_t^i \quad .$$

In this expression the S^i , N^i , and P^i are all positive and the same for each of the six strategies being compared. The a^i and b^i are almost all positive (only eight small negative values among 96) and the same for all strategies. The c^i are the same for all strategies, but many of them are highly negative. The only varying quantities in the different strategies are the G^i , the amounts of Federal aid to the different States.

The "carrot-and-stick" strategies allow G^i to be smaller than in the other plans if the corresponding c^i are less than one (in Plan 3) or less than zero (in Plan 4). When the c^i are negative, the third term in the numerator of Z is negative and reduces the size of the numerator, tending to reduce the size of Z. But Z will be larger the smaller the amount of reduction due to this term. And the third term will tend to be least negative in the two "carrot-and-stick" strategies because the minimum G^i's are combining with the negative c^i's. Further, the minimum G^i's will make the denominator of Z a minimum, thereby contributing to making Z a maximum for these strategies.

How, then, do the authors get from this arithmetic result to the inference that a "carrot-and-stick" strategy is superior to other strategies? I think the authors may have proceeded in the following way. They discovered that the multiple R's for the regression equations were very high (the median R is 0.96). Having made this discovery, the authors turned the regression function into a <u>theory</u> of how budget-making gets done. There were many different theories that would be consistent with the regression functions, but the one that the authors implicitly chose seemed to be the following one.

In making this year's budget a State budget-maker first adds in an amount equal to the size of last year's budget. He (ordinarily) adds an increment to that. He then (ordinarily) adds a second increment to the amount so far cumulated, based on the sizes of last year's budgets in neighboring States. Then he looks at the amount of Federal aid to his State for last year. He expects to get the same amount of Federal aid this year, or possibly more. Now he is at a critical choice point: he can either add in a third amount to his cumulating total that is larger than the amount of last year's Federal aid; or he can add in a third amount smaller than that sum; or he can subtract an amount based on that sum of Federal money. If he takes the first option, he is said to be "stimulated" by the Federal money, because he adds in all the anticipated Federal money in this year's budget and more besides, which must come from the State. If he takes the second or third option, he is said to be "constricted" by the Federal money, for the following reasons. In the second option he adds in a smaller amount of anticipated Federal money than he expects to receive, thereby expecting to have some Federal aid left over to substitute for State money already budgeted. In the third option he actually reduces the size of the total cumulated on the basis of the first two variables, but since he expects to get the Federal aid anyway, all of it can be substituted for State money that would otherwise have to be spent.

While it is true, of course, that budget-making gets done somehow, the authors present no evidence to show that it gets done in the way outlined above. Until that evidence is collected, the authors' conclusion must remain an unwarranted inference. The most that their analysis allows them to say is that the best-fitted linear combination of the three variables has a very high correlation with the size of the education budget.

Suppose the authors could present evidence that supported the theory discussed above. A simulation could then be made of the theory. The six strategies could be compared much as the authors do here, and it is entirely possible that the "carrot-and-stick" strategies would emerge as superior to the others. The conclusion that the authors would like to draw would have more weight than it does at present because it would be based on some kind of theory of budget-making. The accuracy of the conclusion would rest on the degree to which the theory of budget-making captured the actual budget-making process.

Thus, in terms of the evaluation scheme I outlined earlier, the authors have made some progress in the understanding of the substantive problem, but not a great deal. The high multiple R's show that the three variables selected are <u>somehow</u> related to the process of budget-making, but that is as much as can be safely inferred. Further understanding rests on evidence collected by other methods of study than the present one, methods that try to find out directly how budget-makers proceed and why.

But a deeper level of understanding may not be necessary to provide a good <u>practical</u> answer to the question that the authors wish to answer: How should the Federal government spend money for education so as to stimulate the states to

spend the greatest amount? I think the authors' "carrot-and-stick" proposal is a good one to try, even though their evidence doesn't support it.

