(Reprinted with permission from
International Journal of Forecasting, 4 (1988), 449-465)

Research Needs In Forecasting*

J. Scott ARMSTRONG
The Wharton School, University of Pennsylvania, Philadelphia, PA 19104-6371, USA

Abstract

The demand for research on forecasting is strong. This conclusion is based on the high number of citations to papers published about research on forecasting, and upon the number of subscriptions for journals devoted to forecasting. The supply of research papers is also large, following a rapid growth in the 1960s and 1970s. This research has produced important findings. Despite this, a comparison of published research versus the needs expressed in two surveys of academics and practitioners showed that numerous gaps still exist. A review of the literature also supported this conclusion that the research being produced does not match up well against the research desired. Suggestions are made as to what research is needed and how it should be conducted.

Keywords: Acceptance of forecasts, Forecasting audit, Legal aspects, Survey of forecasters.

Introduction

Most cultures have been concerned with forecasting. Sometimes the forecaster was held in high regard, as was the oracle at Delphi. Often, however, forecasting is regarded as a necessary evil and is frowned upon. According to a current sage (Drucker, 1973, p. 124), " . . . forecasting is not a respectable human activity and not worthwhile beyond the shortest of periods." Sometimes it has been illegal. For example, in Rome in 357 A.D. Emperor Constantine issued an edict forbidden anyone " to consult a soothsayer, a mathematician, or a forecaster... May curiosity to foretell the future be silenced forever," In recent years, however, forecasting seems to have become a respectable activity. Research of forecasting has also gained respect. This paper examines research on forecasting. It includes all aspects of forecasting in the social sciences.

The first section of this paper deals with the demand for research on forecasting. An exa mination of the supply of research on forecasting follows. Discussions are then provided on what research should be done, and how it should be done. The conclusions in this paper are based on a 1986 mail survey of members and former members of the International Institute of Forecasters, a survey administered to attendees at the International Symposium on Forecasting in 1987, a quantitative analysis of publications, and a review of the literature.

The Demand for Research on Forecasting

A formal forecast is not needed for all decision-making. Alternatives include insurance, hedging, control, and flexibility. Still, forecasting serves some definite needs; it can help us plan for the future and make more rational decisions. Research should enable us to do a better job of forecasting.

How strong is the demand for *research* on forecasting? One way to assess this is to study the usage of forecasting research by other researchers. I did this by examining citation rates. Unfortunately, these rates are not available by subject area. However, an interesting statistic is that, in its first two years, the *Journal of Forecasting*

---

* The Gunning Fox Index for this paper is about 11. *Acknowledgements:* Ruth Pagell conducted the computer search and the literature counts. Wende Gladfelter and Kenneth Weissman helped with the coding. The readability indices were calculated by Kathy Armstrong and Stuart Neuman. Among the many people who provided comments on earlier drafts were Stuart Bretschneider, Robert Fildes, Baruch Fischhoff, Wilpen Gorr, Shelby H. McIntyre, and Herman O. Stekler.

was the seventh most frequently cited journal among the 84 journals in business, management, and planning, according to the "citation impact" factor *(Journal Citation Reports,* 1983, 1984).

The demand for forecasting can also be inferred from the subscriptions to forecasting journals. Three journals on forecasting have been founded since 1982. None existed prior to that year. The number of subscriptions increased rapidly for the *Journal of Forecasting;* it was founded in 1982, and by 1984 there were over 1,700 subscriptions. The *International Journal of Forecasting,* founded in 1985, had over 1,000 subscriptions by 1987. The *Journal of Business Forecasting,* which is directed to practitioners, was founded in 1982; in 1987, its circulation was 3,500.[1]

In addition to subscribing to the journals, researchers and practitioners *read* them. Respondents to the mail survey of members of the International Institute of Forecasters (IIF) in April 1986[2] reported reading an average of 2.43 papers per issue from the *Journal of Forecasting (JoF)* and the *International Journal of Forecasting (IJF);* this was 27% of the nine papers published in a typical issue. The reported readership was a bit higher for practitioners (2.7 papers per issue) than for academics (2.2), or consultants (2.4).

The respondents said that they *used* the research from the *IJF* and *JoF* Of practitioners, 36% reported making applications, half of which they said were successful. Of academics, 27% said they cited the work in their own research, 19% were led to do additional research on a topic, and 25% used material from the journals in their teaching.

Another way to assess demand is to examine courses on forecasting. Cerullo and Avila (1975), in a survey of the *Fortune* 500 companies, reported that 98% of the respondents thought that forecasting should be taught at business schools. Hanke (1989), in a survey of the American Assembly of Collegiate Schools of Business (34% response rate), found that 62% of the responding schools offered a course on forecasting as of 1987.

The supply of research on forecasting

To assess the supply of research on forecasting, I examined the number of books published on forecasting. This included all books with forecasting as a subject. (Later editions were counted as new books.) Information on

---

[1] Personal communication with Professor C.L. Jain, editor of the *Journal of Business Forecasting,* 19 May 1986.

[2] A mail survey was sent to a total of 1,370 members and former members of the International Institute of Forecasters (IIF) in April 1986. Of this mailing, 908 were sent to the US, 119 to Canada, and 343 to other countries. Sixty-three of the packets were returned due to incorrect addresses, and another sixteen were returned due to a mail embargo in Finland. There were 202 returns. After deducting for the incorrect addresses, the return rate was over 15%. The return rate was low because the questionnaire was sent out as part of a packet including five items. Also, no monetary incentive or follow-ups were used. This low rate of return does not pose a particular problem for many of the responses reported here because they are used only indicate the views of some experts, not to generalize about a class of people. The breakdown of the respondents is as follows:

Field	Practitioners	Academics
Economics	6	19
Finance and Accounting	10	14
Marketing	31	9
Planning	30	4
Production OR/MS	12	14
Research & Development	10	11
Unidentified/ Miscellaneous	16	16
Totals	115	87

this was obtained from the U.S. Library of Congress (see the description of the LC MARC data base[3]). 1 used the original publication dates of all books that dealt with forecasting as a subject. The results are presented in Exhibit 1.

