Their usefulness much precedes and exceeds their measurability.

10. Reformulation in terms of Bayesian odds ratios will not rescue standard tail-area testing procedures, I believe. Such odds ratios for lower-dimensional hypotheses are problematic and prior-dependent. And whatever the relation between Bayesian odds ratios and real modeling issues, standard tail-area procedures will have to be transformed beyond recognition to become well articulated to either.

In usual F tests, 1/F has a pleasant interpretation as a shrinkage factor, at least in the balanced case, but this doesn't rescue P recognizably or interpret R-squared directly. (This interpretation is well known, I understand. I noticed it while reviewing Mosteller and Wallace (1964) for the Mosteller Festschrift, when working through the simple normal-theory counterpart of the difficult nonnormal shrinkage and discrimination problem that they solve.)

Rejoinder

Harry V. Roberts

I keenly appreciate the contribution of all the discussants. I have very few disagreements to record, and I have been stimulated to offer some extensions of the paper.

Dr. Deming says that the business of statisticians is to transform the company goals, not to help the management to pursue theirs. Thanks in large measure to his efforts, some companies have already transformed their goals, or are at least far along in the transformation. In these companies, statisticians need to ply their trade skillfully in pursuit of company goals, and to train parastatisticians.

Unfortunately, many other companies have not heard about, understood, or believed, the need for transformation. What do statisticians do when the organizational climate is bad, when management's goals are misdirected? (Dr. Deming once wrote me that the statistician may only prolong the life of a sick company.)

Some of my students, discouraged by the contrast between what goes on in their own companies and the advice of the Deming 14 Points, ask the same question. My first impulse is to say that it a rare statistician in the middle levels of such a company who can do much to transform the organization's goals. But that is not a good enough answer. Since statisticians often have considerable freedom in defining the data and studies on which they work, they can help to educate management. They are free to suggest, for example, that it might be valuable to study quality, lead time and inventories instead of, or at least in addition to, machine utilization, cost variances and quota fulfillments.

George Box expresses my basic view of statistics in one golden sentence and one splendid metaphor. The sentence: "In my view, statistics has no reason for existence except as the catalyst for investigation and discovery." The metaphor: teaching swimming by theoretical training alone, and the tendency of many statistics teachers to avoid getting wet.

Professor Moore's description of the U.K. situation bears many similarities with that of the United States. For example, what he says about the London Business School could be applied with only minor modification to the business school at which I teach; better use of basic statistics to improve quality and productivity is needed in the United States as well as in the U.K. and Europe; and upward mobility of accountants and lawyers is conspicuous in the United States (where Dr. Deming deplores "creative accounting" and calls litigiousness a "deadly disease").

My purpose in citing management books by Peters and others was not to endorse them in all details but to point out that statistics is but one component of a major management upheaval in many world class companies. By the regression phenomenon alone, it is to be expected that studies confined to successful companies at any one time will be embarrassed by problems encountered later by some of these companies. There's more to it than regression, however. Excellence in quality and productivity is no insurance against major management blunders in other areas, such as unfortunate acquisitions.

The mention of writers on management gives me the opportunity to cite a new book by Richard Schonberger (Schonberger, 1990) that carries the story beyond what I reported in my subsection "Beyond Parastatisticians" at the end of Section 3. The new book, *Building a Chain of Customers*, extends the focus from manufacturing to the entire business firm.

I do believe that work of Raiffa and Schlaifer (and, of course, that of Savage and de Finetti) is seminal.

I say little about decision theory, however, because it seems to have had relatively little impact on business practice. I agree fully with Professor Moore's conclusion: "If statisticians feel inhibited from going any further than the preparation of data and their analyses, there is a danger that their role will increasingly be perceived as that of a technician rather than as an executive, and they will be marginalized."

As to John Neter's "minor quibbles": I am glad that he can report that the use of probability sampling in auditing is no longer relatively rare. Intervention analysis (Box and Tiao, 1975) is not randomized experimentation, but it is a powerful tool for reducing ambiguity about causal inference. Most of the "serious misunderstandings of statistics" in Section 3.6 are to be found in widely used elementary textbooks on business and economic statistics (e.g., tests of sharp null hypotheses as decision procedures). I agree that the "glue that binds together the functional areas of business" is a composite that includes many other compounds than statistics.