In terms of amount and sophistication of statistical testing, the authors do very well. They are aware of the problems involved in doing regression on time series. They use a sophisticated method for constructing simultaneous confidence intervals to compare the different plans. Best of all, they replicate the simulation of each proposed plan thirty times, something not done too often in simulation work.

Finally, in terms of mathematics, I find very little mathematics done. The paper rests entirely on known statistical techniques. There is some minor mathematics involved in adapting the regression function to the different strategies for Federal spending, but none at all on the substantive issues of how budget-makers make budgets.

In summary, I think the authors have convincingly demonstrated the relevance of their three variables in the budget-making process. Beyond that, I am not convinced by their analysis that the "carrot-and-stick" strategy is the best one for Federal people to adopt, yet on commonsense grounds I find it a perfectly plausible strategy.

I turn now to Schulz' paper.

Evaluation of Schulz

As has been mentioned, Dr. Schulz was originally scheduled to be a discussant in this section. At the last minute he was pressed into service to give a paper. The paper he just gave is an outgrowth of research described in his monograph, The Economic Status of the Retired Aged in 1980: Simulation Projections.[3/] My discussion is based on that monograph since I did not have a copy of his paper in advance.

Let me begin by asking the question: Exactly what did Dr. Schulz try to do in this research? He answers in the following way:

> "Specifically, the study projects the pension income and assets of retired persons in 1980 and investigates the role of private pensions and the U.S. social insurance system in providing retirement income." (p. 1)

The problem is similar in form to the problem pursued by Manser et al. Dr. Schulz seeks to estimate the income of a retired person in the future, while Manser et al. seek to estimate a State's expenditure for education in the future.

How does Dr. Schulz proceed? He proceeds by carefully delineating the major features of American society that affect the size of a retired person's income. These include age, sex, race, marital status, present income, employment status, and private pension plan status, among others.

The projections to 1980 were carried out by selecting all married couples where the husband was between 45 and 60, and all unmarried persons between 45 and 60, from the 1/1,000 Census sample of 1960. These persons were then "aged" into the future year by year on the basis of the simulation model.

A variety of things can happen to a person in a twenty-year period, and Dr. Schulz had to take the most important things affecting income into account. Persons can die in the given time period; they can leave the work force, re-enter it, leave it again, etc.; they can change jobs; they can retire early; they may or may not be covered by a private pension system; they may accumulate financial assets, such as equity in a home; they spend varying lengths of time in each job; etc.

The mechanism of the "aging" process is stochastic. The probability of death in a given year is estimated, as are the probabilities of changing jobs, leaving the work force, being covered by a private pension plan, etc. Then the probabilities (not necessarily assumed constant for each year) are applied to the persons in the 1/1,000 sample for each year from 1961 to 1980. Persons who "die" drop out; persons who "live" are surveyed in 1980. The incomes and assets of persons fully "retired" by the simulation by 1980 are then tabulated to form the basic output of the simulation.

Essentially then, each person selected from the 1/1,000 sample is assigned a history for the period 1961-1980 so that his income and assets may be estimated for the year 1980. The great power of simulation can then be put into play by rerunning the simulation under different assumptions about social security benefits, about the degree of vesting in private pension programs, about the rate of early retirement, etc., to see what would happen to the income distribution if these assumptions were put into practice before 1980.

The general conclusion reached by Dr. Schulz is that while there will be improvement in the retirement income picture by 1980, fifteen percent of retired couples would have total money income (pension plus asset income) less than $2,000, and 35% would have total money income less than $3,000. Analogously, seventeen percent of retired unmarried persons would have total money incomes less than $2,000 and 41% would have total money income less than $3,000. I find this conclusion disheartening.

In terms of understanding the substantive problem, I think Dr. Schulz has done an excellent job. It would have been easy for him to reduce the elaboration of his simulation by simply ignoring many refinements. Dr. Schulz created a much larger number of different types of future history than many others would have, and I

commend him for his painstaking efforts to be accurate.

