Reply to our discussants

KEVIN D. HOOVER1, STEPHEN J. PEREZ2

1Department of Economics, University of California Davis, CA 95616-8578, USA.
E-mail: kdhoover@ucdavis.edu; Homepage: www.ucdavis.edu/~kdhoover/
2Department of Economics, Washington State University Pullman,
Washington 99164-4741, USA
E-mail: sjperez@wsu.edu; Homepage: www.cbe.wsu.edu/~sjperez/

Received: 20 January 2000, in final form January 2000

1. THE GENESIS OF ‘DATA MINING RECONSIDERED’

Let us begin by thanking our eight discussants. It is an unexpected blessing to have received careful and extensive consideration from such a constructive and fair-minded group. Our reaction to their comments is perhaps better understood with some autobiographical perspective.

As is probably evident to any reader, we are not econometricians. This is not to offer an excuse for any shortcomings of our work, but to explain that it was motivated by interests somewhat different than the typical interests of econometricians. We are both applied monetary/macroeconomists, and one of us is particularly interested in the philosophical and methodological problems of empirical economics. We have both done applied work on causal inference in macroeconomics, using concepts and techniques substantially different from those that originate in the work of Clive Granger. In our causal investigations, we have used LSE specification procedures. Our experience has been, contrary to our initial expectation, that the validity of these procedures, which we took to be fairly well established, was more often questioned by referees than was the validity of the causal methodologies that we proposed, and which we regarded as novel.

We were faced with a dual problem: on the one hand convincing referees and editors that our applied work made sense; and, on the other hand, assuaging the doubts that were naturally raised in our own minds by intelligent critics. Thus, while we had a natural predisposition towards the LSE approach, we believe that we were genuinely ready to acknowledge whatever results we obtained. Interestingly, while LSE econometricians have generally greeted our paper enthusiastically, opponents of the LSE approach often see the results that we report—even taken at face value—as lending little comfort to LSE practitioners. The key thing to remember is that our goal was not to provide a generalized tool for specification search, but to ask whether it worked in a case in which we thought the LSE approach should work, if it could ever work at all. The five commentaries raise a large number of points. We will not attempt to address them all. Rather, in the remainder of the reply, we would like to consider only some of the larger themes.
2. IS OUR ALGORITHM TRUE TO THE LSE APPROACH?

Bruce Hansen raises this question most pointedly. While our intent was to provide an algorithm that both proponents and opponents of the LSE approach would recognize as sharing key features with it, the algorithm is our own and any mistakes in our characterization should not be laid at the door of best LSE practice. There is, unfortunately, in any complicated research project path-dependence. Some decisions made early on, that in hindsight should have been made differently, get locked in. If one really wants examples of typical LSE practice, one is better advised to look at the commentaries of David Hendry and Hans-Martin Krolzig and of Julia Campos and Neil Ericsson, who write as paid-up members of the LSE school. However, in any case, the LSE approach is not a set of detailed recipes. Rather, as we discuss in Section 5 below, it is a family of related ideas, which is open to progressive refinement.

That said, let us respond to a few of Bruce Hansen’s queries. No doubt it is best LSE practice to use heteroscedastic-consistent standard errors, and no doubt we should have done so. We do not believe, however, in the context of our particular experimental design that this would have made any substantial difference: the error terms were all drawn from homoscedastic distributions. LSE practice as embodied in, for example, the PC-GIVE econometrics package does typically use the Chow, normality and autoregressive conditional heteroscedasticity (ARCH) tests built into our algorithm (see also the examples in Hendry and Krolzig’s and Campos and Ericsson’s commentaries). Our understanding is that normality and ARCH tests are justified as frequently having power against more general specification error. We in fact use Andrews’s (1993) test of structural change in our own work on causality, but note that it is not built into, for example, PC-GIVE.

There seems to be general agreement among the commentators that the algorithm would have been improved through the use of the Bayesian information criterion (Schwartz criterion) or the Akaike information criterion. While this was not an element in our own understanding of the LSE approach when we began this project some time ago, it is evidently now best LSE practice.

3. IS OUR ALGORITHM AN INTRUMENT OF APPLIED RESEARCH?

The reaction to our paper that surprised us the most is the enthusiasm with which Hendry and Krolzig greeted it as a precursor to their own effort embodied in PC-GETS to construct a generally useful package for specification search. (Campos and Ericsson provide a second practical illustration of this package in action.) Our goal was to provide evidence on the question, could the LSE methodology work at all? In David Hand’s useful terminology there was a strong ‘iconic’ element to our model of the LSE procedures. We were conscious of leaving out many salient features (e.g. issues of exogeneity and a role for prior theory, such as stressed by Clive Granger and Allan Timmerman in their commentary). We attempted to compensate for features that we could not model easily (such as the encompassing of rival models) with features that we had not routinely encountered in LSE work (such as multiple search paths). Still, referring again to Hand’s commentary, we thought of our algorithm as a toy car—albeit one with a fairly powerful engine. We are pleased—and a little proud—that Hendry and Krolzig have used our toy as the model for a full-sized prototype. And we value all the detailed suggestions of the various commentators on how best to implement that scaling-up.
4. WHAT IS THE ROLE OF TRUTH IN THE ASSESSMENT OF THE ALGORITHM?