I didn't say much about how a good organizational culture for statistics can be achieved, but the management writers that I have cited say a lot about it. That is one of the things I like about Schonberger's new book cited above. For another good example, see Grayson and O'Dell (1988). Some companies, such as the winners of the Malcolm Baldrige National Quality Award, are well on the way to improvement of the organizational culture. Dr. Deming's 14 Points and the criteria for the Baldrige award go a long way towards a road map.

The development of parastatisticians and the education of future business managers, which I tend to link together, are difficult challenges. Both relate to the vexing question of how to teach statistics so that students will not loathe the subject. In the last half of his discussion, John Neter gives an excellent analysis of ways of meeting the teaching challenge in business schools.

On parastatisticians, I do not know the best way to develop them. But I do know that the emphasis of statistics in training programs in leading companies is likely to lead to some employees being turned on by statistics and wanting to go farther; I know some of these personally. Some companies—for example, Motorola—offer a wide range of short courses to their employees. There is an opportunity for statisticians both in and out of universities to help companies to develop parastatisticians. One good model is the program offered by the Japanese Union of Scientists and Engineers (JUSE). There is wide coverage and, unlike most short courses in the United States, there are often sequences of short courses with substantial periods in between for application at the student's own company.

Let me also extend Neter's discussion of statistical education in universities. I would urge teaching statisticians to "push" rough-and-ready data analysis (their own version thereof, which may differ substantially from the one sketched in my paper) at every opportunity, by personal example and by guidance of students in projects. The advent of excellent computing capability makes data analysis a pleasurable hobby for any statistician who is interested in applications. For example, in the last few weeks I have had a chance to explore, hands-on, data sets bearing on global warming, an alleged leading indicator for stock returns, the effects of a modified stance for putting and principles of robust estimation. In all four instances, simple analysis led to satisfying insights. For example, the leading indicator for stock returns appears in fact to be a leading indicator. Such computer analyses, for me, are much more fun than any computer game.

It is useful also to recognize, as Neter points out, that concepts of quality improvement are applicable to education, and, in particular, to required elementary statistics courses. Teachers of these courses need not only continuing minor improvements but revolutionary improvements as well. At the Xerox Quality Forum, Leesburg, Virginia, in August, 1989, Bob Galvin (former CEO of Motorola) challenged a group of business school academics (mostly statisticians and management scientists) who were complaining about the difficulty of adding courses in quality and productivity improvement to business curricula that were already crowded. Moreover, other new courses, such as business ethics and environmental issues, were also vying for inclusion. Galvin replied, in effect, "Why can't you teach 50% more in a year? We make equivalent improvements in industry. Why be satisfied with the 1943 syllabus?"

One may contend that revolutionary improvements in industry are relatively easy and that education is different. But there are several reasons for thinking that education may not be so different.

1. Work sampling methods in industry, using methodology developed by F. Timothy Fuller (Fuller, 1985), have suggested that the time in which mental workers actually add value may be astonishingly small. For example, a student of mine, a systems programmer, found that he was spending less than 5% of his time adding value; the rest was fixing bugs and other forms of waste. My guess is that students may not do much better when they are sitting in class or studying. Note that an increase of value-added time from 3 to 6%, which seems modest, implies a doubling of output.

2. In industry, it has been found that radical redesign of products and processes can lead to major improvements that would be totally unattainable by minor changes; for example, reducing the number of

parts in an assembly may greatly increase quality. My suspicion is that some elementary statistics courses are like badly written, bug-ridden computer programs that have been repeatedly patched and have become incomprehensible and unmaintainable. Evolutionary changes won't do the job: a complete rewrite is needed! In rewriting elementary courses, one useful principle is that simplicity, not watering down, can pay off handsomely. Consider, for example, how proofs of mathematical theorems become shorter and simpler as the theorems are better understood.

3. Educational innovations in industry have suggested that desirability of relating classroom instruction to the actual job. A continuing alternating pattern of learning and application, even over an employee's working career, can be more effective than intensive short courses that are not followed up with application.

4. If the basic theme of the course is one of data analysis, the alternating pattern of learning and application can take place even within a single statistics course. This is because a series of increasingly complex applications using real data can be studied, each illustrating the sequence of model identification, fitting, diagnostic checking, and, as necessary, re-identification of the model and repetition of the cycle. As the applications became more complex, more statistical techniques are brought into the data analysis, but the underlying strategy remains the same and the basic ideas are repeated and reinforced.