Exhibit 1
Books published on forecasting: Long-term growth*

Years	Forecasting books per year	Per 1,000 published****
(1900-1939)	2	—
1940s	7	0.2**
1950s	14	0.3
1960s	38	0.5
1970s	223	1.9
1980s***	336	2.6

*Search of LC MARC as of February 16, 1988.
**Based on interpolation using total books published from 1947-49.
***Based on 1980 through 1986 (only preliminary data were available for 1987).
****The number of forecasting books in relation to all books published.

The publication of books on forecasting appears to be a recent phenomenon. From a modest start in the 1940s, growth was rapid, especially in the 1960s and 1970s. The number of forecasting books published was also growing rapidly relative to all books published. By the 1980s, of each 1,000 books published, approximately 2.6 were on forecasting.

The rapid rise of book publication in the 1970s can also be seen in the second column of Exhibit 2, where annual figures are presented. (The information for 1986 is preliminary; the final figure will be higher as books continue to be entered into the data base.) Note, however, that the annual publication rate leveled off in the 1980s.

The number of papers published on forecasting was also examined. Three sources were available (see Note 3 for a description of the sources). The most relevant series is the *Social Science Citation Index* (SSCI) because it includes only academic journals. This showed a rapid growth in the 197(followed by a leveling off in the 1980s (see third column of Exhibit 2). ABI/Inform, which include academic as well as practitioner-oriented papers, followed a

[3] The computer data base search was conducted on January 30, 1988, with an update of LC Marc on February 16, 1988. A brief description of each source was abstracted from the *Dialog Database Catalog:*

ABI /INFORM (August 1971-present, Data Courier, Inc., Louisville, KY). This is designed to cover all phases of business management and administration. Approximately 550 primary publications in business and related fields are scanning for articles to be abstracted in included in ABI/ INFORM.

LC MARC (1968-present, U.S. Library of Congress, Washington, D.C.). This contains bibliography records for all books cataloged by the U.S. Library of Congress since 1968.

MANAGEMENT CONTENTS (September 1974-present, Management Contents, Inc. Northbrook, IL). This provides information on a variety of business and management-related topics to aid individuals in business, consulting firms, educational institutions, government agencies or bureaus, and libraries. Articles from over 700 U.S. and international journals, as well as proceedings, transactions, business course materials, newsletters, and research reports were indexed and abstracted to provide information in the areas of accounting, decision sciences, finance, industrial relations, managerial economics, marketing, operations research, organizational behavior, and public administration.

SOCIAL SCISEARCH (1972-present, Institute for Scientific Information, Philadelphia, PA). This multi-disciplinary database indexes items from the most important social science journals. It covers every area of the social and behavioral sciences, and relevant articles selected from 3,000 additional journals in the natural, physical, and biomedical sciences. It also includes many important monographs. It corresponds to the printed Social Science Citation Index.

similar pattern. *Management Contents* with a much broader practitioner coverage, showed continuing growth. Judging from the wide annual variations, the reliability of the measures is suspect. Unfortunately, the coverage of journals varies over time. Nevertheless, the SSCI and ABI/Inform each provides evidence of growth in the 1970s the number of academic research papers. *Management Contents* indicates growth in the 1980s papers for practitioners.

Exhibit 2
Publications on forecasting

Year	Books	Journal Articles		
		SSCI*	ABI/INFORM**	Management contents***
(1)	(2)	(3)	(4)	(5)
1969	79	—	—	—
1970	95	—	—	—
1971	109	—	21	—
1972	139	62	57	—
1973	181	100	98	—
1974	215	123	100	—
1975	246	143	127	349
1976	323	121	163	363
1977	292	152	209	270
1978	317	172	545	272
1979	305	172	440	381
1980	319	172	411	433
1981	334	156	365	485
1982	348	154	266	504
1983	388	142	279	612
1984	357	155	322	706
1985	349	154	280	N.A.
1986	254_p	176	396	N.A.

* Includes all articles with the world "forecasting" in the title.
** Includes al articles with forecasting as a subject or in the title.
*** Includes all articles with forecasting as a subject.
$_p$ Preliminary: the final figures will be substantially higher.
N.A. Number of journals covered decreased greatly so that 1985 and 1986 figures are not comparable. [See Note 3 for a description of the data base.]

While the number of journal articles on forecasting was growing in the 1970s, so were those in other areas. According to the SSCI, the growth rate of publications for all subjects was about 6% per year in the 1970s. Interestingly, there has been little growth in the 1980s. Overall, the growth rate for research in forecasting has been similar to that for research in all of the social sciences in the 1970s and 1980s.

Topic areas for research

Within this overall trend, one might ask what particular areas have shown the greatest growth examined this by taking a 10% probability sample of the papers on forecasting that were listed in t SSCI from 1972 to 1985 and categorized them according to the field of study. The sample size papers) did not allow for a detailed examination of trends. Still, it was noted that none of the sample of 24 papers from 1972 to 1976 dealt with financial forecasting, while eight of the 53 papers from1977-1985 did (statistically significant at $p < .05$). This change in financial forecasting may have been influenced by a proposal by the U.S. Securities and Exchange Commission with respect to mandatory forecasts. (The proposal was not passed.) During that same interval, papers on forecasting in marketing went from 0 to 6, and on politics from 0 to 4 (neither change was statistically significant). In general, there seems to be a growing number of papers in the applied areas.

What research is needed?