In terms of statistical testing, however, Dr. Schulz is much sketchier. He made only one simulation run under a given set of assumptions, and he made only eight simulation runs in all. His results would be more convincing if they were averages from many identical simulation runs.

Finally, in terms of my evaluation scheme, Dr. Schulz uses mathematics very little, mainly to estimate changing probabilities as the years move from 1961 to 1980. He avoids the use of extensive mathematics by baldly assuming that persons' future histories are independent of their past histories. He doesn't worry about whether or not he gets the right history attached to the right person; he only tries to ensure that the right numbers of events in different categories will occur.

In summary, I think that Dr. Schulz has made a convincing case that the income for retired persons in 1980 will be better than it is now, but that large numbers of persons will still be below the present poverty line. The conclusion would be more convincing if the simulation runs had been replicated a number of times.

I turn now to a comparison of the two papers and some general comments.

Comparison of the two papers

In terms of my evaluation scheme Manser et al. and Schulz complement each other in the first two areas. Manser et al. are strong on statistical testing and weaker on substantive understanding. Schulz is just the opposite, being strong on substantive understanding and weaker on statistical testing. On balance, I think I prefer Dr. Schulz' use of the available time. I prefer to spend more time on understanding the substantive problem than on statistical tests of whatever understanding I have, but that is purely a personal preference. I fully recognize the importance of statistical testing.

However, it is the comparison in the third of my areas, the use of mathematics, that leads to what I think is the heart of present difficulties with many simulation studies. Neither paper uses mathematics in more than a minor way to deal with substantive issues. I take the relative lack of mathematics in the present studies to indicate that we do not yet have the requisite substantive understanding necessary to make thoroughly convincing simulations. For example, it seems to me that a strong theory of budget-making, in Manser et al.'s case, or of accumulating personal savings, in Schulz' case, would consist of mathematizations of psychological, sociological, and economic findings about these problems. Looking ahead, I expect that future simulation studies on the present problems will "reek" with mathematics.

Dr. Schulz' paper is instructive in this regard because he spent a great deal of time laying out the substantive issues in the income of retired persons. But at a critical point he brought in the most simplistic of mathematical assumptions, statistical independence, by randomly assigning future histories to persons. (I say "simplistic" only in reference to a strong theory of how persons' futures are tied to their pasts.) I think Dr. Schulz was justified in making this assumption because his task was basically practical and not theoretical: what will the income distribution look like in 1980 closely enough so that certain gross conclusions may be drawn correctly?

I urge, therefore - as discussants have urged from time immemorial - that the authors not lose sight of our deep theoretical problems. It would be easy for Manser et al. to "export" their regression approach to other problems without advancing our theoretical understanding of the substantive issues involved; similarly for Dr. Schulz. And these studies would have practical importance of the kind seen here. Nevertheless, I urge our authors not to take this easy road but to devote some of their talent and energy to the very difficult substantive problems involved. I urge our authors in this way because I believe that essentially new advances in simulation rest on advances in understanding these deep problems.

In conclusion, I want to praise the authors for actually having completed their simulation studies. I have discovered that a lot of enthusiastic talk goes on about the possibilities of simulation, but that relatively few studies get done. I am delighted that the present authors have escaped that easy fate.

1/ Marilyn E. Manser, Thomas H. Naylor, and Kenneth L. Wertz, "Effects of Alternative Policies for Allocating Federal Aid for Education to the States: A Preliminary Analysis," Econometric System Simulation Program, Working Paper No. 34, May 15, 1969.

2/ James H. Schulz, "The Use of Simulation Techniques to Study the Changing Economic Situation of the Aged," August 1969.

3/ James H. Schulz, The Economic Status of the Retired Aged in 1980: Simulation Projections, Research Report No. 24, Office of Research and Statistics, Social Security Administration, U.S. Department of Health, Education, and Welfare, Washington: U.S. Government Printing Office, 1968.