Statistical Science
1990, Vol. 5, No. 4, 372–402

Applications in Business and Economic Statistics: Some Personal Views

Harry V. Roberts

Abstract. Statistical methodology has great potential for useful application in business, but that potential is seldom realized. However, companies are increasingly exploiting simple statistical tools in quality and productivity improvement and developing "company cultures" congenial to effective use of statistics. Statistical and probabilistic thinking is essential for sound decision-making. Only with understanding of statistical variability can managers distinguish special from common causes of variation, intelligently direct efforts to improve processes, and avoid the tampering that can make processes worse. Statistics can be used most effectively in business when many employees—"parastatisticians"—have some grasp of statistical tools and thinking. Fortunately, there is evidence that very elementary tools suffice to make rough-and-ready studies that can illuminate most business problems and facilitate most decisions.

Key words and phrases: Parastatistician, quality improvement, common cause, special cause.

1. INTRODUCTION

I believe that there ought to be many more applications of statistics, in all fields, than there are. My particular concern is applications in business and economics. Since my experience in business is more extensive than my experience in economics, however, I shall concentrate mainly on business and offer brief comments on economic statistics at the end.

It is important to understand an important difference between statistical applications in business and those in science. Both scientific and business applications of statistics rely on a common set of statistical tools. But in business the major emphasis is in solving immediate problems, while in science the major emphasis is on broad empirical generalizations. Scientific statistical applications are usually aimed at diverse, sophisticated audiences, who will often examine the statistical methodology critically and at leisure. Business applications are usually directed to a small group of employees—even a single individual—concerned with a specific problem, who may be unsophisticated about statistics and uninterested in methodology. In business applications, analyses that can be described as "rough and ready" or "quick and dirty" (though not "confused and careless") may suffice for practical goals; there may be less need for refinement of methodology and generality of inferences than in scientific applications.

By a business application of statistics, I mean any use of statistical reasoning in pursuit of company goals. My scope extends beyond published studies or even confidential reports and presentations within companies. For example, when sales managers try to figure out why last month's sales declined when a sales increase had been expected, I regard their efforts as a statistical application, though perhaps a very informal one that may not entail explicit computation. This interpretation of statistical application includes rough and ready applications of statistical reasoning by people who lack substantial formal training in statistics as well as systematic studies by expert statisticians. It concludes even informal application of statistical thinking to problems for which there is little or no specific data bearing on the question at hand, as W. Allen Wallis taught me long ago (Wallis, 1970).

In taking a broad view of what constitutes a business application of statistics, I must go beyond published materials and rely heavily on personal impressions about business practice, so I owe the reader a brief discussion of how these impressions were acquired. I was drawn to statistics over 40 years ago by a business application based upon data from a consumer panel; I tried to measure the effectiveness of advertising expenditures for a consumer product with two dominant

Harry V. Roberts is Professor in the Graduate School of Business at the University of Chicago, 1101 East 58th Street, Chicago, Illinois 60637.

brands. From that time on, I have worked on many applications of statistics to management problems, often in a consulting capacity.

My contacts with business firms have often discouraged me about the way statistics is applied—or not applied—in business. I remember in particular a traveler from one of the leading publishers of books on statistical methodology. I was recommending a manuscript on applied time series analysis, which his company was about to pass over. I said that the book would lead to lots of useful business applications. Thinking that I would clinch my point, I said that even *his* company could make good use of the techniques. For example, by better sales forecasting, they might not run out of stock on textbooks as often as they seemed to. He remarked, sadly, "My company never uses statistics." (The manuscript was published by another company and became very successful).

Similarly, my students working full time in companies have provided feedback about the nonuse—sometimes the misuse—of statistics in their companies. I have for many years required these students to do actual company applications of their own choosing. A typical project would be a study of company monthly sales in several recent years, including exploration of outside factors that affect sales. Students have usually found these projects helpful, but they are surprised that their companies have been doing so little along these lines and that vital company statistical information—such as monthly sales data—is often either not kept for any length of time or, for practical purposes, is irretrievable.

Nor is there evidence that the situation has improved with the development of computerized data bases. Even when the data bases contain potentially useful information—for example, human resources data bases—my experience is that they often need extensive data cleaning before they can be used for serious analysis. Apparently it takes good data analysis to provide the discipline needed to maintain clean data bases.

The somewhat negative tone of the past three paragraphs does, however, need some balancing. Some large companies have made effective use of groups of skilled professional statisticians at high levels who have made major and sophisticated contributions both in consulting with others and in their own projects. Simply for concreteness, I can mention examples from personal knowledge: AT&T Bell Laboratories, Bellcore, DuPont, IBM, Allied Signal, Amoco, and Commonwealth Edison. The electric power industry and its Research Association EPRI have made effective applications. There are marketing research and financial firms who have made sophisticated applications. On the broader landscape, however, I see these as oases. Even in one of the examples just mentioned, I once made a rough approximation that there was one high-level statistician for each billion dollars of sales.

My experience in teaching and consulting have left me with two dominant impressions:

1. It is easier to do technically good statistical work than to serve actual company needs.
2. The company problems to which statistics is applied are only a tiny fraction of those for which statistics is potentially helpful.

In this paper, I have therefore decided to place major emphasis on constructive suggestions for improvement of business practice by more effective use of statistics.

Most potential applications in business, if they are to be realized, must be accomplished by nonstatisticians. These applications may be based on data that are routinely generated by the company and stored in data bases (which may, as I have suggested, require extensive cleaning before they are usable). Often, however, applications will be based on data collected quickly to meet an immediate need. Moreover, typical applications will be small and simple, as opposed to sophisticated surveys and experiments in which specialized statistical knowledge is needed.

In making constructive suggestions for improvement, I am encouraged by the experience of the companies in Japan and the United States who have pioneered in the thrust for quality and productivity improvement that has drawn so much attention in recent years. In many of these companies, there have been successful applications of statistics in process improvement by large numbers of employees at all levels. Typically the tools have been simple ones, such as histograms, scatter plots, and control charts. Not too long ago, I would have been skeptical that such widespread and effective use of statistics could happen. Today, some of my colleagues are skeptical that it is really happening and are inclined to discount the published claims. However, I had enough contact with credible witnesses and enough opportunity to observe directly what is happening in leading-edge companies to be convinced that the claims are broadly accurate and to believe that the quality movement conveys important lessons for all statisticians interested in effective business applications.

The plan of the rest of this paper is as follows. Section 2 attempts to draw an impressionistic picture of business applications (and absence of applications). I point to traditional areas of business, such as marketing research, where statistics has been used and sketch the statistical tools that I believe to be most relevant for business practice. Then I discuss four topics of special relevance: the government infrastructure, quality and productivity improvement, diffusion of statistical knowledge beyond statisticians and the computing revolution.

In Section 3, I systematically discuss what statistics can hope to accomplish in business and present a framework to suggest how statistics can be applied more effectively. I rely heavily on the idea that a major function of statistics is to help in the continual improvement of products and processes. I discuss the kind of organizational culture and management philosophy that is conducive to effective use of statistics and consider the relationship of statistical thinking to the functions of management, especially the vital distinction between common causes and special causes. I sound a warning about certain misunderstandings of statistical concepts that have become endemic among people exposed to superficial statistical instruction. Finally, I endorse Brian Joiner's idea of the "parastatistician" as a key to expanding the role of statistical applications in business, and go beyond it to suggest that even those who do not become "parastatisticians" can make effective use of statistics. In Section 4, I present a few brief comments on statistical applications in economics. Section 5 considers the possibility of tension between statistical theory and application and possible implications for statistical training and education. Section 6 provides a summary and discussion.

2. AN IMPRESSIONISTIC PICTURE OF BUSINESS APPLICATIONS AND STATISTICAL TOOLS

2.1 The Visible Applications

There are several areas of business in which statistical applications have become conspicuous. Marketing and opinion research, including surveys, panels and focus groups, have been widely used for decades. Recent developments in bar coding and computing have extended marketing research to the study of purchasing behavior in stores. Industrial quality control, including process control and sampling inspection, was widely used during and after the second world war, then languished, and in recent years has had a revival. (In this revival, the role of sampling inspection has been reduced and that of statistical experimentation has been expanded.) In response to regulation of the introduction of new pharmaceuticals, drug companies have made extensive use of statistics, including experimental design and analysis. Experimental design and analysis have been used in research and development in other industries, such as chemicals and, in recent years, the automotive industry and other industries affected by the revival of interest in quality control.

Regulation has spawned statistical applications in industries other than pharmaceuticals. Rate cases for public utilities often require sampling studies and statistical analysis of economic data. Litigation centered on antitrust, affirmative action, environmental impact, securities regulation, occupational safety, and other areas has spawned extensive, protracted and sometimes sophisticated statistical studies. These applications are so specialized that outside consultants rather than employees often do the statistical work.

Statistics is also commonly used for communication and presentation of company information, such as sales reports and accounting information, in annual reports and internal reports and presentations. Simple but colorful "presentation graphics" play a central role.

Economic forecasting and long range planning are pervasive functions, at least in large companies, and have long drawn on statistical methods, especially those associated with econometrics and time series analysis.

Sampling is used for many purposes, though it is often relatively unscientific. Probability sampling is relatively rare, either for marketing surveys or for auditing.