One way to determine what research is needed is to ask experts. I took a direct approach to this including the following questions on the April 1986 survey of IIF members (this group includes people who prepare forecasts as well as those who do research in forecasting):

> What are your priorities for research on forecasting? Assume that you were asked to oversee budget of $250,000 for the next five years ($50,000 per year). This budget has been earmark for forecasting. Assuming that you have good personnel, a good library, good computer support, and other resources. Also assume that several organizations have indicated a willingness to have you test your ideas in their organizations. Your role in this project is to define what problems should be studied. List here, in order to priority, those problems that you would study. Please be specific. What questions would you address? How would you study them?

It is interesting that 38% of those who returned the 1986 mail questionnaire did not respond to th4 questions about research needs. (Two respondents said that although they did not have an opinion what research should be done, they did not think that $250,000 would be enough.) The results from the 113 respondents to this item were summarized by the field of the respondent and also by whether the respondent was a practitioner or an academic (see Exhibit 3). For example, among those respondents in marketing, four people proposed research projects that would help in the implementation of forecasts. The most interesting finding from the answers to this question was the wide variety of topics among the more than 220 ideas that were suggested.

Exhibit 3
Selected research needs identified by respondents
(Based on 113 respondents, i = number of ideas suggested).*

Field	Practitioners	Academics
Economics	casual model; survey research; (i = 7)	causal models (2); uncertainty; (i = 21)
Finance & Accounting	environmental forecasting; seasonal variations; (i = 9)	expert system (2); uncertainty; (i = 6)
Marketing	Implementation (4); computerization (3); combine methods (2); competitive actions; evaluation; (i = 19)	Incorporate judgment in models (2); competitive actions (2); combine forecasts (2); compare alternative methods; implementation; (i = 13)
Planning	impact on decision-making (4); expert systems (2); judgmental forecasting (2); computerization; compare alternative models; implementation; scenarios; uncertainty; (i = 30)	compare alternative methods; monitor forecast; (i = 7)
Production MS/OR	new product forecasting (3); combine methods; quality of data vs. method; seasonality; (i = 10)	combine alternative methods (3); uncertainty (2); combine forecasts; compare alternative methods; (I = 4)
Research & Development	new product forecasting (20; outliers (2); causal models; computerization; (i = 10)	compare alternative methods (2); combine alternative methods; impact on decision-making; scenarios; (i = 16)
Other areas	experts systems (2); compare alternative methods; impact on decision-making; implementation monitor forecasts; (i = 10)	compare the alternative methods (2); quality of data vs. method (2); impact on decision-making; scenarios; uncertainty; (i = 14)

Implementation was the most frequently mentioned area for the preceding question. As expected, implementation and related areas such as "impact on decision-making" are more commonly mentioned in the practitioner column than in the academic column of Exhibit 3. Implementation is a broad term referring not only to the implementation of new methods, but also to gaining acceptance of the forecast. How can the forecaster convince the decision-maker that the forecast is correct and that it should be used? This issue is also implied by the response to the question of whether important forecasts were affected by "politics" in their organization. Of the 115 practitioners who responded to this item on the 1986 and 1987 surveys,[4] about half (48%) said that the forecast was subject to the influence of politics.

The use of direct questioning to assess this issue provided a crude assessment. It seemed that respondents had not given much prior thought to what would be the best areas to research. In-depth questioning of the experts or focus-group interviews might be helpful for further assessments of this issue.

[4] A survey was enclosed in the packets of the 559 registrants at the International Symposium on Forecasting in Boston in May 1987. Replies were received from 53 people. See also note 2.

Comparing research needs with supply

Another way to assess what research needs to be done is to compare what has been published versus the expressed interests of the experts. For the latter, I used the interest-level ratings provided by practitioners and academics who responded to the IIF surveys in 1986 and 1987 (See footnotes 2 and 3). Although there were three times as many respondents for the 1986 survey as for 1987, the results from each survey were weighted equally; this reflects the belief that other sources of error are more serious than sampling error.

The results, summarized in Exhibit 4, indicate a strong interest in applications not only by practitioners, but also by academics. The papers by the IIF journals seem to be meeting these needs. Of the 215 papers published to date, almost half dealt with applications. This includes all papers coded as describing applications or potential applications to real-world situations. Furthermore, this orientation to applications has been increasing over the years from a total of 36 papers during 1982 through 1984, to 67 in the 1985 through 1987 period. In percentage terms, that is an increase from 41% of all papers published in the journal in the early period to 67% in the latter.

Exhibit 4
Comparison of interest vs. papers published

	Interest ratings * (1 = lo; 5 = hi)		**Papers published in IIF journals Covering 1982-87**
	Practitioners	**Academic**	
Applications	3.9	3.9	103
Methods			**
??Econometrics	3.8	3.4	56
??Evaluation	3.7	4.0	15
??Uncertainty	3.6	3.6	4
??Gaining acceptance of forecasts	3.5	2.6	3
??Computer methods	3.4	3.2	0
??Judgment	3.3	3.3	15
??Extrapolation	3.1	3.3	73

* Unweighted average of 1986 and 1987 surveys.

** For "methods," each paper was classified according to primary focus. This coding was done by the author and by a research assistant. Commentaries were not counted and some papers fell outside of these categories.

Substantial differences exist between what is desired and what is published with respect to methodology. This occurs not only for practitioner interests, but also for academic interests. The forecasting journals publish many papers on extrapolation, which is the least interesting area for the practitioners. This occurs despite an exceptionally high rate of rejection for papers on extrapolation. In contrast, some high-interest areas such as evaluation and uncertainty receive little attention in the literature. For a recent exception, see the paper on uncertainty by Makridakis et al. (1987); this study provides evidence that formulas that had long been recommended for the assessment of uncertainty in forecasting provide poor estimates.

The interests of academics and practitioners show some agreement (r = .38). The most significant difference is that practitioners are much more interested in research on ways to get forecasts accepted. This is as one would expect. It also is consistent with the results in Exhibit 3.

