Richard Colombo (New York University): It is usually more difficult in the social and perhaps biological sciences to exert control over variables of interest than it is in the physical sciences. Consequently much of empirical social science is largely concerned with identifying variables and contingent factors that are relevant to a phenomenon. When regression analysis is used it is not usually to find a routinely predictable relationship but more modestly to find a routinely predictable set of relevant variables. On this admittedly weak test there are perhaps many regression success stories.

That more ambitious goals are not often attempted may be because statisticians have failed to develop good practical methods for analysing many sets of data (MSOD) (although the skill, effort, luck and insight involved in collecting and analysing MSOD to establish routine predictability should not be underestimated). In marketing, Professor Ehrenberg and his colleagues have successfully established routinely predictable relationships—so it can be done but there are few, if any, other examples. Much of the problem lies in the concern that statistics has with single sets of data, as the authors point out. This legitimizes studies based on one data set to the detriment of well-tried methods of science involving MSOD. For an example of the unsuitability of 'conventional' statistical approaches for tackling MSOD see Colombo et al. (1993).

The authors have showed how MSOD can be analysed when a small number of variables is involved. Where there are more than two or three variables, as is typical in social science, the search for a good model to fit MSOD is likely to be much more difficult. What are we to do in these circumstances?: settle for qualitative statements?; report regression coefficients for the data set but with a warning about the likely absence of predictability? Since much of social science leans heavily on statistical methods, statisticians have an opportunity to influence, perhaps radically, the way that social science research is conducted by developing techniques for MSOD. The plea for statisticians to be more 'scientific' and 'relevant' has often been made. With their focus on routine predictability and MSOD the authors eloquently repeat this plea but also offer a specific prescription about how it might be achieved. This paper should be widely read and pondered.

David Cox (Nuffield College, Oxford): The importance of studying relationships under a known range of different conditions is stressed in the traditional work on the design of experiments. See, for example, Yates and Cochran (1938) on variety trials. The insertion of factors into a design deliberately to achieve a good range of validity is standard practice.

There is much in the present paper with which to agree, but the following remarks concentrate on an issue on which tentatively I disagree strongly with the authors. In some sense science is about 'understanding what is happening'. Whatever the history, Boyle's law as an empirical prediction equation seems of fairly limited interest; its importance lies more as a link in a coherent set of ideas in classical physics involving thermodynamics, the kinetic theory of gases, the virial expansion and so on. Section 6 of the paper addresses these issues. Although a healthily empirical attitude to some of the more imaginative aspects of, for example, modern physics is no doubt desirable, it seems to me that the authors' account massively undervalues theoretical discussion. The authors see as the main challenge the development of more and better empirical prediction equations. An alternative would see the need for analyses that are more strongly based in the subject-matter knowledge of the field, i.e. involving substantive rather than empirical models. Although the balance must depend on the context, I see the second as the more pressing issue.

Finally there is no space to comment on the remarkably pessimistic statement in the final sentence of the second paragraph of Section 1.

D. R. Draper (University of California, Los Angeles) and C. L. Mallows (AT&T Bell Laboratories, Murray Hill): We like the paper's emphasis on prediction. However, the authors' main points provide new support for the old maxim that statisticians working on quite different types of applied problem may come to quite different conclusions about what is important and what has been neglected in the subject. If Professor Ehrenberg and Mr Bound had developed expertise in the field of medical research, say, where prediction of individual patient level outcomes is at the heart of comparison of alternative treatment protocols, rather than consumer research, where prediction of aggregate buying habits seems important, a paper like this might have had two different emphases: the authors would perhaps not have come down so hard on least squares and allied methods, which are about the best that we can often do when working with individual level data, and they would probably have noted that it is much easier to find law-like relationships at the aggregate level. We suspect that the statistical tent is sufficiently large to find room for workers searching for predictability at a variety of levels of aggregation.
Professor Ehrenberg and Mr Bound continue to castigate us and our co-workers (Draper et al., 1993) for insisting that some 'leap of faith' is always present in prediction, and for providing only a circular argument. But their arguments in support of routine predictability also seem circular: a 'law' holds except when it does not. Our differences seem to be merely semantic. Their formulation in the third paragraph of Section 6.3 of the 'basic statement of empirically based induction' can be stated in our language as

'If past observations are judged exchangeable, and the conditions under which they were observed are judged exchangeable with a new set of conditions, then we can predict that the event will recur in the new situation'.

We can predict, but we may be wrong; we need to have faith that no new factor is influencing the outcome. Some sort of leap of faith about the persistence of structure (Hodges, 1987) is always present in prediction.

Stephen G. Hall (London Business School): This paper contains much with which I agree. If we take its main point to be a call for the analysis of all available data then few could argue with it. If in addition it is arguing for a depth of insight which can emerge from using data from different time periods then my own work and recent publications on co-integration would support the conclusions drawn here.

I believe that the data under consideration are characterized by two features. First, it is a panel of data consisting of a number of cross-sections taken at different points in time. Second, because the moments of the distribution of the cross-sections change over time they are drawn from a non-stationary population. This last point is crucial to the procedure proposed by the authors as otherwise the sets of data would only differ by sampling error and so the line drawn through the means of the data sets would be meaningless. Much recent work has been undertaken on appropriate estimation and inference strategies for models involving non-stationary data. Although the procedure proposed in this paper will clearly give consistent estimates of a well-specified model, my concern is that in the absence of any theory of inference the rejection of incorrect models becomes a matter of arbitrary interpretation on the part of the researcher. Also, although not so relevant to the types of data set used by the authors, the technique proposed here is an inefficient way of using the data for small data sets.

The theory of non-stationary regression offers a formal framework for testing models of this type. It also allows the treatment of more complex models, both in terms of the number of variables and the dynamic structure of the model, and a better understanding of the efficiency considerations when data sets are small. This paper points to some serious weaknesses in traditional regression theory and this I fully applaud, but this does not mean that the whole body of statistical theory should be abandoned. The theory can, and has been, extended to account for these problems and I believe that it offers a more comprehensive way forwards than the procedures proposed here.

D. V. Lindley (Minehead): The main idea in this paper that we should often be looking for general rules, rather than models for single situations, is brilliant. The authors are correct to criticize statisticians for the comparative neglect of the topic. I would like to add a more formal structure to their argument, using subjective probability.

For several data sets, let $X_i$ denote the data in set $i$ ($i = 1, \ldots, n$). It is usual to model this by a probability distribution $p(X_i | \theta_i)$ dependent on a parameter $\theta_i$. It is then necessary to assign a distribution for the parameter set $\theta = (\theta_1, \ldots, \theta_n)$. Let this be $p(\theta | \phi)$, where $\phi$ is a hyperparameter. One possibility is to suppose the parameters to be independent and identically distributed given the hyperparameter. This general model is termed hierarchical. The main interest may then be in the hyperparameter, expressing the general rule that is being investigated, whereas the parameters express the behaviour in narrower cases. With a distribution for $\phi$, the model is complete and other probabilities can be evaluated. For example, the general prediction of $Y$ from the data is provided by $p(Y | X)$. The evaluation of this would depend on whether $Y$ was from one of the previous sets or from a new set. Details are standard and omitted.

The advantage of this type of procedure is that it gives an explicit way of calculating all quantities of interest within the framework of specific assumptions that are spelled out. The idea can easily be extended to include nuisance parameters at each stage of the hierarchy. Additional stages can be included if needed. Computation can sometimes be difficult, but progress is rapid here and even fairly complicated models can be handled.