Finally, there is management science/operations research, a specialized function in some companies, that has largely been concerned with optimization techniques such as linear programming, but that has also made substantial use of statistical methodology.

2.2 Statistical Tools

The flowering of statistical theory and methodology in the first half of this century was guided ultimately by the challenges posed by applications, as is well illustrated by the work of Sir Ronald A. Fisher. The mid-century textbooks by Fisher (1954), Tippett (1952), Snedecor (1956) and Yule and Kendall (1947) contained a distillation of important statistical tools that are still relevant today; Snedecor and Cochran (1989) represents a recent updating of this useful line of textbooks. In my judgment, Snedecor and Cochran provide a better coverage of tools of great potential value for business application than do most of the current general textbooks on business and economic statistics.

Since mid-century, there have been many important developments in statistical theory and methodology of special relevance to business application. I now give a brief, personal survey, including selected references to textbooks or treatises; these references are for concreteness only and do not purport to be comprehensive.

1. The development of probability sampling from finite populations (Cochran, 1977) and accompanying techniques for the conduct of surveys.

2. The emergence of ideas of Bayesian decision theory that help to formalize thinking about business decisions and the relationship of these decisions to data (Raiffa and Schlaifer, 1961; Schlaifer, 1959).
3. Developments in time series analysis that have made it possible easily to extract the information needed by management, to avoid widespread popular misinterpretations of autocorrelated time series data and to improve forecasting techniques; (Box and Jenkins, 1976).
4. The increasing focus of experimental design and analysis on problems of business decision making (Box, Hunter and Hunter, 1978).
5. Development of tools for coming closer to sound causal inferences from non-experimental data (Zellner; 1971; Cochran, 1983).
6. The modern movement for quality and productivity improvement, in which statistics is seen as a major contributor to improved management practice (Deming, 1986).
7. Improvement in graphical and tabular techniques for the analysis of data and communication of findings (Tufte, 1983).
8. The recognition that exploratory data analysis is coordinate in importance with confirmatory data analysis (Tukey, 1977; Mosteller and Tukey, 1977).
9. Pooling of information, including empirical Bayes and shrinkage estimation (Mosteller and Wallace, 1964).
10. Simulation, robust inference and bootstrapping (grouped together to suggest that, in my view, these developments individually are of somewhat smaller import for business applications than those described by the first nine headings).

Of these developments, time series analysis has an especially strong claim to attention in business. Many critical business decisions require reasoned interpretation of time series data on sales, market share, quality of product, inventories, work in process, gross margins, profits, promptness of collection of receivables, labor turnover, safety, the price of the common stock, etc. Always lurking in the background are macroeconomic time series of potential relevance to individual companies, such as industry sales, indices of retail sales or industrial production, rates of unemployment and inflation and growth of the Gross National Product. Because business time series data are typically autocorrelated, simple trend and seasonal analysis of the kind developed in elementary textbooks on business statistics have limited value unless extended by autoregression. Moreover, regression of one autocorrelated time series on another, such as company sales on GNP, is very likely to lead to seriously misleading conclusions.

Above I have placed heavy emphasis on the solid core of methodology that has long been available for business applications. I have not addressed innovations, large and small, emerging from continuing methodological research and which are the central interest of many statisticians. The total number of journal pages dedicated to these developments is many times the number of pages in all the core references cited above, which makes it virtually impossible for any statistician to achieve a comprehensive grasp of the methodology that might be relevant to business applications.

The continuing expansion of statistical tools provides a challenge for our best statistical theorists and dissertation topics for the new generation of statisticians. The most relevant of these tools will gradually come into use. But progress in business applications depends much more on effective use of techniques that have long been available than on prompt harnessing of the latest developments, which are often refinements rather than new capabilities.

The one direction of development that may have dramatic implications for business application is that of expert systems. I discuss expert systems very briefly in Section 3.7.

Similarly, the controversies between statisticians on the foundations of statistics or the merits of one analytical strategy over another (for example, the role of nonparametric statistics) seem to me relatively unimportant for business applications. For example, my own orientation is towards the Bayesian rather than the sampling theory approach, but I have observed that theological differences between experienced statisticians rarely have much influence on their approach to business applications.

2.3 Other Related Developments

In addition to the major areas of statistical theory and methodology just sketched, other developments relating to statistics have been important: government statistics, the computing revolution, the movement for quality and productivity improvement based in substantial part on the use of statistics, and the diffusion of statistical knowledge beyond statisticians. I consider each of these briefly.

Governmental statistics: the statistical infrastructure. Of special importance for business application are the advances in government statistics that provide information about the economic and social backgrounds in which businesses operate. Here, although there is room for more progress, the picture is encouraging. An outstanding example is the Current Population Survey, which provides useful current estimates of employment and unemployment. In the 1930s, by contrast, unemployment statistics were more nearly judgmental guesses than statistical estimates, and the dispersion of the guesses was substantial. The *Statistical Abstract of the United States*, prepared by the U. S. Department of Commerce, Bureau of the Census, and updated annually, is an invaluable reference for both economic and demographic information and a guide to many valuable sources of information; it is available through the Superintendent of Documents, U. S. Government Printing Office. In addition to aggregate information, there is a wealth of small-area data, down to the level of block statistics.

Whether or not business firms have generally made effective use of the wealth of government data is less certain. One leading writer on management who has done so, especially in the area of demographics, is Peter Drucker; see, for example, *The Frontiers of Management* (1986).

The movement for quality and productivity improvement. Here I refer to the application in day-to-day practice of ideas of statistical quality control and other statistical methods, most conspicuously in a number of companies in Japan starting in the 1950s but extending in the 1980s to many American companies. Leading statistical writings about this development include, among many others, Deming (1986), Juran (1964) and Ishikawa (1985). This movement deserves to be called a "revolution" because of the widespread participation in statistical training and application by large numbers of employees in the companies that have attempted to implement these ideas. In Japan, statistical concepts such as control chart are not only taught and used in business firms but have even penetrated school curricula, and statistical training of employees by many companies is remarkably extensive. Box (1984) put it this way: "... many millions of people in Japan have some training in the use of statistical methods. These techniques permeate the whole of their industry and are the basic tools for a never-ending incentive towards improvement. Most of the techniques used are very simple but enormous use is also made of designed experiments, many of which are quite complex ..."

It is beyond my scope to assess why this revolution has occurred or to predict how far it will carry, but I can offer a few speculations. In all cases that I know of, the companies that have made major efforts area were initially confronted with serious competitive threats. Not all companies who have embraced "quality" have succeeded; a strong management commitment for the long pull seems clearly to be needed. Statistics alone cannot bring success. For example, employee and supplier involvement is essential, and new production techniques such as JIT (Just-In-Time production, in which parts and components arrive only as needed, with minimal work in process) may be required to exploit the effects of improved quality. Some of the more recent Japanese developments, such as Quality Function Deployment, are not strictly statistical tools but systematic ways of carrying out complex tasks in large organizations.

As I have learned more about the revolution in quality and productivity improvement, I have come to realize that this revolution must be seen as an important component, but only one component, of a broader development in management thinking and business practice. From an engineering perspective, a very similar set of ideas is subsumed under the term "concurrent engineering" (Winner, Pennell, Bertrand and Slusarczuk, 1988). The quality ideas are an essential component, but only one component, of new developments in the theory and practice of manufacturing (Schonberger, 1986; Hayes, Wheelwright and Clark, 1988) and in business management more generally, as reflected, for example, in Peters and Waterman (1982), Peters and Austin (1985) and Peters (1988). Companies caught up in these developments have made sharp changes from standard practices of the past.

Diffusion of statistical knowledge beyond statisticians. It has been true, and continues to be true, that professional statisticians are scarce or nonexistent in most companies. Even in companies with well-established groups of statisticians, statisticians are few relative to the potential problems. Further, statisticians, like other technical professionals, such as engineers, often become involved in work outside their professional specialization, ultimately to end up as managers. It seems likely that in most companies statisticians will be largely involved in special projects of great economic significance, training of nonstatisticians and consulting roles in statistical studies conducted by nonstatisticians, or else they will join the ranks of management.

However, the fraction of business and economic applications done by people who think of themselves primarily as statisticians has been steadily decreasing. This tendency in Great Britain was noted by Baines (1984, page 317), and I have reasons to believe that it is broadly paralleled in the United States: "In the 1970s the number of professionally trained and practicing statisticians in industry ceased to grow and in the past 5 years has declined, particularly in the larger companies." Baines adds: "The most dramatic trend among the practitioners of industrial statistics has been the growth in the number of non-statistician practitioners."

I see a reflection of these trends in the composition of the faculty of the Graduate School of Business at the University of Chicago, where most of today's faculty know more statistical theory and have wider empirical research experience in business and economics than many professional statisticians of a generation ago. This is a reflection of the increased emphasis on empirical research, the need for formal statistical tools in research and the major advances in econometric methodology in recent decades. The same development seems to be taking place in industry as well as in academia. The level of statistical competence in industry, however, may not be as high as in academia. Baines is pessimistic: "This trend will continue with an increasing proportion of statistical analyses in industry being done by scientists, engineers and managers with little or no knowledge of statistical methodology" (Baines, 1984, page 318; Box, in the Discussion, accepts the trend but rejects the prediction).