Trends in the literature

Still another way to approach the need for research is to examine what areas have been studied fruitfully in the past. This draws upon the extensive literature review that I have conducted (e.g., Armstrong, 1984, 1985, 1986).

Exhibit 5
Conclusions from the review of the literature

Topic	Prior research			Future research
	Effort	Payoff	Some conclusions	prospects
Decomposition (Predict parts, then aggregate)	Little	Low	Helpful where uncertainty is high	Good
Extrapolation	Very great	Moderate	Seasonal factors useful Trends should be dampened Minor differences in accuracy among methods	Modest
Intentions	Modest	Modest	Ways to reduce response and non-response bias	Modest
Expert Opinion	Great	High	Limited value of expertise in forecasting large changes Role playing accurate for conflict situations	Excellent
Expert Systems (Bootstrapping)	Modest	Very high	Less expensive & a bit more accurate	Excellent
Causal Methods (Econometric)	Very great	Moderate	Simplicity is a virtue Econometric methods better for large changes	Modest
Combined Forecasts	Modest	Modest	Combination yield substantial gains	Modest
Uncertainty	Modest	Modest	Judgmental estimates are typically overconfident Argue against your forecast	Good
Implementation	Little	Modest	Scenarios can help to gain prior commitment	Excellent
Audit Process	Little	Little	Guidelines needed	Excellent

My conclusions are summarized in Exhibit 5. This summary is an updated version of the one published in Armstrong (1986). The ratings have been revised in light of research published over the past three years.

The areas of greatest promise, according to Exhibit 5, are in expert opinion, expert systems, implementation, and auditing. Some of the possibilities for research in each area are discussed below.

Expert Opinion: The study of defects in expert opinion has made great strides over the past decade. Of particular importance is the work by Tversky and Kahneman (e.g., see Kahneman, Slovic, and Tversky, 1982). Fischhoff (1976) provides evidence that the judgmental findings pertain not only to current events but also to forecasts. The challenge recently has been to find ways to overcome some of the defects (Christensen-Szalanski, 1986; Fischhoff, 1988).

One neglected area that seems to be of particular interest is how to forecast outcomes in conflict situations. For example, what would happen if South Africa went to a one-person/one-vote scheme? Or, how would company X respond if company Y drastically reduced the price of a competitive product?

An especially promising technique for forecasting in conflict situations is role playing. It is expected to help because it provides better insight into the opponent's thinking. This technique was not mentioned by any of the respondents in the IIF surveys as a promising area for the research budget. Furthermore, role-playing has not been widely used. For example, of the 81 respondents to the 1986 and 1987 IIF surveys who claimed to know how their organization forecasted the actions of its competitors, only four (5%) checked "role-playing" as one of the techniques used.

My studies and my review of the literature suggest that role-playing offers significantly more accurate forecasts than can be obtained by alternative methods in situations that involve conflict. In the seven actual situations studied to date, subjects who used unaided opinions were able to predict the outcomes correctly 20% of the time. In contrast, role-playing was correct on 70% of the predictions (Armstrong, 1987). Role-playing predictions were more accurate in six of the seven situations (there was one tie).

Expert Systems: In this paper, the term "expert systems" for forecasting refers to all procedures for translating judgmental forecasts into an explicit, reproducible format that is independent of the expert. It can be done by paper and pencil, or by computer. The creation of an expert system can aid one's understanding of the judgmental forecasting procedures used in a given situation.

Although expert systems is an area of enormous interest currently, little of the effort has been directed toward predictive validity. Do expert systems produce more accurate forecasts than currently-used methods? Do people use the forecasts from expert systems? Most of the validation research on expert system has been done under the term "bootstrapping." Good evidence exists to support expert systems as being low-cost and accurate, relative to unaided judgment (Armstrong, 1985, pp. 274-284). Still, we need to learn more about the best way to develop an expert system and under what conditions it is most useful.

Implementation: Implementation appeared to be a key concern in the survey of IIF members (see Exhibit 3), in the comparison of supply and demand (Exhibit 4), and in my assessment of the literature (Exhibit 5). One implementation topic of particular interest is how to gain acceptance of a forecast.

Forecasts on important topics often meet resistance. Consider the following. Would an accurate forecast have been accepted by the U.S. Supreme Court in *Brown vs. Board of Education?* That is, would it have affect the decision that led to school busing? Stephan (1978), in a follow-up of the predictions provided to the Supreme Court by social scientists, found that many of the forecasts used to support the decision were inaccurate. (He claimed to be among those who had made poor forecasts.) One forecast was that school integration would reduce racial prejudice. That did not occur. Apparently, when people with racial differences are placed in a competitive situation, such as school, it does not reduce prejudice. Would a forecast stating that integration would not decrease prejudice have been accepted and used by the Supreme Court in its decision on this case?

The few studies on implementation indicate that a forecast is often ignored by decision-makers if it implies actions that are contrary to those initially favored by management (e.g., see Griffith and Wellman, 1979).

An alternative to ignoring the forecast is to manipulate it. This is more likely when the forecast is a matter of public record. This was shown to occur in macro-economic forecasts prepared by the government (Brouthers, 1986) and in election polls (Shamir, 1986).

One way to gain acceptance of forecasts is to use scenarios. The scenario involves converting the forecast into a story about the future. Recent research has suggested ways to write scenarios in order to gain acceptance of unusual of threatening forecasts. The use of vivid examples in the scenario is one way to attract attention. Another is to ask the decision-maker to write a scenario where she is a participant in the scenario. Such approaches are likely to lead the decision-maker to take unusual forecasts more seriously. A summary of the research on scenarios is provided in Armstrong (1985, pp. 38-45).