There is one line of criticism, however, that we do not accept. Both Hand and Granger and Timmermann criticize our use of the recovery of the true specification as a standard of assessment and our implicit assumption that the model space includes the true specification. We believe that this criticism reflects a mistaken view of the role of truth in econometrics, a misunderstanding both of our strategy and, as we discuss in the next section, of the spirit of the LSE methodology.

We agree that we cannot use the truth as a reference point in actual empirical investigations. Truth is, none the less, an indispensable notion to useful econometrics. In simulations we can know the truth, and we would find any search procedure that failed to uncover the truth, when it was there to be found, to be lacking to some degree. Hand and Granger and Timmermann wish to replace the notion of the true model with the notion of an adequate approximation, where ‘adequate’ depends on one’s loss function. However approximation means, ‘coming close (or close enough) to the truth.’ To test the adequacy of an approximation, we must create a situation in which we know the truth (or some bounds on the truth).

Hand proposes a simulation standard and a measure of similarity, but similarity is a standard that requires the regulatory notion of truth just as much as the experiment that we proposed. We believe that Hand’s proposal that the search algorithm can be evaluated by its predictive performance alone and that the true structure is irrelevant is clearly mistaken. Judgements of predictive performance are always based on past (or examined) data. Genuine predictions are always of future (or unexamined) data. Any inference that says that genuine prediction of the future (or unexamined cases) is warranted by the performance in the past (or examined cases) must presuppose that the process that generates the data (that is, the unobserved structure) remains constant. The assumption of constant structure may be based on faith alone. It may be based in extra-statistical knowledge. Or there may be statistical evidence that data are at least consistent with constant structure, even if such evidence cannot positively demonstrate constancy. The critical point is that genuine prediction must be grounded in a claim (implicit, at least) about constant structure, which is a claim about what is true about that structure. Prediction does not avoid assumptions about truth; it requires them.

There is nothing novel in our position. It is the well-known basis for the Robert Lucas’s (1976) policy non-invariance argument, which was not original to Lucas. The Lucas critique is probably not that important empirically, but its logic is correct. Policy evaluation requires a notion of true structure (or adequate approximation thereto). Forecasting may not require knowledge of the true structure, but it certainly must presume its constancy. And how could one know that it was constant without some attempt to characterize it? We would never claim in any practical empirical work that the model space in which we conduct a specification search necessarily contains the truth. Every possibility mentioned by Hand and Granger and Timmermann, such as omitted variables and non-linearities, would occur to us as well. While we stick by the claim, challenged by Hand—indeed we think that it is trivially true—that a sufficiently complex model can in principle characterize the true data-generating process, we emphasize the qualification ‘in principle.’ The necessary variables may not have been observed; the functional forms may be terribly complex. The art of applied econometrics is partly to guess what can be omitted and what can be simplified, and it is partly to gather evidence that these omissions and simplifications do not do too much damage.

We do not claim that the truth is nested in the model space of our search. Instead, we ask, if it happens that the truth is in fact nested in the model space (as it can be in a simulation in which
there are a limited number of variables and functional forms), does the search procedure recover it? The argument is, if it fails to recover it in such a favorable case, then a fortiori we should doubt its usefulness in less favorable cases. The ability to recover the true specification when it is actually known is, we believe, an essential property of a good search procedure and an interesting thing to know about it, but it is not the only thing to know.

5. THE SPIRIT OF THE LSE APPROACH

We believe that the criticism of the search procedure for focusing on the truth misunderstands the fundamental nature of the LSE approach. Here we run the risk of going too far: others are better qualified to speak for the LSE view than are we. This is, none the less, our own understanding of what is essential to it.

The critics of truth present a static picture of econometric investigation. If a specification search were once and for all time, then it could be a serious matter if the search were conducted on a space in which the true model did not appear. However, the LSE approach is, in our view, a critical (perhaps dialectical) methodology. The key feature is not the general-to-specific search, which is largely a way of dealing with the unmanageably large number of possible specifications, but rather the adjudication between competing specifications. The reason that encompassing tests feature so prominently is that they are instruments of adjudication. The point is not dogmatic attachment to particular implementations of encompassing tests; rather it is the spirit of adjudication. If two competing specifications claim to characterize the data, they cannot both be correct. They might both be incorrect or it may not be decidable given the particular sample which specification is correct. And, even if one specification encompasses another, it may still not encompass another as-yet-unproposed alternative; so all judgements are tentative. The correct response, then, to anyone who thinks that something is omitted from the model space is to seek adjudication—e.g. test whether either the selected specification and a proffered alternative encompasses the other or expand the model space appropriately and search again. The pattern of criticism and adjudication is, we believe, what Hendry meant in other contexts when he referred to the LSE approach as a ‘progressive research program.’ The details of the implementation of the LSE approach are all potentially dispensable; so long as it maintains the pattern of criticism and adjudication, a search procedure will remain recognizably part of the LSE tradition. Thus, there is no reason why the suggestions made in Granger and Timmermann’s Section 5 or in Hansen’s Section 5.2 could not be integrated into the LSE approach if they proved to be successful practically. The assimilation of a large number of ‘foreign’ influences might undermine the rationale for referring to the syncretic product as the LSE approach, but then that is only a convenient label—and one that many of its practitioners dislike anyway. The real point is to do better econometrics. And we believe that each of the commentaries has given us lessons in how to do that.

REFERENCES