Some statisticians regard the increased use of statistical tools by nonstatisticians as an erosion of their own influence and importance. The problem of competence of the nonstatisticians, suggested by Baines, is real, but the wider use of statistical methodology seems to me to be desirable. I support the view expressed by Shewhart (1939, page 49): "The long range contribution of statistics depends not so much on getting a lot of highly trained statisticians into industry as it does on creating a statistically minded generation of physicists, chemists, engineers, and others who in any way have a hand in developing and directing the production processes of tomorrow." I would extend Shewhart's view to all areas of business, not just production.

There is, of course, room for an increased role for professional statisticians, a role set forth in detail by Marquardt (1987). But even though the ranks of professional statisticians grow, as Box hopes, and are augmented by "statistically minded" people with primary professional allegiance to other fields, the numbers of statistically sophisticated people will still be small compared to the potential applications of statistics in business. In Section 3.7, I suggest possible ways to meet this problem.

The computing revolution. Interactive statistical computing, with emphasis on analytical graphics, began with the minicomputers of the late 1960s and early 1970s and is now increasingly flourishing in the world of microcomputers, where character graphics are being displaced by high resolution graphics. As a result, it is possible for applied statisticians to draw more effectively on the vast and steadily growing body of statistical theory and methods. The regression computations that required weeks during my study of advertising effectiveness in 1946 would now require a few seconds on a personal computer. The microcomputer revolution of the 1980s has brought new statistical packages and friendlier versions of old ones, and also a wealth of superb supporting software for statistical applications, such as word processing, spreadsheets, graphics and data base managers. The floppy disk has greatly reduced barriers to the movement of data from one computing system to another, a task that seems simple in principle but that is often difficult in practice.

Improved computation is not an unmixed blessing, however; it also makes it possible to misapply statistics more easily and on a wider scale. Moreover, I must confess some disappointment that the rate of growth of useful statistical output from computers seems to lag substantially behind the rate of growth of computing power. Also, it is noteworthy that many of the excellent statistical applications in quality and productivity improvement have been done without the aid of sophisticated computing equipment. Automation is no more a guarantor of improved productivity in statistical applications than in any other function of business.

3. WHAT STATISTICS CAN ACCOMPLISH FOR BUSINESS

In this section I will sketch how statistics is actually applied in business and present a framework to suggest how it can be better applied.

3.1 Objectives of Statistical Applications in Business

My view of what statistics should seek to accomplish in business has been greatly influenced by the developments in quality and productivity improvement, sketched in Section 2 and amplified below in Section 3.4. I see the central objectives to be monitoring and improvement of processes and improvement of product and process design. The terms "process" and "product" are to be interpreted broadly; they refer to services as well as manufacturing. I include both continual, gradual improvement (kaizen) and breakthrough (new products and technologies). The search for these objectives presents challenges for effective use of statistical techniques from the most elementary to the most advanced.

Statistics can also further other objectives. For example, there is need for descriptive summary and reporting of important information. This is generally straightforward, entailing elementary descriptive techniques. Moreover, statistics is a tool of inquiry in basic scientific research that may contribute to company goals, and it facilitates management tasks, such as forecasting, that are only *indirectly* related to process improvement. For example, there is need for statistical support of studies of basic strategic and policy questions, such as long range planning.

3.2 Why Statistics Is Potentially Useful in Business Applications

Statistics provides essential guidance in collection and evaluation of evidence to help to achieve company survival and profitability (long or short run) and to manage the processes that make up a business firm. The statistical tools used to provide this guidance need not be highly sophisticated. Although there are opportunities to apply sophisticated tools (such as experimental design and analysis and multiple time series analysis), simple tools, some of them almost prestatistical rather than statistical, can play an important role. In applications for quality and productivity improvement, for example, Ishikawa (1985) emphasizes Pareto charts, cause-and-effect diagrams, stratification, check sheets, histograms, scatter plots and control charts. Given proper training, almost any literate employee can use these techniques effectively.

More broadly, statistics provides tools for the implementation of scientific method and a language—the language of probability—for thinking systematically about the inevitable uncertainties of the business world. I have argued in Section 2 that differences of opinion about the interpretation of probability have relatively little effect on applied work by statistical professionals. The interpretation of probability in the practical business world is another story. If a business person tries to formalize practical assessments of uncertainty (and not to bypass the problem, as some do, by saying that the future is completely uncertain), probability as degree of belief provides the only anchor point I know. Insofar as statistical theorists press seriously, say, the frequency interpretation of confidence intervals, a virtually unbridgeable gap of communication between statistician and business person will arise. In practice the problem seems moot; in my experience, business people interpret confidence as degree of belief, and it is not likely to be helpful for statisticians to talk them out of it. (A more serious problem arises with tests of hypotheses; I discuss this problem in Section 3.6.)

One other aspect of scientific method is worth mentioning in any overall view of statistics: the general presumption—Occam's razor—in favor of simpler over more complex models of reality, when other considerations are equal. In particular, this presumption reminds us that statistical models need not explain everything in order to be useful in the solution of business problems.

3.3 How the Potential of Statistics Can Be Exploited in Business

It is useful to distinguish six potential specific contributions.

1. Description of past performance. All but the smallest or most casually managed business firms generate written information about past performance, and some of this information is retained, at least temporarily, in data bases. Organization and management of data bases is not primarily a statistical function, but tabular and graphical methods of statistical summarization and display are useful in making data accessible to managers. One such display is a time series plot (or "runs chart") of past performance, such as sales by months for the last five years.

2. Forecasting. (I distinguish unconditional and conditional forecasting: *Unconditional forecasting*: how a process is likely to perform, assuming continuation of current management actions and policies. *Conditional forecasting*: what effect actions of management are likely to have on the future performance of a process. See also paragraph 4 below on the study of cause and effect.) The future is always unknowable in any exact sense, but informed conjectures about the future, even when subject to great uncertainty, are essential. For example, management has to have some idea about likely sales next month or next year for scheduling of manufacturing, ordering from suppliers, transportation, promotion and advertising, recruitment of personnel, etc. Whether or not explicit numerical forecasts of sales are made, forecasting is an inescapable function of management.

Forecasting is often judgmental rather than statistical, in that managers may base forecasts on a judgmental evaluation of all available information rather than on formal statistical computations applied to systematically collected numerical data. But judgmental forecasting, undisciplined by statistical analysis, has weaknesses. Judgments are often unduly influenced by wishful or fearful thinking. Unaided judgment may not properly allow for information provide by the past history of sales; in fact, there is reason to believe that simple statistical forecasts based only on analysis of past sales will often be better than purely judgmental forecasts. (The reason is that the statistical forecasts are automatically calibrated against actuality in the process of identifying and fitting a statistical model, while purely judgmental forecasts are not so calibrated.)

3. Performance evaluation. Management should always know how well a company and its subordinate units are doing. Summarization and description of past performance is an essential need. Long run profitability, either in the accounting sense or in valuation of common stock by the market, is the ultimate test, but intermediate measures of performance such as sales, product quality, customer satisfaction, costs, work in process, inventories, cash flow or personnel turnover often are of practical concern. In focusing attention on these intermediate measures, managers should avoid stressing immediate performance at the expense of long term profitability. For example, overemphasis on current earnings or delivery schedules can lead to underemphasis on quality, and long term profitability may suffer as a result.

For management decisions, knowledge of comparative as well as absolute performance is required. It is not enough, for example, to know current sales. Managers want to know whether or not sales are improving; hence they want to know not only sales for, say, the current quarter, but also sales for the previous quarter and the same quarter a year ago. Similarly, they may want to compare company sales with sales of competitors, the aggregate sales of the economy or a measure of the general price level.

For evaluation of performance, one can compare actual performance against forecast performance; forecasting provides benchmarks for evaluation of actual performance. For example, if current sales are substantially below forecast sales, and if the forecasts are judged to have been sound in the light of available information at the time they were made, direct management attention—trouble-shooting—is appropriate to try to discover an "assignable" or "special" cause that can either be corrected or compensated for.

Performance evaluations of individuals and groups sometimes are used to help to determine compensation. Whether and how it is desirable to take performance into account in compensation policy is more a question of management philosophy than of statistics. The function of statistics is to furnish meaningful and accurate performance evaluations.

4. Study of cause and effect. Insight into causation can help to improve processes. There are two broad approaches: (1) *observational studies*, i.e., passive observation and (2) *experimental studies*, i.e., observation after direct management intervention designed to improve process performance.

Good statistical analysis can sometimes give causal insight even from happenstance or observational data, that is, data arising in the ordinary course of business. Thus if a company's sales decline during a recession, statistical analysis can suggest that part of the sales problem is to be found in economic factors that are effectively beyond the control of any individual company; but if they decline during an expansion, the problem is more likely to be found within the company.

By the use of statistically designed experiments, managements can go a step beyond observational studies. They can intervene to try out new ways of doing things in such a way that the effectiveness of intervention can be assessed statistically. The opportunities for successful experimentation are most obvious in Research and Development (R&D) and manufacturing. In other areas of business, it is less often feasible to conduct statistically designed experiments, but effective use of experimentation has been sometimes made in marketing, as in tests of direct mail copy appeals and time-of-day pricing of electricity. Also, many service areas of business, such as billing or check processing, resemble manufacturing in their opportunities for experimentation. Related strategies such as "evolutionary operation" (EVOP) (Box and Draper, 1969) and intervention analysis (Box and Tiao, 1975) are often useful.