The way in which the forecasts is made can influence decision-making. For example, forecasts are more likely to have an impact if the decision-maker is involved in the forecasting process. An interesting application of this is to ask people to forecast their own behavior. They tend to forecast that they will act in a rational and socially responsible manner. Then, when faced with similar situations, they try to fulfill their predictions (Sherman, 1980, and Greenwald et al., 1987). On the other hand, this involvement is likely also to bias the forecast.

Finally, the use of computers can have a great influence on implementation. Procedures for good practice can be built into programs. For example, many forecasting packages now contain useful procedures for validation. I expect, in the future, these programs will guide the user through the model development process. Many aspects of the forecasting process can be automated.

Auditing the Forecasting Process: A growing interest is expected in ways to audit forecasting. This is particularly interesting because of the attention paid recently to legal issues in forecasting. Some of the more important legal cases are discussed here.

An error led to a forecast of a severe drought in the Yakima Valley in Washington. This caused farmers to undertake expensive actions. When the predicted drought did not occur, farmers took legal action against the U.S. Bureau of Reclamation to recover their losses. The farmers claimed that it was not really a forecast; they claimed that it was merely a calculation. The government did admit that a mistake had been made. However, they had been under no contractual agreement to provide the forecast; as a result, the court ruled against the farmers (*Schinmann vs. U.S*., 618 F. Supp. 1030, September 18, 1985). This decision was upheld by the appellate court (*Schinmann vs. U.S*., unpublished opinion, U.S. Court of Appeals for the Ninth Circuit).

In Brown vs. U.S., four Massachusetts fishermen were lost at sea on November 21, 1980 because, their families claimed, of an incorrect weather forecast. Three families brought suit and won an initial judgment on the ground that the National Weather Service was negligent in failing to repair a weather buoy that might have provided useful data. The decision was overturned by the First U.S. Circuit Court of Appeals, and the U.S. Supreme Court refused to take the case (*Brown vs. U.S*., 599 F. Supp. 877 (1984), 790 F.2d 199 (1st Cir., 1986)). The key issue in this case was not the fact that the forecast was wrong, but whether the National Oceanic and Atmospheric Administration (NOAA) failed to take reasonable steps to obtain accurate data. Another issue was that when key information was no longer available, NOAA did not notify the users of the forecast. The court ruled that the government was not required to report on the process used. This ruling implies that it is reasonable for a forecaster to make tradeoffs between the cost of the forecast and its benefits. A ruling against NOAA might have had a chilling effect on forecasters.

In a British case, *Esso Petroleum vs. Mardon* (London, 1966 E. No. 2571), Mardon entered into a contract with Esso to own and operate a gas station. A critical part of the negotiations was the forecast that the station would sell 200,000 gallons of gas per year by the third year. The actual sales fell well short of the forecasted figure, and Mardon went out of business. Esso sued Mardon for unpaid bills. Mardon then countersued on the basis that the Esso forecast misrepresented the situation. In effect, Esso had originally forecast the 200,000 gallon figure under the assumption that the gas pumps would face the road. After a zoning hearing, they were forced to change the design so that the pumps were not visible from the road. Despite this significant unfavorable change, Esso then used the 200,000 gallon forecast in the original contract with Mardon. Mardon won; the court concluded that Esso misrepresented the situation.

In a case that is currently being tried, *Beecham vs. Yankelovich*, Beecham alleged that an inaccurate market forecast prepared by Yankelovich Clancy Shulman resulted in a $24 million loss. Yankelovich, on the other hand, claimed that Beecham provided incorrect inputs to the forecasting models, and that they failed to follow the marketing plan; for example, they changed advertising claims and reduced promotional expenses *(Adweek's Marketing Week*, 7 December 1987, pp. 1, 4). The suit, if successful, could expose forecasters to a broad range of lawsuits.

To date, only a few cases have been of direct relevance to forecasters. They imply that if you do not have a contract to provide forecasts, you are unlikely to be held liable. Furthermore, the courts recognize that forecasts contain uncertainty; reasonable attempts to balance costs and benefits should provide good protection. Finally,

forecasters can be held liable if it can be shown that the forecasts were not obtained by reasonable practice, and only if the poor practice was intentional so as to bias the forecast.

In addition to their use in legal cases, an agreement on good standards of forecasting would also be useful in auditing public projects. For example, does the government use adequate procedures to forecast the outcome of various projects for mass transportation, nuclear power plants, synthetic fuels, or convention centers? One approach to such an audit was demonstrated by Stewart and Glantz (1985). They audited the U.S. National Defense University's forecast of climate change to the year 2000. They concluded that the climate change forecasts were suspect because the study was not well-designed, particularly in its failure to follow good practices for judgmental forecasting.

How should the research be carried out?

What research strategies are relevant to forecasting? This section examines historical trends in research strategies, and then describes strategies for future research.

Historical development of research on forecasting

Before 1960, research on forecasting was characterized by descriptive studies that were mostly verbal. During the 1960s the descriptions became more mathematical. In the 1970s, there was an emphasis on empirical research. The 1980s have shown much attention to validation research. The M-competition (Makridakis et al., 1982) provides a well known example of validation research in forecasting.

Ideally, validation research calls for empirical comparisons to assess alternative hypotheses. Such research provides an excellent foundation for scientific advances. To assess the use of validation research in forecasting. I coded papers from the *Journal of Forecasting* (1982 to 1985, Issue 1) and for the *International Journal of Forecasting* (1985-1987). This showed that almost half of the papers were empirical. Of these, 58% used the method of multiple hypotheses (see Exhibit 6). These results compare favourably with papers in *Management Science*, where only 22% of the empirical papers used the method of multiple hypotheses (Armstrong, 1979). Thus, the forecasting field strongly emphasizes scientific validation.

