Comment

David Freedman

Hodges has written a thoughtful and scholarly essay on the strengths and limitations of statistical models in policy analysis. Although there may be some differences in emphasis, his views are quite close to mine. I think the two main points are as follows:

(i) Models are usually chosen on the basis of familiarity and tractability; the degree of correspondence with reality is seldom of primary concern. Hodges' formulation:

> "Certainly the range of possible model selections is strongly conditioned by the set of models the analyst's software can handle and by the analyst's desire or ability to spend time and money developing custom software. Models favored by readily available programs tend to allow only linear causal relationships, and random variables are usually members of exponential families. The dominant position of these models notwithstanding, they are little more than conventions: they have become conventional through constant exposition in service courses and textbooks, through availability in popular software packages, and because their mathematical tractability makes them inviting examples for scholars seeking to propagate new theory and methods."

(ii) Policy analysts usually assess only one component of the uncertainty in their results; in effect, they partial away uncertainty about structure. As Hodges dryly says,

> "this creates an inherent tendency for analyses to understate uncertainty about predictions—about what is known—which can lead to invisible biases in policy considerations based on those analyses and can obscure the role of judgment and convention in the conclusions they produce."

These generalities aside, what really caught my attention in Hodges' paper (as may be only natural) was the sympathetic discussion of my own work. He quotes me—rightly—as saying that in many contexts, ad hoc analyses by experts may be better than modeling (write essays rather than fit models); recognition that some questions cannot be answered at all sensibly by the analysts within their contract performance period may be the wisest course of all.

Hodges then asks, "In what sense are the energy policy models that Freedman attacks *not* ad hoc expert analyses, albeit elaborate ones?" This is a rhetorical question, but I'll respond to it. Models are often defended—not by Hodges—as being state of the art, objective, scientific exercises, with assumptions made explicit. By contrast, analytical essays informed by data are old fashioned, arbitrary, unscientific, with crucial premises left unstated.

Some may find this defense of models an attractive fantasy, but fantasy it is. My experience includes risk assessment and econometrics. (So I do not comment on the air force logistic models discussed by Hodges.) To make contemporary model-based policy analyses in Washington, the analyst has to introduce dozens if not hundreds of fairly arbitrary assumptions. Rather than being articulated and defended, these are buried in the statistical estimation procedures, or even deeper in the computer code.

Because first versions of models seldom give plausible results, the analyst has to massage inputs, outputs and model innards, until these are more or less in balance. Indeed, one well-known modeling group is famous for the "add factors" that must be applied to regression intercepts in order to get sensible-looking macroeconomic forecasts. Such Rube Goldberg contraptions are models by courtesy only.

Hodges' question implicitly acknowledges the arbitrariness of current policy models and their weakness as formal arguments. He seems to be asking whether I would be more sympathetic to the models if they were relabeled as informal argument. The answer is, a little.

Computer code often functions as a decent veil of technical obscurity covering up some basic silliness. Articulating models in English rather than FORTRAN tends to make their problems—the multiplicity of arbitrary and unreasonable assumptions—more visible. That is why the code was there in the first place.

In summary, essays can be more objective, scientific and explicit than computer models. Consider, for example, the debate on capital punishment. Hodges cites (not approvingly) the model in Ehrlich (1975); for a devastating critique of such models, see Leamer (1983). By contrast, Zeisel (1981) is a fascinating and persuasive essay based on data. For more general discussions of modeling issues see Freedman (1987), Freedman and Zeisel (1987), Kolata (1986) and National Academy of Sciences (1984).

Continuing his review of my position. Hodges also asks, "how should data and data reduction techniques be used to inform all these judgments?" Then his

David Freedman is Professor of Statistics, University of California, Berkeley, California 94720.

unkindest remark (which is not so unkind): "It is difficult to see how these questions can even be posed within the frequentist framework." This seems wrong. There is no difficulty in posing the questions, in either the frequentist or Bayesian framework; Hodges just did it. The problem is finding answers.

Now there comes a shade of difference between us. He is a little more optimistic than I am about the potential usefulness of Bayesian techniques for properly integrating judgments about uncertainty. For example, he discusses predictive distributions starting from (i) a prior on models and their parameters and (ii) a likelihood function for the data given the model and parameters.

This is quite sensible, provided there is a sound basis for choosing the prior and the likelihood. Unfortunately, Bayesian policy analysts can be just as slaphappy in such matters as us frequentists. For discussion of this issue, see Freedman and Navidi (1986) or Hill (1985).

Good statistical analysis can be done in either the frequentist or the Bayesian framework. However, for either approach to succeed, the analyst has to get the model right, or close enough. That idea may seem ridiculously old fashioned. As policy analysts can be heard to sputter, "Models be right? How can they be right? They're all approximations. Even Newton was wrong. And a mystic besides." Because nothing is perfect, anything goes.

Hodges wants "to bring de Finetti to . . . practitioners." As I understand him, for de Finetti a prior represents a major intellectual commitment to be adopted only after serious investigation of the subject at issue. If policy analysts followed that percept, we would all be better off. The real issues here are of science, not statistical technique.

ADDITIONAL REFERENCES

FREEDMAN, D. (1987). As others see us: A case study in path analysis (with discussion). *J Educational Statist.* To appear.

FREEDMAN, D. and ZEISEL, H. (1987). From mouse to man: The quantitative assessment of cancer risk (with discussion). Technical Report, Dept. Statistics, Univ. California, Berkeley. *Statist. Sci.* To appear.

HILL, B. (1985). Some subjective Bayesian considerations in the selection of models (with discussion). *Econometric Rev.* **4** 191–288.

KOLATA, G. (1986). Asking impossible questions about the economy and getting impossible answers. *Science* **234** 545–546.

LEAMER, E. (1983). Taking the 'con' out of econometrics. *Amer. Econom. Rev.* **73** 31–43.

NATIONAL ACADEMY OF SCIENCES. (1984). *Improving Energy Demand Analysis.* Washington.

ZEISEL, H. (1981). Race bias in the administration of the death penalty: The Florida experience. *Harvard Law Rev.* **95** 456–468.

Comment

Seymour Geisser

Now comes James Hodges to inform us on some of the larger issues of statistics. And what are these issues? They are the ones that statisticians have dealt with—lo these many years—uncertainties from various sources. And there are other issues besides—is it an observational study? a controlled experiment? a retrospective investigation? a haphazard collection of items? Is what is measured or observed actually what one defines as measured? Are there flawed observations? Was the experiment or trial carried out according to the protocols? Is there a temporal imperative with regard to an action or a decision? There is, to say the least, limited interest (other than procedural validation perhaps) in the prediction of events that already have occurred and been observed.

What is the point then? The point is that we have here a lucid and trenchant exposition vividly reminding us of three of the principal sources of uncertainty or variation. What is more novel than most previous explications is that the sources are related to predictivism, which is stressed as the penultimate aim when taking an action is the ultimate goal. Hence, from my point of view, there is really nothing to quarrel with. But it is the job of a discussant if not to be quarrelsome to be at least quibblesome—to coin a neologism.

Hodges intimates that for proper application of statistical methods, the implementation of de Finetti's approach is required. He also states that the approach "lacks a crucial connection to real problems." I would like to quibble with both these points. In regard to the latter point, we have only to realize that de Finetti was involved in applications especially in finance,

Seymour Geisser is Professor and Director, School of Statistics, University of Minnesota, 270 Vincent Hall, 206 Church Street, S.E., Minneapolis, Minnesota 55455.