5. Serendipity and outliers. The word "serendipity" refers to accidental discovery of good things. In the process of data analysis, managers often find useful surprises. For example, a single data point may appear to be statistically inconsistent with the rest of the data—a so-called "outlier." Outliers can be viewed as challenges for statistical analysis and opportunities for process improvement. The outlier may simply represent a mistake of some kind, but this should never be assumed. For example, in a study of gasoline blending, one particular blend gave much higher octane ratings than expected in light of the overall pattern of the data. The reported octane rating for that blend was not a mistake, and its discovery and investigation led to large cost savings and the issuance of a patent to the statistician who found it.

6. Sampling in obtaining information. Relevant information for business is never free. Statistical sampling offers an approach to the collection of partial information that is subject to a range of sampling error that may be expressed in probabilistic terms.

3.4 Statistics, Organization Culture and Management Philosophy

Most statisticians understand that statistical theory, though general in its formal structure, must be applied with knowledge of the subject matter of the field of application. A fascinating recent illustration to a quality problem in manufacturing has been given by Bisgaard (1988). If Bisgaard had not been trained in engineering as well as statistics, his efforts to solve the problem would have been fruitless.

But another aspect of statistical application is less well appreciated; at least I have come to appreciate it only in recent years. The movement towards quality and productivity improvement, sketched in Section 2, has led to better understanding of how to make statistical ideas relevant to the needs of managers. Statistical applications cannot be made in an organizational vacuum. Statistical virtuosity and knowledge of subject matter are not sufficient for successful application of statistics to business, and virtuosity often is unnecessary. In the absence of favorable organizational culture, the need for statistics may not be clearly perceived; if statistical work is done, the findings may not be properly implemented.

Leading writers on quality (Deming, 1986; Juran 1964; Ishikawa, 1985) have contributed to the understanding of organizational cultures conducive to effective statistical application. Although I know statisticians who feel that pronouncements on nonstatistical matters such as company cultures go beyond the technical competence of statisticians, even to the point of having an evangelical tone, understanding of organizational culture explains a great deal of the frustration I have long felt about inadequate use of statistics in business. The problem has often been this: since statistical analysis will often lead to conclusions incompatible with past company decisions and practices, it constitutes a threat to those responsible for the past decisions and practices. This threat must be sublimated into a positive search for improvement, without assessment of blame for past mistakes, and that can be done only with a favorable organizational culture. Favorable organizational cultures, however, do not arise spontaneously. They must be shaped by top management. This, I see, is the central message of Deming, Juran, Ishikawa and the others.

Deming and the others point out that the keys to these cultures include: a shared dedication throughout a company to problem solving based on evidence rather than armchair opinion; the recognition that all employees, not just managers and technical specialists, can contribute to problem solving; and the assignment of preeminent importance to pleasing customers, immediate and ultimate. The desirable organizational culture can be sought by the implementation of a management philosophy that shares many common elements with principles advocated by writers on management theory and business practice, such as Peters (1987).

The philosophy of Deming and the others also ties into statistics. The key link is the notion of a *process*: a particular method of achieving a desired goal, usually involving a series of more or less well-defined steps. Businesses and other organizations can be thought of as bundles of many interrelated processes, such as sales, manufacturing and assembly, employee training and customer service. Processes typically go on repeatedly, and the outcomes can be measured and recorded. For example, we can observe the percentage scrap per day in production of a particular molded rubber part. There is no real distinction between processes in manufacturing and in other areas of business, such as finance, marketing, purchasing, accounting, engineering, and research and development, or between processes in businesses and in other organizations.

According to the philosophy, management's central task is the continuing attempt to improve processes, and hence overall company performance, and to use statistics effectively in doing so. Statistics is seen as a set of tools and a way of thinking that are useful in improving processes in all areas of business. (The aim is not an attempt to improve *all* processes *simultaneously*. Rather, priorities must be set by what Juran has called the "Pareto principle," which means that one aims first at processes for which improvement will make the greatest contribution to overall company profitability. Thus priority is given to the "vital few" rather than the "trivial many.")

Statistics is potentially useful because process improvement usually can not be achieved solely by armchair reasoning, with reliance on general technical knowledge and skills. Much depends on the evidence bearing on each individual application, and statistics is a tool for collecting and interpreting the evidence. The stress on statistics can be viewed as a significant updating and revision of the scientific management philosophy founded by Frederick C. Taylor. Taylor's great insight was to see that scientific method is a powerful tool for improving processes, but the systematic use of statistics in implementation of scientific method became possible only after Taylor's time. Since Taylor's career ended at about the time when the career of Sir Ronald A. Fisher was beginning, Taylor did not know very much about efficient implementation of scientific method, which has been the central contribution of Fisher and the other statisticians who have shaped the development of modern statistics. Brian Joiner has pointed out (personal communication) that Taylor's scientific study of metal cutting, which extended over more than two decades, might have been accomplished in much less time had he known of the ideas of statistical experimentation that Fisher would introduce.

Taylor's approach was essentially authoritarian: he assigned the study of process improvement to specialists. The leaders in the movement for quality and productivity improvement (Deming and the others) believe that it is essential to involve all employees—managers, foremen, workers—in the improvement of processes. This management philosophy is essentially democratic and thus antithetical to Taylor's conception of management. As Brian Joiner has put it, scientific method (that is, statistics) cannot be successfully applied in an organization unless the organization is dedicated as "all one team" to the application of statistics to improvement of all processes.

3.5 Statistical Thinking and Management

Business statisticians often say, correctly in my view, that "statistical thinking" is more important for management than specific techniques. Although the precise meaning of "statistical thinking" may vary somewhat from one statistician to another, I believe that all would agree that statistical thinking requires familiarity with ideas such as the regression phenomenon, sample selection, location and variability, distributions, conditional probability and the hazards of measurement error in data, ideas that are useful even in the absence of formal statistical evidence. The regression phenomenon makes it clear, for example, that a decline from top ranking to third ranking of a sales territory from one year to the next is not necessarily an indication of poorer sales performance. Statistical thinking includes also an attitude expressed by Winston Churchill's injunction "to pay no attention to anything but realities" (Churchill, 1948, page 468). Finally, as emphasized especially by George Snedecor (Wallis, 1970), statistical thinking should be a *habit*, not just an exercise reserved for certain formal statistical studies.

I see statistical thinking as central to effective management: statistics is the glue that binds together the functional areas of business. Among the writers on quality and productivity improvement, Deming (1986) has most clearly spelled out the connection of statistical thinking and management. In particular, he emphasizes that understanding of variability is central to management, and that statistical training is needed for proper understanding of variability.

For example, nonstatisticians who view time-series measurements such as stock price changes from day to day, yields of a chemical process in a state of statistical control or the batting performance of a major league baseball player are very likely to misunderstand what they see. They are prone to contrive *ex post* explanations of extreme observations, even if in fact these extremes are consistent with simple chance models. Nonstatisticians are likely to perceive "hot streaks," "cold streaks" or a "shift of momentum" in athletic performances that are essentially in a state of statistical control (that is, behaving as realizations of IID random variables). Equally, nonstatisticians may discern complicated systematic patterns in stock market data that conform to a random walk and believe that these patterns have predictive value.

Managers who do not understand variability will have similar difficulties in understanding what is happening in their companies. Even though annual performance ratings (say, growth of sales across sales territories) may effectively behave as drawings from a lottery, they will be prone to reward the lottery winners and punish the losers, with adverse consequences for employee morale. If the average performance level of a process is short of an arbitrary target, these managers often assume that exhortation and sloganeering will close the gap. They will tamper with a process when there are cogent statistical reasons to believe that the tampering will only make the process worse. Deming (1986) calls such tampering "overadjustment" and cites examples where the substitution of benign neglect for tampering has led to large quality improvements and cost savings. On the other hand, when there is statistical evidence suggesting that in a true "out-of-control" situation, benign neglect is not appropriate; active investigation and corrective action are in order.

Deming's (1986) distinction between common and special causes of variation is also essential for clear management thinking. Special causes are transitory and localized, while common causes are inherent in the process. For example, if errors in filling out a form are traceable to a single employee who was ill on the day the form was explained, we have a special cause. On the other hand, if most or all employees contribute errors, we have a common cause: inadequate instruction in how to fill out the form, or a poor form.

Special causes can be corrected by prompt *ad hoc* action, by workers (if they are permitted to do so) as well as by foremen, managers or engineers. Common causes usually demand management attention and a fundamental study of how the process really works. It is counterproductive to blame workers or supervisors for unsatisfactory performance stemming from common causes. Skilled statistical analysis can help to sort out special causes and common causes; this was one of the great contributions by Shewhart (1931) to statistical process control. Regression analysis can also aid in the understanding of common causes even for processes that are in a state of statistical control by relating process performance to other variables that may have been measured.

Interpretation of the latest data point—last month's sales, the most recent closing stock price or the yield of a production batch—brings out clearly the need for statistical thinking by managers. Given a statistical model and analysis thereof, a statistician can assess where the point lies in the predictive distribution based on statistical analysis of past performance and how the point alters the predictive distribution. A manager not understanding statistical thinking can only guess and pontificate. Only a manager who understands variability can understand how statistical analysis can help to put the latest data point into proper perspective.