Much can be learned by seeing how different disciplines and different organizations have attacked the same problem. To assess the extent to which the research is coming from different perspectives, I examined the papers published in the *Journal of Forecasting* (1982 to 1985, Issue 1) and the *International Journal of Forecasting* (1985-1987). As one indication of diversity, 42% of the papers had at least one author from outside the U.S. Also, there was some diversity in the institutional affiliations; while 76% of the papers were by authors from universities, 10% had ties to non-profit organizations of government, and 14% with for-profit organizations (see Exhibit 6).

Exhibit 6
Research by type and source (entries are number of papers)

Types of research*	Author's organization***			Totals (%)
	Academic	Non-profit	Profit	
Theory	74	11	10	95 (44)
Review	14	1	0	15 (7)
Empirical**				
No hypothesis	12	1	3	16 (8)
Single hypothesis	20	2	6	28 (13)
Multiple hypothesis	44	7	10	61 (28)
	164	22	29	215 (100)

* Papers were coded by a research assistant in terms of primary thrust. In addition, 128 of the studies were also coded by the author. For 11% of the articles, interpretations were different; these differences were resolved by discussion.

** Mere illustrations with data were not considered to be empirical studies. Also, in order for an article to be classified as having an hypothesis, the author must have presented the hypothesis as being made prior to doing the analysis.

*** For multiple authors, papers were coded as non-profit (or profit) if at least one author was in this category. No overlap occurred between nonprofit and profit authorship.

Source: Papers published in the *Journal of Forecasting* (1982-85, Issue 1) and in the *International Journal of Forecasting* (1985-1987).

It is apparently difficult for research journals to publish papers from those outside academia: for example, Heck and Bremser (1986) report that less than 3% of the papers published in *The Accounting Review* from 1966-1985 were from non-academic authors. They also reported that, over time, the trend has been toward fewer papers by those in business and government. The *Journal of the Operational Research Quarterly* has been more successful in gaining input from those outside academia. During the 1982-1987 period, 13.2% of the authors of papers were non-academic (Hough, 1987).[5] Relatively speaking then, the *IJF* and *JoF* have significant inputs from non-academic researchers; almost one-fourth of the papers had at least one author from outside academia.

Standards for research on forecasting

This section provides a set of standards for scientific research and then examines how research in forecasting might better meet these standards. The standards relate to:

?? objectivity
?? replicability
?? importance
?? competency
?? generalizability
?? intelligibility, and
?? efficiency.

The rationale for the selection of these criteria is provided by Armstrong (1982), with the exception that generalizability is stressed in this current list. As part of this examination, I considered what steps can be taken by journals to promote publication of research that does well on these criteria.

Objectivity: The selection of an hypothesis often leads to bias as the researcher can become an advocate of this hypothesis. The above-mentioned trend towards multiple hypotheses helps to promote objectivity. Another

[5] The JORL statistics are not directly comparable of our forecasting figures as the percentages were based on all authors, whereas we were interested in the papers that had at least some outside influence. Also, the JORL case studies were not included; these tend to attract a higher participation by non-academics.

approach is that, rather than stating the issue in terms of which hypothesis is correct (or which method is best), it seems preferable to indicate under what conditions a given hypothesis is relevant (Greenwald et al., 1986). For example, we might test whether bootstrapping is inferior to judgment in cases where the data on the causal variables are poor and the rules used by the judges are based on irrelevant variables. Gardner and McKenzie (1985) provide a good example of such work by defining the conditions under which a damped trend will provide more accurate extrapolation forecasts than those provided by an undamped trend.

Replicability: This activity is vital to academic progress. It serves as a check on spurious results, puffery, mistakes, and cheating. Interestingly, for reasons not totally understood, a substantial percentage of published results cannot be replicated by other researchers. Reid, Soley, and Wimmer (1981) found that 40% of the 30 published replications they examined in advertising research conflicted with the original results. While these studies have not dealt with studies on forecasting, I suspect that the situation may be similar in forecasting.

Replication can be hampered by the reluctance of researchers to provide full disclosure. Researchers may not want others to use the data that they have collected. Recent evidence, however, is encouraging. Eaton (1984) found that most psychologists provided the necessary data for replication. Journals should take an active role in insuring that the necessary data are disclosed. For example, such procedures by the *Journal of Money, Credit and Banking* allowed Kramer et al., 1985, to replicate 11 of 12 studies. Kramer's study also illustrated the importance of replication in its conclusion that over half of the papers that used regression contained seriously flawed models. Journals should give preference to replication studies.

Those who have published empirical studies in forecasting have a tradition of making the data available to other researchers (e.g., Dalrymple and King, 1981; Makridakis et al., 1982; and Schnaars, 1986). This aids replication. In fact, there have already been a number of replications of the Makridakis study.

Importance: Importance is seldom used as an explicit criterion for deciding what should be published (Armstrong, 1982). Grassmann (1986) concluded that more emphasis should be placed on importance of research in the management sciences.

Journals can ask referees to rate the importance of the papers submitted. Certainly referees will differ on what constitutes importance. The *IJF* asks its referees to rate explicitly the importance of papers to practitioners as well as to other researchers. The *IJF* also asks referees to rate how surprising the findings are on a scale from 1 ("not at all surprising") to 5 ("Very surprising"). Presumably, surprising findings are more important, all other things being equal. Research on forecasting has produced many surprising results. For example, Fildes and Lusk (1984) show how the results from the M-Competition were surprising in light of prior opinions of many experts in the field.

Competency: Obviously, researchers should try to avoid errors. Gardner's (1984) study of errors in exponential smoothing models suggests that errors may not be unusual. He found 23 errors in journal articles and textbooks.

Refereeing helps to prevent errors. Nevertheless, errors will occur and journals should be active in reporting corrections. One should not conclude, for example, that journals are error-free even though they do not report corrections. Unfortunately, neither the *JoF* nor the *IJF* have been reporting corrections.

