DISCUSSION

Bette S. Mahoney, Department of Health, Education, and Welfare*

Metcalf's paper is an interesting and useful one and there is much for a discussant to comment upon. Some of what he says applies to social experimentation in general and not just to the income maintenance experiments. There is much in it about which I will not comment and much with which I agree.

I disagree with the proposition that "the methodology of controlled experimentation has taken firm hold as the focal point of analyses of major changes in social programs." Experimentation is an important methodology for developing knowledge. It is both costly and limited in its results and total reliance upon it as the focal methodology, I think would be an expensive error. I see little evidence that such reliance has occurred. One has only to examine the recent Welfare Reform analyses for support of my contention. And since the proposition is not central to the discussion of the paper or the subject of the role of statisticians in experimentation, I will not discuss my concerns about social experimentation here.

The role of a statistician in social experimentation ought to be the design of the most efficient and effective methods to meet the purposes of the experiment. Metcalf describes a prototype model without regard to the purposes or the hypotheses to be tested. Because of this he presents a "strawman" prototype as the "contribution" of statisticians. One could expect statistiticians as well as economists to be more sophisticated in experimental design.

The paper's description of the difficulties in determining the "policy space" is useful. The additional questioning of the purposes of social experimentation is better suited for another discussion. The difficulties described in the Metcalf paper led to an innovative design, the Watts-Conlisk model, the merits of which are still being discussed. As Metcalf notes the design has its costs.

Metcalf calls our attention to the fact that the sample designes in the experiment are non-orthogonal and properly warns that the data therefrom must be analyzed with caution. I want to talk for a moment about this lack of orthoganality and the Watts-Conlisk sample allocation model which is responsible for it.

Cost, as Metcalf notes, is one important factor. Some observations cost more than others. The differences in cost raise two quite different problems. First, cost is not known a priori; it is, in fact, a major objective of the experiments to determine what the cost is. Thus, differences in cost per observation cannot be perfectly accounted for in the sample design unless one already has the knowledge that would make the experiment much less valuable, if not altogether unnecessary.

The second problem has to do with value, not with the experiments themselves (at least not in this context) but with the individual observations. One would not consider filling up the cells that are likely to be the cheapest unless one thought that an observation anywhere in the design space was equal in value to all the other observations.

Metcalf lists both cost and value among the four major components of the Watts-Conlisk model but does not dwell on them. He does comment about a third major component, the specification of the assumed structural relationship being tested. He observes that, if it is being properly specified, it can be a valuable tool in increasing the efficiency of an experimental design. He goes on to point out, however, that some social scientists have criticized the experiments for incorporating structural assumptions within the experimental design since those assumptions may turn out to be incorrect. And this leads me to my point: the assumptions about both the cost and the value of individual observations may also be incorrect and lead to errors in the sample allocation.

This is probably not the place to start a discussion about the value of knowledge. I will simply note that the Watts-Conlisk model assumed that policy makers were more likely to prefer some policy parameters than others. But to my knowledge, no one when using the model bothered to ask people running welfare programs which parameters were preferred before assigning weughts to the objective function. One result is that the experiments have generated virtually no information about the potential impact of plans with tax rates much higher than 70%. While many economists and other social scientists believe that policy makers should prefer plans with lower tax rates, there are others, including many administrators of current welfare programs, who disagree. They believe instead that plans with high tax rates are more efficient and effective.

Differences in value among different observations need not result only from the relative interests of policy makers. They may also be a function of how the data will be used. For example, suppose one important use of the data will be in estimating the cost of national programs. Other things constant, one would want greater relative estimating precision for plans with high tax rates than plans with low rates. This is so for exactly the same reason that high income taxpayers wish to be more accurate in estimating their annual income than low; to wit, the same relative error will be more costly. An objective function might specify that the dollar cost of plans with 50% tax rates and the dollar costs of plans with 70% tax rates should be estimable with equal absolute precision would assign more observations to the latter plan.

This might be the place to observe that data from the experiments have been used in estimating

the cost of the Administration's welfare reform plans. The data have also been used to buttress the contention that the labor withdrawal effects of the proposed plans will be within acceptable limits. But they played almost no role in the deliberations over which plan would be preferred and, as far as I can tell, which tax rate is to be preferred for its labor supply effects.

In fact, the choice of a 50% tax rate in the Administration's proposals are made despite the findings from the experiments of small labor force withdrawals and potentially higher program costs at these rates.

The administrative evidence which several years ago was felt to be of major importance to the implementation of a negative income tax has also been of limited use. One has only to compare sample sizes of 800 to 5,200 families in the experiments to 10 million households in the program to understand why.

There should be a role in the design of social experiments for the statistician. Hopefully it will be to do more than to design a "prototype model" like that described in the Metcalf paper. It seems likely that the constructive interaction of several disciplines might do more for the development of the methodology than the dominance by a single discipline even if it is my own.

As to the future of social experimentation, expensive as it is, I agree with Metcalf's view of it "as an augmentation" to other methods. As such, it can be very valuable but appropriate care should be taken to use it wisely.

*This discussion owes much to W. Michael Mahoney of the Social Security Administration for his invaluable assistance.