Without this statistical framework, reactions to current information are typically confused, as is illustrated by comments of government spokespersons about the economic relevance of the latest macroeconomic data on unemployment, the trade deficit or retail sales; or the explanation by market analysts that yesterday's stock market decline was explained by "profit taking." Clever *ad hoc* explanations can be and usually are advanced, no matter what has happened. Only those understanding statistical variation can understand that such *ex post* "explanations" of the last data point are suspect unless they can be supported by systematic statistical analysis.

The need for background knowledge in the interpretation of data is not confined to background knowledge in statistics. There is need for background knowledge in economics to understand better the implications to the firm of events in the industry and the overall economy, to suggest variables that can be used in statistical studies aimed at better evaluation of company performance and to aid in the formulation of statistical models. Equally, background knowledge in empirical science helps one to look for explanations of performance failures that go beyond the assignment of blame to particular workers or supervisors or such convenient excuses as "union restrictions" or "inadequate educational background."

3.6 Danger: How Statistical Concepts Can Be Counterproductive

I have argued for the relevance and usefulness of statistics in business. A warning is also necessary: misunderstandings and misuses of statistics in business are common and serious.

In my view, the most serious misunderstanding is interpreting tests of sharp null hypotheses as business decision procedures. It is very difficult to find a realistic business decision that is clarified by formulation as a statistical test. For example, the decision as to whether a proposed process modification should be adopted has nothing to do with the null hypothesis that the modified process has *exactly* the same yield (or quality) as the standard process. If utility is linear in yield, the question is whether the mean of the posterior distribution of the difference in yield—often equal to the estimated difference in yield—exceeds an economic breakeven point. For a decision-theoretic formulation of the problem, see Schlaifer (1959) or Raiffa and Schlaifer (1961). Tests do play a role for the statistician: they can be useful for diagnostic checks and simplifications of tentative statistical models (Box, 1980), but significance levels are irrelevant to the manager who must make the business decision.

A closely related misunderstanding is the view that diagnostic checking of a statistical model, or choice between two competing statistical models, can be reduced to examination of a single test statistic, or a small number of test statistics. The two major approaches to exploratory data analysis (Tukey, 1977; Box, 1980) place a common stress on graphical analysis of data. In this context, it is interesting that three of Ishikawa's "Seven Management Tools" (Ishikawa, 1985) are graphical: histogram, scatter plot and control chart. A fourth, the Pareto chart, is a special graphical display; a fifth, "stratification," refers to cross-tabular displays.

Sometimes there is no need to make a choice between apparently competing statistical models, since it may be possible, and advantageous, to embed the models in a broader model that includes the individual models as special cases.

There are other pervasive misunderstandings that I have often encountered among nonstatisticians who have had some exposure to statistical instruction.

1. It is often believed that an approximately normal histogram means that the process producing the data is in a state of statistical control in the sense of independently and identically distributed variables (IID). More broadly, there is the failure to recognize that examination of histograms and summary statistics provides a very incomplete perspective on any data that arise originally as a time series.

2. There is a common impression that the R-squared statistic is the most important output of a multiple regression analysis and that there is some level of R-squared below which the regression can be pronounced unsatisfactory. Of course, the R-square statistic has useful interpretations, both in the sense of "variance explained" (a suggestive expression that is unfortunately subject to misinterpretation) and in its relationship in Bayesian formulations of posterior odds ratios for tests of the hypothesis that the regression coefficients are all zero.

3. It is often believed that statistical procedures can be safely applied, and standard inferences drawn, without diagnostic checking of the assumptions or specifications of these procedures. For example, if the data comprise percentage defectives in successive production runs, then it is often assumed, without any examination of the data, that "P-charts" (based on the assumption of binomial variation within production runs) are the appropriate tool for statistical analysis.

4. It is often believed that statistical inference is exclusively concerned with inferences about parameters rather than predictions about future observations.

5. The importance of operational definitions and careful measurement is seldom sufficiently appreciated.

Finally, although statisticians think of statistics as a potentially powerful tool to aid management in making good decisions and improving processes, statistics can be misused, through ignorance, wishful thinking, or intent, to distort reality and create false impressions. Statistical misuses are most evident in political campaigns, but, in my experience, they are equally common in business.

3.7 Statisticians and Parastatisticians

Whatever the area of investigation, there are two essential links in a statistical application: the relevance of the statistical findings to the problem and the statistical support for the findings themselves. Experience suggests that neither link is easy to forge.

The soundness of the first link, relevance to the problem, turns on common sense and subject matter knowledge as well as statistics. Thus the causal interpretation of a female salary shortfall in a regression study in a sex discrimination case must be based on information beyond that of the data being analyzed. Statistical principles alone cannot settle the question of whether discrimination has occurred, although they can help to delineate the potential effects of omitted variables on the regression outcome.

The soundness of the second link, the statistical implementation, is clearly within the domain of technical statistics. Evaluation of soundness raises questions of appropriate statistical models, the basic approach to testing or estimation, including possible approximations and diagnostic checking of model adequacy. A full consideration of these questions is a job for a professional statistician, working in close association with subject matter specialists, not an amateur. Almost every application presents unique challenges that, if pursued, lead to at least a minor modification of standard, off-the-shelf methodology.

But, from an economic perspective, only very important business applications warrant full examination of the technical statistical questions and the time of one of the relatively small number of professional statisticians (within or outside companies) who consult on applications. Time and cost constraints dictate that at least a degree of improvisation is necessary even for an experienced and well-trained statistician. The hope in most business applications is that simple statistical models will be good approximations to reality; that these models can be analyzed by standard methods, preferably implemented by readily available statistical software, to extract most of the relevant information; and that further refinement of analysis would not change the practical conclusions. One good illustration is the widespread and successful use of the linear discriminant function in the selection of credit risks even though the assumptions of the standard model are violated by the data; for example, ordinarily some of the classification variables are categorical, even dichotomous, so that the assumption of multivariate normality for these variables is violated.

The further hope is that the statistical tools can be applied, though admittedly with fear and trembling, by well-trained nonstatisticians ("parastatisticians" in Brian Joiner's felicitous terminology), possibly backed up to a limited degree by the opportunity to consult with professional statisticians. If I am right in this belief, there is an opportunity to expand the range of statistical applications in business by wider employee involvement, that is, by training parastatisticians and encouraging them to apply their training to company problems. In quality and productivity improvement, some companies in both the United States and Japan have shown that this wider involvement is possible, given an appropriate company commitment to statistical training and the appropriate organizational culture for statistical problem solving.

(There is, of course, no reason that a parastatistician cannot develop a fully professional stature through time. Arnold Zellner has reminded me that most of the great builders of modern statistics, for example, Sir Ronald A. Fisher and Sir Harold Jeffreys, were not trained as statisticians. Ordinarily, however, Joiner and I have in mind a more modest aim.)

One can think of parastatisticians either as resource persons within an organization, as does Joiner, or as managers with statistical capabilities that can be drawn upon in doing their jobs. In either event, the aim is to facilitate more effective use of statistics within an organization.

The feasibility of relying substantially on parastatisticians turns on the correctness of my view that a relatively small core of simple statistical techniques will be adequate for a wide range of business applications. (I do not, however, claim that a similar small core of simple techniques would be applicable to areas of statistical application other than business or even to all areas of business.)

I do believe that the core must be larger than Ishikawa's Seven Management Tools, however useful these may be. With the aim only of conveying what I have in mind, and not of trying to spell out details, I refer again to the coverage of the textbook by Snedecor and Cochran (1989). With some supplementation from applied time series techniques, mainly approached through regression and autoregression, parastatisticians could contribute usefully to many applications, applications that in many instances would never have a chance for statistical illumination.

The reason that so much can be done with so little is that many business problems can be viewed from the perspective of the standard regression model, a perspective made clearer to nonstatisticians by the fact that the assumptions can be illustrated and checked by simple graphs. It is natural for theoretical statisticians to dwell on the limitations of the regression model and to devise ways to overcome these limitations, including, for example, generalized least squares and robust estimation. But the standard regression model, including even the specification of normality, often serves as a rough description of business reality. Failures of the standard model may be a signal for substantive investigation rather than more sophisticated statistical modeling. Nonnormality can point to special causes that should be removed. The presence of heteroskedasticity may be an opportunity to learn more about process variability, since variability is often as important as level.

In suggesting the possibility of basing practical statistical work on a small core of techniques, I do not mean to preclude the possibility of going beyond that core when the skills of the analyst permit and when substantial improvements are in prospect. In major statistical studies of important problems, it will indeed be desirable to bring to bear the same degree of sophistication that is characteristic of academic research. Similarly, when smaller studies are done repeatedly, refinements of technique are likely to be cost effective. In these circumstances, the core techniques may serve only for the first statistical reconnaissance of the data. Thus, for example, a public utility like Commonwealth Edison can approach forecasting at different levels: purely judgmental, trend lines, regression models, simultaneous equation models and dynamic simultaneous equation and time series models.