Generalizability: Does the research allow one to generalize? Inasmuch as one of the main concerns is to generalize to forecasting practices in organizations, it is unfortunate that so little field research has been done to date. Instead, reliance is placed on theory and on empirical analyses outside the organizational context. Some recent work shows promise. Murphy and Daan (1984) showed how improved feedback procedure improved the subjective forecasts made by four weather forecasters. Bretschneider and Gorr (1987) statistically analyzed different procedures used by state governments and found evidence on a number of methodological and organizational issues; for example, government organizations that used simple econometric methods had more accurate forecasts then those using complex econometric methods. Hogan (1987) also obtained useful results in this study using students groups to forecast for a medical center. Field research can improve the ability to generalize.

Given the growth of research in the field, there is a need for systematic ways to integrate information. Traditional unstructured reviews are subject to bias and they may lead to unfounded or at least uncertain generalizations. The procedures of meta-analysis are relevant here.

In meta-analysis, the unit of observation is the *finding from a study*. The key steps in meta-analysis are as follows. First, search for studies using an extensive, systematic search with all steps documented; second, screen the studies using clearly-defined pre-specified criteria to determine whether a study should be included; and, third, conduct a systematic coding and statistical analysis of the studies.

Meta-analysis lends itself well to hypothesis-testing. For example, Armstrong (1984) used meta-analysis to test whether complex extrapolation methods were more accurate than moderately complex methods by analyzing results from 39 studies. Previous unstructured reviews had yielded conflicting generalizations. The meta-analysis showed that complexity did not provide more accurate forecasts.

One procedure for improving the results of meta-analysis is to request authors of previous studies to identify additional studies and to confirm that the coding of their studies has been properly done.

The coding is not always a trivial step. For example, the M-competition is a study where the authors themselves disagreed about the findings (see the commentary in Armstrong and Lusk, 1983). This is an extreme case. Often, however, it is far from simple for a reviewer to draw conclusions from prior research studies, especially if that researcher is an advocate of one method.

Intelligibility: It is not only the doing of the research that is important, but also the way in which it is communicated. Journals should consider the clarity of the papers. To aid in this, the IJF had a goal of publishing papers that did not exceed a Gunning Fog Index of 16. (This is equivalent to material that could be understood by someone with an educational level of grade 16, or the completion of college). We calculate the fog index by selecting samples from each paper published. During the first year, 1982, the fog index was about 17. From 1985 through 1987, the average fog index was less than 16. The Gunning Fog Index provides a crude measure of readability. As a further check, the 1986 survey of IIF members asked readers of the *International Journal of Forecasting* and of the *Journal of Forecasting* to rate the papers that they read from 1 ="poorly written" to 5 = “extremely well-written.” The 190 respondents to this item (from the 1986 survey) gave an average rating of 3.7. In other words, they felt the papers were relatively well-written. The average rating by practitioners was 3.6, which is almost as high as the 3.8 rating by academics. This indicates some success with the goal of bridging the gap between theory and practice inasmuch as the journal is communicating with both audiences.

Efficiency: The current reward system in academia encourages authors to divide a study into small pieces. This is inefficient for journal space, referees, and for readers. To address this problem, the IJF encourages submission of complete studies. No page length is imposed. Authors are asked to list similar papers that they have published. Nevertheless, it is not unusual to receive submissions that are similar to papers already published. I will not cite examples here.

Summary

This paper reported on two surveys of forecasters and analyses of research literature. The latter included counts of publications over time, a coding of the type of research published, and a review of the literature.

Research on forecasting since 1960 has been moving toward an empirical basis - using multiple hypotheses - across a variety of disciplines - with relatively clear reporting of results. The effects of this research to date have been impressive. Some of the more important advances have been:

1. econometric methods provide more accurate long-range forecasts than do judgmental or extrapolation methods,
2. structured methods significantly improve the accuracy of judgmental forecasts,
3. expert systems (bootstrapping) provide slight improvements in accuracy and a lower cost for repetitive forecasts that are now being made judgmentally,

4. expertise in the subject area is of little value for long-range judgmental forecasts of change, and
5. high complexity in quantitative models does not yield greater accuracy.

Findings # 1 and # 2 are reassuring. Findings # 3, # 4, and # 5 are surprising.

Despite the numerous advances, a comparison of publications against interests expressed in surveys of experts suggests that we are not producing research in areas where practitioners and researchers think it is needed. Some key omissions are in the areas of implementation, computer methods, evaluation, and uncertainty. In contrast, most papers are on extrapolation methods, where practitioner interest is relatively low. On the other hand, applications rated high for both practitioners and academics, and about half of the papers did address applications.

The surveys of forecasters and the reviews of the literature indicated numerous important areas that are largely ignored. Among these are:

1. What standards can be used to define acceptable practice in forecasting?
2. How can we better predict the outcomes of conflict situations?
3. What methods are best for estimating the uncertainty associated with a forecast?
4. How can we more effectively gain managerial acceptance of unfavourable forecasts (e.g., can scenarios help)?

These represent only a few of the many topic areas that need to be studied. Where should the resources come from? Perhaps we are overinvested in research on extrapolation methods.

How should this research be conducted? To date, one of the strengths of our field is that we produce many empirical studies employing the method of multiple hypotheses. We should build on this strength. I also recommend . . . selecting topics for their importance to researchers and practitioners . . . testing the conditions under which a method is most useful . . . emphasizing field research . . . replicating studies . . . and using meta-analysis when a sufficient amount of prior research has been published.

References

Armstrong, J. Scott (1979), "Advocacy and objectivity in science," *Management Science*, 25, 423-428.

Armstrong, J. Scott (1982), "Research on scientific journals: Implications for editors and authors," *Journal of Forecasting*, 1, 83-104.

Armstrong, J. Scott (1984), "Forecasting by extrapolation: Conclusions from twenty-five years of research," *Interfaces*, 14, Nov.-Dec. 52-66.

