

This PDF is a selection from an out-of-print volume from the National Bureau of Economic Research

Volume Title: Evaluation of Econometric Models

Volume Author/Editor: Jan Kmenta and James B. Ramsey, eds.

Volume Publisher: Academic Press

Volume ISBN: 978-0-12-416550-2

Volume URL: <http://www.nber.org/books/kmen80-1>

Publication Date: 1980

Chapter Title: Some Comments on the Papers by Kagel and Battalio and by Smith

Chapter Author: Frank P. Stafford

Chapter URL: <http://www.nber.org/chapters/c11713>

Chapter pages in book: (p. 407 - 410)

Some Comments on the Papers by Kagel and Battalio and by Smith

Experimentation and Tests of Economic Hypotheses

FRANK P. STAFFORD

DEPARTMENT OF ECONOMICS
UNIVERSITY OF MICHIGAN
ANN ARBOR, MICHIGAN

From reading these papers one can conclude that experimental methods can be used to evaluate a wide range of economic hypotheses ranging from models of individual economic behavior to market processes. Given the rapid growth in models of market equilibrium and adjustment, e.g., signaling, rational expectations, statistical models of discrimination, and given the paucity of data on the particulars of the market adjustment process, experimental methods may gain in popularity as a research methodology.

An experimental sample can range from a nonrepresentative collection of nonhuman subjects to a probability sample of actual people who, under the conditions of the experiment, are making decisions about economic variables with which they normally concern themselves. At one extreme we have estimation of income and substitution effects for nonprimates and at the other are the responses by households and firms to experimental conditions which are part of their normal decision processes. Smith appears to support the view that experimental conditions closer to reality are preferred, though Kagel and Battalio subscribe to the potential usefulness of experiments in "unreal" contexts.

One problem which is not addressed in either of the papers is whether or not alternative behavioral assumptions, e.g., nonmaximization, lead to observations comparable to those in a given set of experiments. The use of

experiments to distinguish among alternative hypotheses is not set out. Although we learn early in graduate school that it is important to test alternative models of observed behavior, in practice econometric research on field data rarely does so. The null hypothesis of no effect of a variable is scarcely an interesting alternative! It could be supposed that an advantage of the experimental method is that one would pay more attention to establishing conditions which would more clearly allow comparisons of models. Yet, for example, the Kagel-Battalio paper does not discuss the issue of whether or not models alternative to the consumer demand model can be viewed as inconsistent with token economy results.

Both Kagel-Battalio and Smith discuss experimental results not so much from the point of replicability but from the point of view of credibility. Kagel-Battalio utilize the terms of internal and external validity, the latter meaning the ability to generalize experimental results to nonexperimental settings. Smith's term for this issue is "parallelism" which is that the propositions demonstrated experimentally hold in nonlaboratory settings. To assert that parallelism is a useful precept is not sufficient, and Smith does not offer advice on mechanisms to effect a translation of laboratory experimental results to nonexperimental settings. Just because college sophomores do not collude in the prisoner's dilemma game does not mean that firms will not; even Alphonse and Gaston know better. If we define a reality scale for experiments varying from zero for responses by lower life forms to experimental variations to one for responses by individuals to real variables which are a part of their normal decision processes, then it would seem that "parallelism" would require that experiments be replicated across different levels of the reality scale, particularly toward the upper end of the scale.

An approach to research in economics alternative to experiments can be called the phenomenological approach. This approach is based on direct observation, perhaps by chance, of some empirical regularity in the real world for which we then seek to organize an explanation. I would contend that the phenomenological approach is, in fact, the major organizing theme for new ideas in economics. Examples include the following:

- (a) the Phillips curve a simple empirical regularity, a curve, which led to a variety of conceptual models and empirical research;
- (b) observed lifetime earnings patterns, rising precipitously in early years (and more so for those with more education) and then leveling off, which motivated a whole literature on human capital theory and associated spinoffs.
- (c) detailed census data, which have shown that married women who are educated work more consistently in the labor market except when there

are preschoolers in the household; this helped to motivate a literature on quantity and quality of children or the "new economic demography" and the "new home economics";

(d) the consumption/income ratio, which was observed to be secularly stable at about .9 while the apparent marginal propensity to consume from cross-sectional studies is .75; this had led to a variety of economic hypotheses and field research.

In these examples, of course, formal econometric models have been used to test hypotheses, but the initial interest was spurred by the discovery of an empirical regularity or inconsistency rather than by a premeditated effort to test a hypothesis. I am not arguing that a prespecified theory is not helpful, for without a theory or expected result there can be no phenomena; all outcomes are equally likely! I am simply arguing that induction has a major role in the discovery process. It may be that experiments are a technique fairly well suited to evaluation of a prespecified hypothesis rather than for discovery, although I am not too strongly convinced of this. It is probably true that *any* prescribed procedure or methodology is not particularly suited to greater possibilities of discovery, though perhaps experiments could be conducted to see whether innovation is more likely with experimental research rather than with *ad hoc* field research which most economists do.

The token economy results as discussed by Kagel and Battalio appear similar in level of insight to that offered by W. A. Radford (1945). At this point experiments appear to be most successful in demonstrating basic principles but are less likely to be useful if we want to know, say, the price elasticity of demand for gasoline. As Smith demonstrates, however, for system design we may want to know qualitative differences in response under alternative settings as in his examples of bidding. But one would not want to constitute a nationwide negative income tax based on the experimentally determined labor supply parameters of deer mice. Yet, a positive substitution elasticity on work hours with respect to wages for deer mice may be useful in convincing some economists of the general validity of the standard labor supply hypothesis, and in some areas we *do* make resource allocation decisions based on quantitative responses by animals to experimental variations. We do formulate public policy for people that are based, in part, on experiments with mice in the health area. Federal regulations on occupational health and safety are often based very heavily on responses by mice since *all* the costs of uncontaminated inferences from field studies or experiments with human subjects are viewed as prohibitive. This underscores Smith's contention that the question of using various experimental or nonexperimental techniques is largely a matter of cost.

At this point there is no new way of thinking about economic problems derived from experiments nor any evidence that economists have been persuaded of the validity of a new hypothesis largely on the basis of experimental results. Consider the Negative Income Tax experiments. The major "findings" were (i) the net result of substitution and income effects appears to be small with respect to married males but (ii) appears to be substantial in the case of women. Of course these were the accepted "truths" based on field research prior to the experiments. Would experimental results inconsistent with these priors have been credible for economists, or would there have been *ex post* references to "Hawthorne effects" and the like? My point is that experimental results may have reinforced our beliefs about hypotheses but have not as yet resulted in new beliefs or in new ways of conceptualizing economic problems.

I do not believe that limited credibility for experiments is a *necessary* result but that it derives primarily from the lower level of professional energy expended by economists in this form of research. A well designed experiment on peak load pricing of electricity could generate empirical evidence to convince most economists that demand substitution between peak and off-peak periods is presumably substantial for some uses. Further, *if* the experiment were performed on a probability sample so that one could generalize to statements about *market* demand parameters at least as well as econometric models designed to analyze "field" evidence, then experimental evidence would be a useful addition to the economist's tool kit. When experiments are designed in the context of a cost-benefit analysis statistical decision theory can be used to determine the scale of evaluation, such as sample size, types of treatments, and so on. Hence experimentation may have a future in public policy evaluation as well as in general research.

REFERENCE

- Radford, W. A. The economic organization of a P.O.W. camp. *Economica*, 1945, 12(48), 189-201.