Purely as a concrete illustration of what I have in mind, and with full realization that there is ample room for improvement, I refer to the techniques used in my own elementary book, *Data Analysis for Managers with Minitab* (Roberts, 1987, 1988), which are unified by ordinary least squares. The coverage of that book has been shaped by my aim, discussed in Section 1, to have students do individual projects entailing effective use of statistics for company problems. In developing the book over many years, I tried out the core techniques (implemented either by Minitab or a similar minicomputer package called IDA) on a wide range of business data sets from serious applications, with the aim of seeing whether I could *quickly* capture the essential practical message conveyed by these data sets, using only the tools accessible to my students. My impression is that it is indeed possible to do so in a large majority of the applications, and that my students, with some guidance and supervision from me, have also been reasonably successful. Most of the data sets presented special features that would have invited refinement of these rough and ready analyses, even the development of novel techniques, but I doubt that the practical conclusions would often have been substantially altered.

There are areas in which I am uneasy when I am trying to see how far my core tools for rough and ready analysis can be extended. Some examples: applications in which errors in (independent) variables, missing observations, transfer-function relationships, shocks with lagged effects or omitted variables are potentially important. I have ways of thinking about and grappling with such problems, but it is easy to pass over the border between a satisfactory rough and ready analysis and an analysis that misses something important that is happening in the data.

In thus evaluating my experience, I am of course grading my own papers, and I would not blame a reader for being skeptical about my claims of success. My purpose is to urge other statisticians to try to replicate my experience, using, of course, the core techniques that they find most useful, which may differ substantially from mine.

Practical skill in using even a small core of statistical techniques is not an automatic byproduct of training in theoretical statistics, and it may not be attainable in the time allocated to statistical training by most companies. But I believe that many employees, including especially middle managers, can benefit from statistical training at the level exemplified by *applied* statistics courses in undergraduate and graduate business schools. (For additional discussion of statistics in business schools, see Easton, Roberts and Tiao, 1988.)

Even if the majority of business applications can be handled effectively (though hardly ever optimally) by the simple techniques that parastatisticians can master, there will also be applications that require more sophisticated treatment. For example, a parastatistician with a working understanding of regression by ordinary least squares may encounter a problem calling for logistic regression. In some instances, but not always, use of ordinary least squares with a dichotomous dependent variable may suffice to meet the practical requirements. But it is well for parastatisticians to have backup support from professional statisticians. If the number of parastatisticians increases, there will also be need for expansion of the number of supporting statisticians, especially those with the needed subject matter expertise.

There is another reason for an expanded base of professional statisticians: the need for expert assistance at the planning or design stage of statistical applications. Statistical design is usually harder than statistical analysis. Given adequate instruction, almost any literate person can make a control chart; not anyone can figure out how often the process should be sampled. Anyone can design a randomized two-group experiment; not anyone can select a good two-level fractional factorial design.

This discussion of statisticians and parastatisticians would be incomplete without some consideration of statistical computing packages and expert systems. The development of computing packages has made business applications easier for all, professional statisticians and amateurs. Not only have these packages extended the range of techniques and the possibilities for display of results, but they have to some degree automated statistical analysis, as is illustrated by the package AUTOBOX (Shumway, 1986), which offers users the choice of automatic identification of appropriate ARIMA time series models. Efforts are under way to advance from such automation of special statistical procedures to expert systems, in which the program helps the user to emulate the line of attack of expert statisticians. Successful development of expert systems could do much to extend the range of effective applications in business.

As with computation in general, however, I am cautious about expert systems. One reason is that the difficulty of defining statistical problems, which requires, in addition to statistical expertise, intimate knowledge of the particular subject matter at hand. Another is that expert systems are likely to try to duplicate the expertise of a statistician in the use of sophisticated methodology that may be unnecessary for solution of the practical problem and hard to understand without being an expert oneself. I therefore see a continuing need for the rough and ready analytical capabilities outlined above. Perhaps these capabilities themselves could be embodied in an expert system dedicated to quick clarification rather than to emulation of an expert statistician, a system that might provide warnings when rough and ready methods do not suffice.

Beyond Parastatisticians. A remarkable book by Richard Schonberger (Schonberger, 1986) has convinced me that even the idea of "parastatistician" places unnecessary restrictions on the potential scope of statistical applications in business. Schonberger surveys modern developments in manufacturing management, which extend well beyond the points stressed by statisticians interested in quality and productivity improvement; for example, they include just-in-time production (very little work-in-process), flexible manufacturing (quick changeovers), multiplicity of machines (as opposed to one or a few very large machines), new methods of cost accounting, linearity of scheduling (producing at a steady rate per day and using extra time for maintenance and problem solving) and cellular organization (compact layouts for manufacture of each individual product). These developments are unified by simplicity and elimination of waste. They entail enhancement of the role of line operators and emphasis on staff as support for line operators rather than as master planners and analysts.

These developments have important implications for the role of statistics in manufacturing, and, by extension, in other areas of business. The following excerpts convey something of Schonberger's position.

> ...The jobs of everyone in the factory must be changed. Most of the line jobs were direct labor (operator or assembler), nothing more or less. The new line jobs are direct labor plus a variety of indirect duties—like preventive maintenance—plus some activities that have *always been done by managers and staff specialists.* I refer to *data recording, data analysis,* and *problem solving.*
>
> Data recording comes first. The tools are cheap and simple: pencils and chalk. Give those simple tools for recording data to each operator. Then make it a natural part of the operator's job to record disturbances and measurements on charts and blackboards. The person who records the data is inclined to analyze, and the analyzer is inclined to think of solutions. Success depends on recording the right kind of data at the right time.
>
> One approach is for operators to record a piece of data each time there is a work slowdown or stoppage. The vital piece of data to be captured is the *cause* of the slowdown or stoppage. [pages 18–19]

> We need a simple, natural way to wean factory people off a diet of pure direct labor and onto a mixture of direct and indirect duties. The way to begin is with charts and graphs. ... [page 36]
>
> Solving problems, continually and rapidly, is everyone's business. Collecting the ... data has been thought of as the front-end, routine part of problem-solving. The vital core, the glamorous part, has been analysis and decision-making, which has required the talents of well-paid staff experts.
>
> Neglect of the front end [line operators] has been a chief reason why Western industry has done so poorly at problem-solving. ... The line operator measures and records problem data. In so doing, the operator naturally reflects on the data and tries to diagnose the trouble ... [and] sometimes come[s] up with obvious, simple, commonsense solutions—the best kind.
>
> ... A central, not a peripheral, role of staff people is to be on call. [pages 39–40]
>
> The computer does have a bright future in the factory. That future is mostly in *direct process control* and not so much in information to support staff and management. [page 45]

These developments in manufacturing have obvious implications for the hope expressed in my opening paragraph of this paper for more applications of statistics, and they support my broad definition of a "business application of statistics" as any use of statistical reasoning in pursuit of company goals. They suggest enormous opportunities for statistics, usually simple statistics, as a pervasive tool in business. They call for prompt analysis by operators of current information as opposed to leisurely analysis by staff of past data from computerized data bases or company archives. They lead to immediate actions and brief summaries, perhaps just notations on a control chart, rather than formal reports. They lend themselves to simple statistical tools such as run charts and control charts, evolutionary operation (EVOP) (Box and Draper, 1969) and intervention analysis (Box and Tiao, 1975). Computers may not always be necessary.

Changes in business practices of the kind now occurring may impact the use of statistics or other quantitative tools in business. For example, electronic links between manufacturers and mail order customers may permit prompt transmission of information to suppliers about current sales to *ultimate* customers, thus lessening the need for statistical forecasting by the suppliers. Simplified manufacturing processes may permit direct experimentation rather than computer simulation to investigate the effects of possible changes. Diminished emphasis on traditional inspection of shipments from suppliers may greatly reduce the need for acceptance sampling.

4. COMMENTS ON STATISTICAL APPLICATIONS IN ECONOMICS

As I stated at the beginning, my primary focus is on business rather than economic applications of statistics. However, my discussion applies directly to economists working within business firms on company problems.

The empirical research that is essential to economics as a scholarly discipline raises additional questions, but there are some parallels to the issues I have raised about business applications. In economic research as in other scientific fields, there has been a substantial expansion of both theoretical and applied statistical work by researchers who do not think of themselves primarily as statisticians. The discipline of econometrics is a conspicuous example. So are the specializations and offshoots of economics of special interest to business: finance, marketing, government regulation, industrial organization and labor. In these fields, the level of statistical sophistication among economists is often very high. And the empirical work that they have done has often been of very high quality and has influenced business practice, as in the instance of research in finance on efficient markets.

I shall offer impressions about statistics in economic research that are much less grounded in experience than the impressions I have offered about business applications of statistics. The most important of these is that many theoretical economists are unconcerned about the need to confront theory by data. At the University of Chicago, where I have spent virtually all my professional life, the dedication of economists to empirical validation of economic theory—ultimately, if not immediately—is strong and essentially universal. But dedication to empirical testing of economic theory is not uniform in the economic profession (Morgan, 1988).

Among the economists who work with data, I have the impression of generally high statistical standards, certainly by comparison with the other applied fields to which I have had some exposure, mainly behavioral science, medicine and health care and law. But I also have the impression of a certain narrowness of statistical outlook in empirical economic research: the view, criticized in Section 3.6 in the context of business, that empirical testing of scientific hypotheses is equivalent to the statistical testing of sharp null hypotheses. (I hasten to say that economists are not alone among scientists in holding what I consider an important misconception.) What empirical economists ultimately accomplish, in my view, is the formulation and estimation of statistical models. In doing so they are of course guided by economic science and limited by the inability to design randomized experiments. As Friedman (1953, page 40) put it,

> The necessity of relying on uncontrolled experience rather than controlled experiment makes it difficult to produce dramatic and clear-cut evidence to justify the acceptance of tentative hypotheses. Reliance on uncontrolled experience does not affect the fundamental methodological principle that a hypothesis can be tested only by the conformity of its implications or predictions with observable phenomena; but it does render the task of testing hypotheses more difficult and gives greater scope for confusion about the methodological principles involved. More than other scientists, social scientists need to be self-conscious about their methodology.