Armstrong, J. Scott (1985), *Long-range Forecasting: From Crystal Ball to Computer*. Wiley, New York

Armstrong, J. Scott (1986), "Research on forecasting: A quarter-century review, 1960-1984," *Interfaces* ,16, 89-109.

Armstrong, J. Scott (1987), "Forecasting methods for conflict situations," in George Wright and Peter Ayton, eds., *Judgmental Forecasting*. Wiley, Chichester, 157-176.

Armstrong, J. Scott and E. Lusk (1983), "Research on the accuracy of alternative extrapolation models: Analysis of a forecasting competition through open peer review," *Journal of Forecasting*, 2, 259-311.

Bretschneider, Stuart 1. and W. L. Gorr (1987), "State and local government revenue forecasting," in Spyros Makridakis and Steven C. Wheelwright, eds., *The Handbook of Forecasting: A Manager's Guide*, 2nd ed. Wiley, New York, 118-134.

Brouthers, Lance Eliot (1986), "Parties, ideology and elections: The politics of federal revenues and expenditures," *International Journal of Public Administration*, 8, 289-314.

Cerullo, Michael J. and A. Avila (1975), "Sales forecasting practices: A survey," *Managerial Planning*, 24, 33-79.

Christensen-Szalanski, Jay J .J. (1986), "Improving the practical utility of judgment research," in B. Brehmer, J. Jungerman, P. Lourens and G. Seven, eds., *New Directions in Research on Decision Making*. North-Holland, Amsterdam, 383-410.

Dalrymple, Douglas and B. E. King (1981), "Selecting parameters for short-term forecasting techniques," *Decision Sciences*, 12, 661-669.

Drucker, Peter F. (1973), *Management*. Harper and Row, New York.

Eaton, Warren Q. (1984), "On obtaining data for research integrations," *American Psychologist*, 39, 1325-1326.

Fildes, Robert and E. Lusk (1984), "The choice of a forecasting model," *Omega*, 12, 427-435.

Fischhoff, Baruch (1976), The effect of temporal setting on likelihood estimates," *Organizational Behavior and Human Performance*, 15, 180-194.

Fischhoff, Baruch (1988), "Judgmental aspects of forecasting: Needs and possible trends," *International Journal of Forecasting*, 4, no. 3, 331-339.

Gardner, Everett S., Jr. (1984), "The strange case of the lagging forecasts," *Interfaces*,14, 47-50.

Gardner, Everett S., Jr. and E. McKenzie (1985), "Forecasting trends in time series," *Management Science*, 31, 1237-1246.

Grassmann, W. K. (1986), "Is the fact that the emperor wears no clothes a subject worthy of publication?" *Interfaces*, 16, 43-51.

Greenwald, Anthony G., C. G. Carnot, R. Beach and B. Young (1987), "Increasing voting behavior by asking people if they expect to vote," *Journal of Applied Psychology*, 72, 315-318.

Greenwald, Anthony G., A. R. Pratkanis, M. R. Leippe and M. H. Baumgartner (1986), "Under what conditions does theory obstruct research progress?" *Psychological Review*, 93, 216-229.

Griffith, John R. and B. T. Wellman (1979), "Forecasting bed needs and recommending facilities plans for community hospitals: A review of past performance," *Medical Care*, 17, 293-303.

Hanke, John (1989), "Forecasting in business schools: A follow-up survey," *International Journal of Forecasting*, 5, 259-262.

Heck, J. Louis and W. B. Bremser (1986), "Six decades of *The Accounting Review*: A summary of author and institutional contributions," *The Accounting Review*, 61, 735-744.

Hogan, Andrew J. (1987), "Combining forecasts: Some managerial experiences with extrapolation," *Socio-economic Planning Sciences*, 3, 205-211.

Hough, John (1987), "Editorial," *Journal of the Operational Research Society*, 32, 1107-1108.

Kahneman, Daniel, P. Slovic and A. Tversky (1982), J*udgmentUunder Uncertainty: Heuristics and Biases*. Cambridge University Press, Cambridge.

Kramer, Walter, H. Sonnberger, J. Maurer and P. Havlik (1985), "Diagnostic checking in practice," *Review of Economics and Statistics*, 67, 118-123.

Makridakis, Spyros, et al. (1982), "The accuracy of extrapolation (time series) methods: Results of a forecasting competition," *Journal of Forecasting*, 1, 111-153.

Makridakis, Spyros, M. Hibon, E. Lusk and M. Belhadja (1987), "Confidence intervals: An empirical investigation of the series in the M-Competition," *International Journal of Forecasting*, 3, 489-508.

Murphy, Allen H and H. Daan (1984), "Impacts of feedback and experience on the quality of subjective probability forecasts: Comparison of results from the first and second years of the Zierikzee experiment," *Monthly Weather Review*, 112, 413-428.

Reid, Leonard N., L. C. Soley and R. D. Wimmer (1981), "Replication in advertising research: 1977, 1978, 1979," *Journal of Advertising Research*, 10, 3-13.

Schnaars, Steven P. (1986), "An evaluation of rules for selecting an extrapolation model on yearly sales forecasts," *Interfaces*, 16, 100-107.

Shamir, Jacob (1986), "Preelection polls in Israel: Structural constraints on accuracy," *Public Opinion Quarterly*, 50, 62-75.

Sherman, Steven J. (1980), "On the self-erasing nature of errors of prediction," *Journal of Personality and Social Psychology*, 39, 211-221.

Stephan, Walter G. (1978), "School desegregation: An evaluation of predictions made in Brown v. Board of Education," *Psychological Bulletin*, 85, 217-238.

Stewart, Thomas R. and M. H. Glantz (1985), "Expert judgment and climate forecasting: A methodological critique of `Climate change to the year 2000'," *Climatic Change*, 7, 159-183.