5. Another theme from industry is the need for managers and staff specialists to emphasize the role of coach over that of "boss"; the new role is that of facilitator. Statistics teachers might be helped by thinking of themselves as "facilitators."

John Pratt's comments on parsimony and time series deserve very serious consideration. My own experience is that screening out "unnecessary" variables in regression seems to work. One clear example is that of two-level fractional factorial experiments where many practitioners have come to believe that only a few of the possible effects—the particular ones cannot be predicted with confidence—are likely to be substantial. In industrial experimentation, at least, confirmatory experiments seem to support parsimony. It is even possible to speak of a "Pareto effect": a few variables account for most of the action.

With observational data, where neither balance nor randomization is present, there are indeed many potential pitfalls besetting attempts at causal inference. Pratt is a major contributor to the understanding of these problems. I believe, however, that the difficulties have to do less with parsimony *per se* than with the general problem of the correlation of omitted variables—measured or not—with the variables included in the model, and that obvious mistakes and maybe some of the subtler ones can be avoided in common applications entailing time series data or successive cross sections; see Roberts (1988) for some discussion along these lines. In single cross sections, as illustrated by many employment discrimination studies or health studies, the problems of causal interpretation are overwhelming, but this has little to do with parsimony. I agree, however, that "smuggling in sharp constraints" is in principle less desirable than coming to grips directly with the assessment of reasonable nondiffuse prior distributions. It would be interesting to explore this task in the context of two-level fractional factorial experiments.

For time series data, the most important outputs for management are the predictions and prediction errors (looking forward) or the fitted values and residuals (looking backwards). These suffice reasonably well for the key task of sorting out special causes from common causes. The fact that several different-looking models fit a given time series about equally well is not important for this purpose. When one wishes to draw causal inferences from time series data, the principles of intervention analysis (Box and Tiao, 1975) are very helpful. In passing, I strongly endorse the desirability of statistical packages that combine full-blown regression models with ARIMA disturbances.

On the educational climate for statistics, Pratt's comments well represent my own views in my discouraged phases. But I am not always discouraged, mainly because of the leadership of some American companies: I have heard stronger endorsements of statistics from CEOs of leading companies than from most statisticians. And there is very great business interest in statistics, evidenced, for example, by the large numbers of short courses and seminars in applied statistical topics.

Moreover, statistical methods are being increasingly used. For example, not all statisticians are completely happy about "Taguchi methods," but these methods have diffused rapidly in the automotive industry because they seem to help to solve important problems. Much of the teaching of statistics, however, has not focused on solving problems; see the Box quotation.

John Pratt's "other points" are all interesting and useful. I have only one specific remark to make: On point 6, one simple and useful way to combine statistical forecasts with judgment is to use judgmental point predictions as additional regressors.

ADDITIONAL REFERENCES

BROWN, R. V. (1987). Doctor Logic. *Illustrated London News*, October, 28–31.

BROWN, R. V. and LINDLEY, D. V. (1986). Plural analysis: Multiple approaches to quantitative research. *Theory and Decision* **20** 133–154.

CHERVANY, N. L., COLLINS, R. O., JR., FIENBERG, S. E., JOHNSON, P. E. and NETER, J. (1977). A framework for the development of measurement instruments for evaluating the introductory statistics course. *Amer. Statist.* **31** 17–23.

COX, D. R. (1958). *Planning of Experiments.* Wiley, New York.

FISHER, R. A. (1935). *The Design of Experiments*, 1st ed. Oliver and Boyd, London.

FULLER, F. T. (1985). Eliminating complexity from work: Improving productivity by enhancing quality. *National Productivity Review* **Autumn** 327–344.

GRAYSON, C. J., JR., and O'DELL, C. (1988). *American Business, A Two-Minute Warning: Ten Changes Managers Must Make to Survive into the 21st Century.* Free Press, New York.

HAMAKER, H. C. (1977). Bayesianism: A threat to the statistical profession. *Internat. Statist. Rev.* **45** 111–115.

NETER, J. (1986). Boundaries of statistics: Sharp or fuzzy? *J. Amer. Statist. Assoc.* **81** 1–8.

PRATT, J. W. and SCHLAIFER, R. (1988). On the interpretation and observation of laws. *J. Econometrics* **39** 23–52.

SCHONBERGER, R. I. (1990). *Building a Chain of Customers.* Free Press, New York.