But economists, like researchers in other areas of scientific application, often tend to assume that statistical null hypotheses can be directly identified with scientific hypotheses and to emphasize sophisticated test statistics for elaborate null hypotheses over informal diagnostic checks and data analysis. The lack of emphasis on data analysis often leaves me uneasy about appropriateness of the statistical model within which the test statistic is formulated. A null hypothesis includes a statistical model specification that makes an economic theory operational. The conclusion about the null hypothesis depends on the model specification as well as the theory. In sophisticated economic research the evidence on the model specification itself is often examined via a test statistic. But in the absence of simple data analysis, I am uneasy about putting so much reliance on a test statistic.

Over the years I have learned a useful way of quickly reading empirical economic studies. First, skim the initial theoretical section, which sets forth, usually in mathematical language, the economic theory to be "tested." Second, read carefully the empirical section, which often includes multiple regression, not necessarily confined to ordinary least squares, applied to available data in order to "test" the models. From the empirical section, I can see how close the statistical model comes to the economic theory. But because of the emphasis on test statistics and p-values, I have much greater difficulty in forming a judgment as to how well the statistical model is supported by the data.

Some of my difficulty centers on reservations about the standard tail-area testing procedures; this could be alleviated by reformulation in terms of Bayesian odds ratios. Some of the difficulty centers on the frequent assumption that asymptotic distributions give adequate guidance; this could be alleviated by more work on finite-sample distributions. Some of the difficulty centers on the emphasis on inferences about parameters; this could be alleviated by more work on the predictive distribution of the observations. But my basic concern is the tendency to rely on high-tech inferential procedures and pay too little attention to the simple methods of data analysis, including especially careful examination of data plots.

There are also encouraging developments in economic applications. There is an increased emphasis on predictive tests based on forecasts derived from economic or econometric models. There is increased awareness of the problem of quality of data in general and on the limitations of data for particular studies. For example, comparisons of real income through time may be affected by biases in price indices; study of the health of American manufacturing may be affected by the details of construction of the national income accounts; and measurement problems may make less obvious than appears on the surface such seemingly simple questions as to whether the United States is a debtor nation or whether the U.S. savings rate is lower than that of other countries.

5. THEORY, APPLICATION AND EDUCATION

Any discussion of statistical applications, even though focused on business and economics, should consider the field of statistics in general and statistical education and training in particular. A clear delineation of the issues is provided by Anscombe in the discussion of Wallis's paper (1980) on a highly successful experience in statistical applications, "The Statistical Research Group, 1942–1945." First, an extended comment by Anscombe (Wallis, 1980; page 331):

> In some ways the subject of statistics is flourishing. . . . Yet I think many of us are aware that not evrything is right with . . . statistics. Consider the literature. The journals are ever growing in size and number, and there are ever more new books. Much of this literature is highly technical and narrow in scope, written apparently only for the two or three other specialists in the topic referred to, presenting a forbidding aspect to any more general reader unwise enough to tackle it. On the other hand, the elementary textbooks in their profusion mostly reflect absence of fresh thinking, understanding or experience. Like many other sciences, statistics has developed greatly since, say, the end of the Second World War, but it has not developed in the profound and spectacular way that biology and astronomy and pure mathematics have developed. Current research appears to proceed in many unrelated directions. Some major movements in the subject can be perceived, but there are some dreadfully many minor movements going on as well.
>
> One of the contributory causes of all this confusion is the usual practice that persons who obtain a PhD immediately go into teaching. There they are expected to be excellent teachers and to be highly productive in research, but one thing is not expected in the academic world, namely that the research should be relevant to anything; that would violate academic freedom. I believe our subject would be in better shape if we could return to a former tradition, if it were the usual practice for the most exciting new PhD's to spend several years in a research team that had some definite mission, before, perhaps, reentering the academic world. Fisher developed the science of statistics at Rothamsted, not in a university.

Next is part of Wallis's reaction (1980, page 334):

> I am disposed to defend, though not without limit, irrelevance in academic research. In some ways, the most valuable contribution that universities can make and one that is unique to universities is to support research and scholarship that are not subject to tests of relevance. Research that is clearly relevant is supported by many other institutions, including profit-dependent and vote-dependent institutions. But the life of the mind

has an intrinsic value of its own. At the least, its products are relevant to a clearer or deeper understanding, if only of logical implications and relations. And who can say what may be the relevance and practical applications of intellectual understanding and of the processes through which it is obtained? ... [Wallis then discusses, as an illustration, contributions to the Statistical Research Group of work by distinguished theoreticians Wald, Wolfowitz, and Savage, and the influence of "highly technical and academic work" by Von Neumann on Abraham Wald's development of sequential analysis.]

These remarks to the contrary notwithstanding, I agree strongly with Anscombe's emphasis on the value in statistical training of fairly deep exposure to one or more fields in which people are struggling to unlock nature's secrets and are interested in statistics not for its own sake but as an implement. Indeed, the preceding remarks buttress his position.

It is hazardous to try to add to two such excellent statements, but I must express the opinion that the overwhelming thrust of statistical education has concealed from students, at all levels from elementary courses to advanced doctoral study, the vital role of statistics and data analysis in real (as opposed to imagined) applications in business and economics. There is a little evidence of recognition of this problem within academia (Easton, Roberts and Tiao, 1988), but I am not optimistic about prospects for internally generated change given the current incentive structure in universities.

I look for help from another direction, the "other institutions" mentioned by Wallis. The increasing use of statistics and emphasis on statistical training in business firms responding to increased international competition—for example, Ford and General Motors—is one example. The proliferation of short courses and seminars is another; although not all these offerings may be of high quality, they do have to meet a relevancy test, a test that tends to weed out poor quality as well as irrelevant content. Encouragement of applied statistics by the Japanese Union of Scientists and Engineers (JUSE) and other organizations in Japan is also heartening. Developments like these may send the right signals back to academia. See also my discussion of parastatisticians in Section 3.7.

Another interesting possibility for help might come from organizations such as the Center for Business and Economic Statistics (CBES) that has been proposed by a group of statisticians led by Arnold Zellner. Zellner points out (1988, page 4):

...To achieve the interrelated goals of improved data generation, data bases and effective use of improved data in applied modeling, forecasting and decision making as well as the goals of achieving improved modeling, forecasting and decision-making techniques, ... top flight theorists working in close conjunction with very talented applied researchers are required. ... [The CBES] program incorporates a coordinated effort by applied and theoretical researchers from the areas of statistical measurement methodology, statistical time series modeling and forecasting, econometric modeling, forecasting and policy-making, business and economic theory and practice, statistical experimental design, marketing analysis, behavioral science, statistical product quality management and improvement and computer science. The CBES organizational structure will facilitate and coordinate interactions among theorists, applied researchers and users, and provide for prompt, thorough testing of new research methodologies and speedy implementation of those that are found to be dependable and reliable.

6. SUMMARY AND DISCUSSION

Statistical methodology has great potential for useful application in all areas and at all levels of a business firm. That potential is seldom realized. For the most part, statistical applications in business, such as marketing research surveys, have met intermittent, specialized needs and have not been made an integral part of day-to-day problem solving and decision making. They have been isolated from the management process.

However, in recent years in the United States and for a longer time in Japan, some companies have effectively exploited statistical tools on a broad scale in quality and productivity improvement, guided by the objective of continual improvement of all processes and products. For the most part the statistical tools have been simple ones, such as histograms, control charts and scatter plots, but even these simple tools have helped greatly in improving a wide range of processes within the companies. Often these simple tools have been applied by nonstatisticians, trained at company expense. A company culture that accepts evidence over unsupported opinion seems essential. I believe that the experience of these companies offers great hope and specific guidance for all business applications of statistics, not just those thought of as "quality control."

Statistical and probabilistic thinking, even in the absence of data, is essential for sound decision making. For example, a key task of managers at all levels is to distinguish special causes from common causes of variation in the output of processes. Only with understanding of statistical variability can managers intelligently direct efforts to improve processes and avoid the tampering that can make them worse.

Statistics can be used most effectively in business when large numbers of employees—in principle, all—have some grasp of elementary tools. Fortunately, there is evidence both that these elementary tools can be grasped and that they are useful, at least for the rough-and-ready studies that suffice to iluminate most business problems and facilitate most decisions. Beyond the most elementary tools, the body of methods unified by the standard regression model suffice to cope with many applications, provided that they are applied with awareness of the need for informal data analysis, including generous reliance on graphics, to check the appropriateness of model assumptions and to locate unusual features, such as outliers or nonconstant variance, that may convey essential information about the business process under study. There is also need for sophisticated methodology, including especially modeling of single and multiple time series, econometric modeling, design of experiments and sample surveys.

Statisticians have an important role to play in business applications, but they cannot carry out, or even supervise, all of them. Rather, the statisticians will mainly have to consult with nonstatisticians and to concentrate their own statistical work on advanced applications. Beside the statisticians, however, a veritable army of "parastatisticians" is needed. These parastatisticians can turn to statisticians for specialized support and backup, but for the most part they are on their own. And neither statisticians nor parastatisticians can be effective unless there is full support by top management for statistical problem solving and for the kind of company culture in which statistical reasoning can be effective.

There is a danger that common misunderstandings of statistics may slow or even misdirect the movement towards more effective use of statistics for business applications. One pernicious misunderstanding is the belief that tests of sharp null hypotheses have any direct connection with the decisions faced by the management of business firms. A closely related misunderstanding, which sometimes affects empirical research in economics as well as practical applications in business, is that the belief that the fundamental objective of a statistical investigation is a test statistic or a p-value, or even an R-square. A more fundamental misunderstanding is the idea that any concept of probability other than degree of belief is relevant to a decision maker dealing with the uncertainties of the business world.

The development of modern computing has enormously raised the potential for statistical applications in business and economics, but computing power per se, like any form of automation, will not automatically lead to enhanced performance.

The current incentive structure in universities has led to much less emphasis on business and economic applications of statistics than is warranted by the potential of statistical methodology. Possible corrections may come from special statistical centers, like the one proposed by Zellner, and from the example set by institutions outside academia, such as business organizations, that make effective use of statistics and offer statistical training of their own.

REFERENCES

BAINES, A. (1984). Present position and potential developments: Some personal views. Industrial statistics and operational research (with discussion). *J. Roy. Statist. Soc. Ser. A* **147** 316–326.

BISGAARD, S. (1988). The quality detective: A case study. Report Number 32, Center for Quality and Productivity Improvement, University of Wisconsin-Madison.

BOX, G. E. P. (1980). Sampling and Bayes' inference in scientific modeling and robustness. *J. Roy. Statist. Soc. Ser. A* **143** 383–430.

BOX, G. E. P. (1984). Comment on "Present position and potential developments: Some personal views. Industrial statistics and operational research" by A. Baines. *J. Roy. Statist. Soc. Ser. A* **147** 325–326.

BOX, G. E. P. and DRAPER, N. R. (1969). *Evolutionary Operation: A Method for Increasing Industrial Productivity.* Wiley, New York.

BOX, G. E. P., HUNTER, W. G. and HUNTER, J. S. (1978). *Statistics for Experimenters: An Introduction to Design, Data Analysis, and Model Building.* Wiley, New York.

BOX, G. E. P. and JENKINS, G. M. (1976). *Time Series Analysis: Forcasting and Control,* Revised ed. Holden-Day, San Francisco.

BOX, G. E. P. and TIAO, G. C. (1975). Intervention analysis with applications to economic and environmental problems. *J. Amer. Statist. Assoc.* **70** 70–79.

CHURCHILL, W. S. (1948). *The Gathering Storm.* Houghton Mifflin, Cambridge, Mass.

COCHRAN, W. G. (1977). *Sampling Techniques,* 3rd ed. Wiley, New York.

COCHRAN, W. G. (1983). *Planning and Analysis of Observational Studies.* Wiley, New York.

DEMING, W. E. (1986). *Out of the Crisis.* MIT Press, Cambridge, Mass.

DRUCKER, P. F. (1986). *The Frontiers of Management.* Harper and Row, New York.

EASTON, G., ROBERTS, H. V. and TIAO, G. C. (1988). Making statistics more effective in schools of business. *J. Business and Economic Statist.* **6** 247–260.

FISHER, R. A. (1954). *Statistical Methods for Research Workers.* Hafner, New York.

FRIEDMAN, M. (1953). The methodology of positive economics. In *Essays in Positive Economics.* Univ. Chicago Press.

HAYES, R. H., WHEELWRIGHT, S. C. and CLARK, K. B. (1988). *Dynamic Manufacturing.* Free Press, New York.

ISHIKAWA, K. (1985). *What is Total Quality Control?* Prentice-Hall, Englewood Cliffs, N.J.

JURAN, J. (1964). *Managerial Breakthrough.* McGraw-Hill, New York.

MARQUARDT, D. W. (1987). The importance of statisticians. *J. Amer. Statist. Assoc.* **82** 1–7.

MORGAN, T. (1988). Theory versus empiricism in academic economics: Update and comparisons. *J. Economic Perspectives* **2** 159–164.

MOSTELLER, F. and TUKEY, J. W. (1977). *Data Analysis and Regression.* Addison-Wesley, Reading, Mass.

MOSTELLER, F. and WALLACE, D. L. (1964). *Inference and Disputed Authorship: The Federalist.* Addison-Wesley, Reading, Mass.

PETERS, T. (1988). *Thriving on Chaos.* Harper and Row, New York.

PETERS, T. and AUSTIN, N. (1985). *A Passion for Excellence.* Harper and Row, New York.

PETERS, T. J. and WATERMAN, R. H., JR. (1982). *In Search of Excellence.* Harper and Row, New York.

RAIFFA, H. and SCHLAIFER, R. (1961). *Applied Statistical Decision Theory.* Division of Research, Harvard Business School, Boston.

ROBERTS, H. V. (1987). Data analysis for managers. *Amer. Statist.* **41** 270–278.

ROBERTS, H. V. (1988). *Data Analysis for Managers with Minitab.* Scientific Press, Redwood City, Calif.

SCHLAIFER, R. (1959). *Probability and Statistics for Business Decisions.* McGraw-Hill, New York.

SCHONBERGER, R. I. (1986). *World Class Manufacturing.* Free Press, New York.

SHEWHART, W. A. (1931). *Economic Control of Quality of Manufactured Product,* Van Nostrand, New York. (Republished in its entirety in 1981, with a dedication by W. Edwards Deming, by the American Society for Quality Control, Milwaukee, Wis.)

SHEWHART, W. A. (1939). *Statistical Method from the Viewpoint of Quality Control.* Dover, New York.

SHUMWAY, R. H. (1986). AUTOBOX (Version 1.02). *Amer. Statist.* **40** 299–300.

SNEDECOR, G. W. (1956). *Statistical Methods,* 5th ed. Iowa State Univ. Press, Ames, Ia.

SNEDECOR, G. W. and COCHRAN, W. G. (1989). *Statistical Methods,* 8th ed. Iowa State Univ. Press, Ames, Ia.

TIPPETT, L. H. C. (1952). *The Methods of Statistics,* 4th ed. Wiley, New York.

TUFTE, E. R. (1983). *The Visual Display of Quantitative Information.* Graphics Press, Cheshire, Conn.

TUKEY, J. W. (1977). *Exploratory Data Analysis.* Addison-Wesley, Reading, Mass.

WALLIS, W. A. (1970). Statistics in nonstatistical contexts. Special Address Invited by the President of the American Statistical Association at its 130th Annual Meeting, Detroit, 27 December 1970.

WALLIS, W. A. (1980). The Statistical Research Group, 1942–1945 (with discussion). *J. Amer. Statist. Assoc.* **75** 320–335.

WINNER, R. I., PENNELL, J. P., BERTRAND, H. E. and SLUSARCZUK, M. M. G. (1988). The role of concurrent engineering in weapons system acquisition. Institution for Defense Analyses, Alexandria, Va.

YULE, G. U. and KENDALL, M. G. (1947). *An Introduction to the Theory of Statistics.* Charles Griffin, London.

ZELLNER, A. (1971). *An Introduction to Bayesian Inference in Econometrics.* Wiley, New York.

ZELLNER, A. (1988). Center for Business and Economic Statistics. Internal report, Graduate School of Business, The University of Chicago.

Comment

George Box

Reading this stimulating paper, I was particularly intrigued by the section on Theory, Application and Education. As Harry Roberts points out, experience with the modern quality movement highlights many important issues. It underlines the necessity for a much wider definition for theoretical statistics. It makes explicit the necessarily iterative nature of investigation as exemplified by the Shewhart Cycle, by the goal of *never-ending* improvement and by the complementary roles of Tukey's exploratory and confirmatory data analysis. It demonstrates how simple graphical techniques, such as Ishikawa's Seven Tools, although ludicrously simple mathematically, can be enormously powerful scientifically because they are devoted to the inductive step of hypothesis *generation* that mathematical statistics so often ignores.

Most subjects have a theoretical as well as an applied side, and ideally each nourishes the other; but for statistics I believe this has not always been true. In my view, statistics has no reason for existence except as the catalyst for investigation and discovery. If this is true then, above all else, the proper study of the statistician is scientific method and therefore statistics should serve the needs of that study. An understanding of the process of investigation involves such things as the roles of induction as well as deduction, the nature of scientific learning, the importance of subject matter knowledge, the psychology of investigators and the management of data acquisition and experiment. The *theory* of statistics should be concerned with all these things.

Unfortunately, its domination by mathematics has led to the teaching and propagation of ideas that I believe are in some cases actually antithetical to good statistical practice. Consider the process of investigation itself as exemplified, for example, by any good detective novel, by any reasonably honest account of scientific research (such as Watson's account of the discovery of the double helix) or by the process of finding out why a manufacturing system is producing low quality product. This process of investigation employs induction and deduction in an iterative

George Box is Director, Center for Quality and Productivity Improvement, University of Wisconsin–Madison, 610 Walnut Street, Madison, Wisconsin 53705.